

DEVELOPMENT OF A DISTRIBUTED MACHINE LEARNING  
PLATFORM WITH FEATURE AUGMENTED ATTRIBUTES  
FOR POWER SYSTEM SERVICE RESTORATION

Miftah Al Karim

Supervisor: Professor Tek Tjing Lie

Co-Supervisor: Dr. Jonathan Currie

June 2018

## **Abstract**

Modern power systems are gradually adopting the philosophy of autonomous and distributed means of dynamic event detection processes, by facilitating system operation using intelligent algorithms. Dynamic data has different forms, such as voltage, current, active-reactive power, frequency, rotor speed and angle, status of a breaker, electricity demand, electricity price, operation schedules and many more. Collecting and storing these dynamic data from a system has recently become much more feasible for engineers. Along side that, computational resources have also been increased in multitude. Implementing these resources can bring practical benefits for post-event operation under uncertain conditions. The market drive towards decentralization and deregulation is constantly introducing problems which are stochastic. Thus, techniques like machine learning are observing a fair share of its applications in the field of power engineering. Though in recent years, application of machine learning algorithms in the power industry has become a prominent domain, understanding a dynamic data set requires additional features. Developing such features is yet an interesting research genre where significant contributions can be made. The concept of feature analysis using power system dynamic data is relatively novel.

This research, in multiple steps, applies different machine learning methods to understand the dynamic nature of power system data. The idea is to search for effective methods to analyse this power system data. The dynamic nature of any power system demands efficient training and testing facilities in a pattern recognition system. Thus, recognizing the impacts of different dynamic techniques on a continuous flow of power system data would be a priority in this work and it requires a prelude to realizing the capacity of data analytics to understand power system dynamics.

However, investigating the dynamic capabilities of the machine learning techniques is only the beginning. The power system is a complex network that has lots of segments and interconnections,

with predefined priorities assigned. Multiple pattern recognition techniques may, therefore, be required to effectively understand the interactions between events that govern system operation and control. Rather than analysing the impact of a method on all types of scenarios, preparing different effective methods for different events would be a logical approach. The question arises whether a centralized scheme can be of any use while understanding events that are occurring in different autonomous zones. For example, the dynamic behavior of a system near a wind farm, reacting to different wind power penetrations, could effectively be understood with a decentralized system that is already installed in a nearby substation. Under such a scenario, collecting all the system data and storing it to the central station for further analysis should be computationally expensive, considering that some events are taking place at that very moment. It is another motivating factor for this work to investigate whether distributed data analysis with dynamic data handling capacity can lead towards better prediction, and consequently, the operation of a system.

However, power system dynamics and machine learning algorithm-based data analysis are two independently broad domains. Analysing the impacts of data analytics on power system dynamics, would therefore be another broad discussion, ranging from a generic towards a specific study. In this thesis, machine learning platforms have been considered as the intelligent means of data analysis and the self-healing framework of a microgrid has been used for classification of the dynamic power system events. The key contribution of this research is developing a feature-extraction-based algorithm that has the capacity of detecting power system events and facilitating decision making.

# Contents

<b>1</b>	<b>Introduction</b>	<b>11</b>
1.1	Research Background . . . . .	11
1.2	Research Objectives . . . . .	15
1.3	Research Contributions . . . . .	16
1.4	Chapter Introductions . . . . .	17
<b>2</b>	<b>Literature Review</b>	<b>19</b>
2.0.1	Power System Data . . . . .	19
2.0.2	Applications of Machine Learning in Power System . . . . .	22
2.0.3	Existing Methods for Dynamic Data Analysis . . . . .	37
<b>3</b>	<b>Stochastic Data Model</b>	<b>53</b>
3.1	Preparation of the Data Models . . . . .	53
3.2	Machine Learning for Power System Stability: Traditional Approach . . . . .	59
3.3	Understanding the Limitations of Traditional Approaches . . . . .	64
<b>4</b>	<b>Machine Learning for Supervised Hierarchical Platform</b>	<b>69</b>
4.1	Post Fault Restoration . . . . .	69
4.1.1	Microgrid Model . . . . .	72
4.2	Proposed Secondary Control . . . . .	72

4.2.1	Data Preparation . . . . .	75
4.2.2	Preparing the Distributed Algorithm . . . . .	77
4.2.3	Results and Discussions . . . . .	78
4.2.4	Analyzing the Limitations . . . . .	84
<b>5</b>	<b>Simplified Feature Selection for Dynamic Data</b>	<b>86</b>
5.1	Addressing the Limitations with Feature Selection . . . . .	86
5.2	Proposed Methodology . . . . .	89
5.2.1	Data Preparation . . . . .	91
5.2.2	System Dynamic Responses . . . . .	93
5.2.3	Selected Features . . . . .	98
5.2.4	Multiclass Classifier . . . . .	104
5.2.5	Testing the Overall Hypothesis . . . . .	109
5.3	Performance of the Feature Augmentation Based System . . . . .	112
5.4	Understanding the Limitations . . . . .	116
<b>6</b>	<b>Data Augmentation for a Segmented Grid</b>	<b>118</b>
6.1	Features in Larger Networks . . . . .	118
6.2	System Under Consideration . . . . .	121
6.3	Proposed Secondary Control . . . . .	124
6.4	Result Analysis . . . . .	128
6.5	Analyzing Limitations . . . . .	132
<b>7</b>	<b>Selection of Appropriate Feature Space</b>	<b>133</b>
7.1	Different Types of Features . . . . .	133
7.2	Machine Data and Additional Features . . . . .	135
7.3	Corrective Voltage Control Data . . . . .	137

7.4	The Machine Learning Platform . . . . .	137
7.4.1	Features for Pre-processing . . . . .	138
7.4.2	Multiclass Classifier . . . . .	141
7.5	Results . . . . .	141
7.6	Overall Performance Evaluation . . . . .	144
<b>8</b>	<b>Ensemble Method for Multi-objective Application</b>	<b>146</b>
8.1	Decision Making With Multiple Constraints . . . . .	146
8.2	The Microgrid Model . . . . .	149
8.3	Data Model . . . . .	152
8.3.1	Wind Energy Model . . . . .	152
8.3.2	Solar Energy Model (Photovoltaic) . . . . .	154
8.3.3	Electrical Load Model . . . . .	155
8.3.4	Security Index Probability of Stability . . . . .	156
8.3.5	Data for Fault Analysis . . . . .	157
8.4	Classification Process for Optimal Voltage . . . . .	158
8.4.1	Preparing the Classification Tree . . . . .	161
8.4.2	Process for Economic Dispatch . . . . .	163
8.5	Results . . . . .	166
8.5.1	Estimation of Instability . . . . .	166
8.5.2	Selection of Restorative Action . . . . .	168
8.6	Analyzing the Limitations . . . . .	169
<b>9</b>	<b>Conclusion and Future Work</b>	<b>172</b>
9.1	Concluding Remarks . . . . .	172
9.2	Future Scopes . . . . .	175

# List of Figures

2.1	Proposed forecasting model for smart grid [1]	26
2.2	Proposed second approach [2]	31
2.3	DW and DM Based Decision Support System	33
2.4	Structure of RNN [3,4]	44
2.5	Five-layer Two-input Four-Rule fuzzy model (Sugeno) [1]	48
2.6	Fuzzy Reasoning [5]	49
3.1	Normal Probability Distribution of Temperature (Degree Celsius)	55
3.2	Random temperature values (Degree Celsius)	55
3.3	Wind Power Plant	56
3.4	Solar Plant	57
3.5	Calculated Electrical Load	58
3.6	Two Area AGC Based System	59
3.7	Hydro Turbine Governor	60
3.8	Diesel Turbine Governor	60
3.9	MSE of ANN training for hydro power plant	61
3.10	MSE of ANN training for diesel power plant	62
3.11	NARX network	62
3.12	Wind speed prediction [6]	63

3.13	Solar radiance prediction [6] . . . . .	63
3.14	Predicted load 24 hour ahead (Fitting Model) . . . . .	64
3.15	Predicted load 24 hour ahead (NARX Model) . . . . .	64
3.16	Prediction of hydro turbine governor (Transient Part) . . . . .	65
3.17	Prediction of hydro turbine governor (Steady Part) . . . . .	65
3.18	Prediction of diesel turbine governor . . . . .	66
3.19	Prediction of hydro turbine governor (with sub networks) . . . . .	67
3.20	Prediction of diesel turbine governor (with sub networks) . . . . .	68
4.1	The Microgrid Model a Modified Approach from [7] . . . . .	73
4.2	Secondary Control Scheme . . . . .	75
4.3	Operating Near the Voltage Instability Point (Without Compensation) [8] . . . . .	76
4.4	Fluctuation During Voltage Instability . . . . .	77
4.5	K-means Cluster based Neural Networks [9] . . . . .	79
4.6	Work-Flow of the Proposed Algorithm . . . . .	80
4.7	Binary Classification of the System Stability (Units in watts) . . . . .	81
4.8	Performance of the Proposed Secondary Control Scheme (Cluster-4) . . . . .	83
4.9	Performance of the Proposed Secondary Control Scheme (Cluster-3) . . . . .	84
5.1	Workflow of The Proposed Algorithm . . . . .	90
5.2	The Microgrid Model . . . . .	91
5.3	Generator Data Under Different Dynamic Events (50 samples/second) . . . . .	96
5.4	Generator Data Under Different Dynamic Events (50 samples/second) . . . . .	97
5.5	Peak Prominence and Width Inside a Predefined Data Window of the Terminal Voltage of Generator-1 . . . . .	101
5.6	(a) Frequency Factor at Different Window Length and (b) Transition Between Two Events . . . . .	102

5.7	Sample Data and Feature Selection . . . . .	103
5.8	Similarity among multiple events . . . . .	105
5.9	Classification Error Vs Number of Trees . . . . .	107
5.10	Preparation of the Multi-class Ensemble Trees [10, 11] . . . . .	108
5.11	Classification With and Without Features (fictitious Data) . . . . .	110
5.12	Classification Without and With Features . . . . .	111
5.13	Accuracy Chart After 400 Simulations . . . . .	114
5.14	Decision making with and without the proposed algorithm . . . . .	115
5.15	Post fault operation of Generator-1 with the Nearby SC in Bypass Mode . . . . .	116
6.1	High Fluctuation in Rotor Angle due to a Critical Fault . . . . .	122
6.2	The Microgrid Model . . . . .	123
6.3	Workflow of The Proposed Method . . . . .	125
6.4	Preparation of the <b>Nine</b> Stochastic Clusters . . . . .	126
6.5	Sensitivity $\partial P_m / \partial V_{fbus}$ Curve of Fault-11 . . . . .	126
6.6	Five-layer Three-input Four-Rule fuzzy model (Sugeno) [1] . . . . .	129
6.7	ANFIS Based Prediction for Generator-7 . . . . .	130
6.8	Stabilizing the Segmented Microgrid with the Proposed Algorithm . . . . .	131
6.9	Generator-7 Data . . . . .	131
6.10	Comparison of Voltages at the Fault Bus With Different Methods . . . . .	132
7.1	The Microgrid Model . . . . .	136
7.2	Time Series Data of Generator-7 Under Different Scenarios . . . . .	137
7.3	Peak prominence and width inside a predefined data window of the Terminal voltage of Generator-1 . . . . .	138
7.4	Available Frequencies . . . . .	139
7.5	Available Frequencies . . . . .	140

7.6	Sensitivity $\partial P_m / \partial V_{fbus}$ Curve . . . . .	140
7.7	A Part of the Classification Tree . . . . .	142
7.8	Events . . . . .	142
7.9	Prediction Accuracy with the Test Data . . . . .	143
7.10	Ramification of the Classification Error . . . . .	144
8.1	The Microgrid Model . . . . .	150
8.2	Stable and Unstable Post Fault Responses . . . . .	152
8.3	Implementing the Data Model into the Proposed microgrid . . . . .	153
8.4	Normal probability distribution of temperature (Degree Celsius) . . . . .	154
8.5	Influence of the Ambient Temperature on the Controllable Load . . . . .	155
8.6	Controllable Load Model . . . . .	156
8.7	Kernel Density Estimation . . . . .	157
8.8	The Decision Tree Classifier (Limited Scenarios) . . . . .	161
8.9	Hierarchical Cluster . . . . .	162
8.10	Work-Flow of the Proposed Algorithm . . . . .	166
8.11	POS of the <b>100</b> Clusters and the Monte Carlo method of convergence . . . . .	167
8.12	POS of the Clusters . . . . .	167
8.13	Top Three Decisions after 10k simulations . . . . .	168
8.14	Results With and Without the Proposed Method . . . . .	170
8.15	Different Decision Making Processes and Impacts (The Remaining Clusters) . . . .	171

# List of Tables

4.1	Data Table With Clusters (Units in watts)	78
4.2	Centroids of the Clusters	79
4.3	Predictions Using Neural Networks	82
5.1	Features In Three Similar Events	106
5.2	Accuracy for <i>Event1</i> , <i>Event3</i> and <i>Event9</i> in Percent %	113
5.3	Online Decision Making based on the feature data collected from different locations (Observed from different generators)	115
6.1	Stochastic Data Preparation Stage for Generator-7&9 (Selected Scenarios)	129
7.1	Decisions	143
8.1	Data Table With Clusters Centroids	157
8.2	Sixteen Probable Restoration Schedules	159
8.3	Data Table With Clusters Centroids	160
8.4	The Hierarchical Ranking Table	162
8.5	Three Different Actions	168
8.6	Normalized Operational Costs	169

# Declaration

I hereby declare that the content of the following work is original, except where specific reference is made to the work of others. I also declare that this work has not been submitted in whole or in part for consideration for any other degree or qualification, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others.

.....

Miftah Al Karim

July, 2018

# Acknowledgments

I am highly grateful to almighty Allah, the creator and sustainer. His infinite blessing and mercy made me successfully go through this arduous journey. I am thankful to my supervisor Professor Tek-Tjing Lie and co-supervisor Dr. Jonathan Currie. Their guidance and support showed me the right path. Finally, I would like to express my gratitude to my parents and lovely wife. Their supports and sacrifices helped me get through many difficult challenges.

# Chapter 1

## Introduction

In this chapter, the motivation behind this thesis, has been discussed and presented as research background. The research objectives are introduced, alongside which the contributions are also brought out.

### 1.1 Research Background

A data-driven self-healing structure is just the crux of the so-called system of systems smart grid. To achieve a state where a power system can be defined as a smart grid, it must undergo several configuration changes. These changes may be represented as, online monitoring and reactive capability, appropriate forecasting capability and capacity of time to time rapid isolation [12]. These attributes can bring an end-to-end resiliency that not only reduces outages but also achieves system-wide operational objectives.

Ideally, the control strategies in a self-healing smart grid can be divided into four categories: emergency, restorative, corrective and preventive [12]. Therefore, a self-healing structure is introduced to reduce operational cost due to service interruption, and optimality can also be sought out inside the boundary of a self-healing architecture [13]. Such optimization problems often lead

the analytical approach from a centralized towards a decentralized one, which sometimes is referred to as a distributed control. The distributed architecture mostly contributes to the field level device such as synchronous machines. However, controlling synchronous machines to perform self-healing functionalities requires additional decision-making tasks. Earlier research focused on introducing intelligent schemes such as machine learning-based systems, to make such decisions, as a means of rapid-response, to face the external stimuli. These decisions at a higher level, are more sophisticated and thus, require centralized coordination [14]. It is not an overstatement to say that, the notion of a self-healing smart grid encompasses low-level system control influenced by high-level decision making. Such a hierarchy and multidisciplinary structure make self-healing a very complex process. Therefore, a timely step would be to discuss different approaches to introduce intelligence in the self-healing functionality. Studies have already discussed the concept of an intelligent restoration decision support system (RDSS) [15]. However, at a major scale, most of those schemes are on an offline basis, which is quite impractical from an economic standpoint, especially, when the integration of renewable energy sources has already introduced higher degrees of uncertainty. Besides, any offline restorative method is time-consuming. Another critical limitation observed in those methods is that the service restoration in a self-healing grid is often not considered as a multi-objective function. Therefore, researchers are addressing diverse issues that can eventually be solved through a single decision support system [16]. One solution to such approach can be an offline-online self-healing platform. Where the offline controller deals with the situational parameters and optimizes those, while the online controller regulates those parameters in a near real-time to real-time basis based on the proposed optimization [17].

The self-healing process requires balance, and balancing the active and reactive power in a stand-alone microgrid is a critical task. A microgrid without energy storage capability is even more vulnerable to energy imbalance and stability issues. Realizing self-healing is not possible, without understanding the core concept of energy balance, through a system parametric point of

view [14, 18]. On top of that, due to the nature of a broader aspect, energy balance itself is a large domain. Individually, all the techniques associated with energy balance can contribute to self-healing. Therefore, it can influence different strategic decision-making processes, based on the technique implemented. It is preferred to investigate one solution at a time to avoid this conundrum. One of those solutions is a distributed supervised secondary control, to maintain the rated voltage in a stand-alone microgrid. This control scheme is hierarchical. The supervision can be a machine learning-based system. As the machine learning (ML) algorithms can be implemented both with supervised and unsupervised data sets, applying ML in a hierarchical decision-making structure is more relevant, as such hierarchy can incorporate external information, like weather data, in decision making. Another motivation for ML is that, supervised control schemes are mostly applied where primary and secondary control schemes are insufficient to maintain system stability after a sudden and major change in load or fault. Once a severe contingency occurs in a microgrid, it is imperative for the grid to monitor its status, act based on the level of severity, and restore the system once the contingency has been cleared. Based on the above philosophy, scenarios in which self-healing was mostly proposed were model specific but generic regarding labels. Therefore, the overall decision making can be a stochastic-data driven intelligent action. Under this context, machine learning is a superior candidate for data analysis [19].

One key trend is to deploy such hierarchical control in a centralized architecture. With an increasing number of microgrids and distributed generation stations, deploying centralized control is costly. In an interconnected network, it is important to detect the underlying events taking place in each of the distributed stations; otherwise centralized decisions become non-coherent. In power systems, event detection is a problem, where data analysis can play a vital role [7, 20]. Ideally, the load-flow data machine parameters are used to detect the underlying events. But, modern system analysis understands that other ambient factors such as weather, interconnectivity and costs should also be considered. Such an approach demands further incorporation of dynamic and stochastic

data analysis. Once the database becomes large, centralized control schemes no longer remain the most efficient methods. Therefore, the distributed operation becomes more relevant. Usually, the performance of any distributed and supervised control is monitored by the substations.

The key motivations for developing a distributed control process are cost-effectiveness and speed of operation, especially in islanding modes. The complex machine learning process introduced in the earlier work is offline, time intensive and not quite suitable for a near real-time application. But, achieving a near real-time basis decision support system is also resource intensive. Therefore, an offline-online structure has more flexibility. In past years, a couple of offline-online methods of post-fault service restorations have been tested on different types of microgrids, that have distributed and also on renewable energy generation schemes [21,22]. Microgrids are smaller, thus, quite suitable to be used as test beds. A microgrid is usually operated at its design limits. Such an operational method is even more prone to post-fault instability, even though the fault that causes instability is cleared within a short duration. With intelligent decision-making tools, such instabilities can be addressed if those tools can analyse and classify stochastic dynamic events. However, the dynamic data obtained from any power system is time series in nature. Thus, recognizing patterns from the individual scenario is complex. So, a keen focus is required to understand, what are the impacts of different types of features on a decision-making tool.

Finally, it should be addressed that, a self-healing process is a multifaceted framework. Different definitions can be found for a self-healing system. Therefore, a frame of reference can always help understand the point of view proposed in any study. This study uses a corrective voltage control (CVC) framework that maintains sufficient load margin, as a point of reference in the ‘N-1’ self-healing strategy. Standard CVC measures are based on an active and reactive dispatch from generating units [23]. However, in a post contingent scenario, it is often critical to select appropriate parameters. This study undergoes (undertakes) such observations and proposes a distributed machine learning platform to address the issues.

## 1.2 Research Objectives

The first research objective is to understand the limitations of a traditional machine learning approach for a self-healing microgrid. The assessment is based on different contingencies or stochastic parameters. Different machine learning algorithms are used in the process, which has also been previously used to solve meaningful power system problems. While analysing the traditional approaches, it was observed that the distributed generators, such as the wind and solar plants, as well as the residential loads, have some degree of randomness.

The second objective is to propose a modification of the traditional machine learning approach to address the self-healing phenomenon. The introduction of this proposed algorithm is small scale. It means, in this phase, the algorithm is only judged based on its problem-solving capabilities. This stage does not consider the cost of deploying the algorithm in a centralized or distributed platform, nor does it consider interoperability. However, one of the objectives of this initial algorithm development is also to find out the limitations of this algorithm, for larger networks.

The third objective of this study is to propose, a novel feature selection-based distributed machine learning algorithm, to detect the dynamic signatures of different power system events. The purpose is to facilitate a post-fault decision-making process to restore a stand-alone microgrid without the intervention of a central station. At this stage, a more complex decision architecture will be used.

The fourth objective would be to implement the proposed algorithm in a larger network and analyse its potential.

The fifth objective is to fine-tune the algorithm, by selecting appropriate features and by comparing those features to solve a self-healing problem.

The sixth or final objective of this research is to apply the proposed machine learning method in a multi-objective scenario, considering cost reduction as one of the vital goals.

The overall concept of developing the above mentioned, machine learning driven data analysis strategy is based on the following relevant and recent literature. The literature review is prepared to bear a theme of dynamic data analysis, keeping the proposed use cases in consideration. The use cases are explained chapter by chapter as the development process progresses.

### **1.3 Research Contributions**

This research is carried out to investigate dynamic data analysis for self-healing microgrids. The contribution observed in this work are as follows.

- Understanding dynamic events during a contingency, is crucial for a self-healing microgrid to act fast. This research develops a feature-selection based method, that can assist a machine learning platform to identify specific events in a microgrid.
- Distributed control often introduces loss of synchronism, while a system is restored. This issue of loss of synchronism has been addressed by the hierarchical machine learning platform. The hierarchy is developed by both unsupervised and supervised machine learning algorithms.
- Finally, a decision support system is developed based on an ensemble of machine learning algorithms, to perform a multi-objective function in a self-healing microgrid.

Over the course of this work, the following articles have been published.

- Dynamic Event Detection Using a Distributed Feature Selection based Machine Learning Approach in a Self-Healing Microgrid. IEEE Transactions on Power Systems, 2018.
- A machine learning based optimized energy dispatching scheme for restoring a hybrid microgrid. Electric Power Systems Research, Volume 155, February 2018, Pages 206-215.

- A distributed machine learning approach for the secondary voltage control of an Islanded microgrid. Innovative Smart Grid Technologies-Asia (ISGT-Asia), 2016 IEEE, Pages 611-616.
- A Feature-based Distributed Machine Learning for Post Fault Restoration of a Microgrid Under Different Stochastic Scenarios. Innovative Smart Grid Technologies-Asia (ISGT-Asia), 2017 IEEE.
- Distributed Machine Learning on Dynamic Power System Data Features to Improve Resiliency, 2018, IET Research Journals. (Submitted, waiting for review)

## 1.4 Chapter Introductions

In this section, an introduction to all the following chapters is provided. In the chapter **Literature Review**, studies relevant to dynamic data analysis is presented. How power system data is a candidate for dynamic data analysis has also been addressed. In the second chapter **Stochastic Data Model**, stochastic data generation is discussed. The following chapter **Machine Learning for Supervised Hierarchical Platform**, discusses the impact of stochastic parameters in a microgrid, and how machine learning algorithms can be trained using a simple triggering mechanism. In the next chapter, **Simplified Feature Selection for Dynamic Data**, different data preparation techniques for event-detection is introduced. This data preparation technique involves three feature selection mechanisms. In the following chapter **Data Augmentation for a Segmented Grid**, a larger power system with more complex events has been analysed. The underlying objective of this chapter is to develop a self-healing strategy. The next chapter **Selection of Appropriate Feature Space** further explores the idea of feature selection in a complex environment to fulfil an objective. This objective is the previously introduced self-healing mechanism. Chapter **Decision Making With Multiple Constraints** discusses ensemble methods for multi-objective functions and a relative analysis on

addressing misclassification problems for a self-healing network.

The purpose of such distribution of the chapters is to create a road-map, for demonstrating the evolution of concept in applied machine learning methods to solve self-healing functions in microgrids.

# Chapter 2

## Literature Review

This chapter discusses the background literature that motivated and supported the arguments presented in this work. Rather than making it an exhaustive search, a key focus has been put, on discussing the chronology of designing the research steps. This chronology includes the relevance of power system data, the importance of machine learning and the handling of dynamic data. The following sections will provide hints and strong case studies to advocate preparing machine learning algorithms to address dynamic events.

### 2.0.1 Power System Data

Deregulation and decentralization have substantial impacts on power system analysis. Modern methods are compelling engineers and operators to work with a vast amount of information. The growing needs for power consumption are also resulting in introduction of more data [24]. Wide area measurement systems (WAMS), therefore, have evolved and can accumulate the so-called big data. Processing such high-volume data requires additional support on top of traditional methods. In reference [25], Equation-2.1 is used to measure the scalability of WAMS data the following equation is used;

$$M_T = \left( \sum_{i=1}^N ((8p_i + 4a_i + 4s_i) + 20) + 12 \right) v_k T + \delta \times C \quad (2.1)$$

where  $M_T$  is the size of the storage space,  $N$  is the number of phasor measurement units PMUs,  $P$  is the number of phasors,  $v_k$  is the transmission rate of the dynamic data, and  $T$  is the specified storage time length,  $\delta$  is the additional space required to store dynamic data that can have a value of zero, and  $C$  is the split number of files required for the dynamic data. It shows the magnitude of data volume, which future operators will have to deal with.

A major contribution in data volume in the distribution nodes is observed due to the frequent implementation of advanced sensor-based systems, such as phasor measurement units, high-frequency voltage and current sensors, advanced metering infrastructure (AMI), etc. Research predicts that the installed base of smart meters would surpass the number one billion by the year 2022. The cost of data storage is going down, thus, the usage of data analytics is gaining popularity. The phasor measurement units (PMU) have the capability of providing high resolution and real-time data. Such capacity enables advent of new types of data and new means of analysis [26]. It is not without challenges — in power system analysis, data science can bring fruitful results [27, 28]. In the consequent chapters of this thesis, different power system models have been introduced. The data collected from those models are phasor type, and the data structure is synchronized. Therefore, this thesis is based on PMU data.

Traditional data analysis methods cannot be overthrown in a day or two. Reference [29] addresses the problems regarding the slow data processing of the SCADA system and fast data processing capacity of the PMUs. The authors there proposed a multi-rate multi-sensor data fusion. This challenge also provides a subtle hint as to why the centralized system control may not be as effective as it has been with traditional systems. Thus, data analysis gets a new perspective. In this thesis, a distributed control scheme has been adopted as the intended case study. The data analysis platform, therefore, is also proposed for use cases deploying distributed control schemes.

While considering a distributed control approach, system reliability, security and stability become the key aspects [30,31]. Thus, practical and modern operation schemes must be carefully utilized to address a scenario and take necessary response based on the effective information. Keeping that in mind, in Chapter-7 a comparison of different parameters regarding different events is presented.

Even if it is quite evident that power system data is a prime candidate for different data analytics, the overall concept of applying algorithms, to understand high volume data, would only make sense if data-driven operational schemes are practiced. The modern energy industry is not shying away from this concept. Scrutiny is required to get some idea on how the industry utilizes high volume data.

### **Data Driven Operation and Control**

Large volume, high velocity, and increasing variety are the characteristics of modern power systems, that are driving the industry towards efficient, condition-based asset management and real-time operation for outage prevention. Streaming synchrophasor measurement system (SMS), automated revenue metering (AMR) and intelligent electronic devices are compelling operators to use intelligent control and operation schemes for situational awareness, billing and condition-based management in a typical modern power system. Data outside of the technical realm such as weather data, lightning data, financial data are also having a fair share of contribution in system-operation. Though the existing system is coping with modern data-driven aspects of power systems, the decision-making process requires finesse. Understanding large data requires management techniques such as sensing, measuring, information exchanging, knowledge extracting and visualizing. Statistical methods also need to be implemented in order to derive the relevant data set for efficient decision-making that influences corrective, predictive, distributive and adaptive actions [32]. The on-line voltage stability study is important for interconnected systems like PJM, ERCOT etc.

Reference [33] summarizes the efforts of *CIGRE'* Task Force 38.02.20 for advanced power system control using intelligent systems. Online dynamic security assessment using a neural network to predict dynamic frequency response, intelligent systems for emergency control, control centre operation under deregulation are some of the focused topics of the task force. The key argument placed in reference [33] lies in figuring out supplements for the conventional control approaches. To achieve high-level adaptation and quick decision-making features, system-wide control schemes must be capable of working with imprecise information, that is generated in a modern yet uncertain and less structured environment. This issue can somewhat be addressed by implementing effective data pre-processing techniques. In Chapter-3 such a data pre-processing technique has been analysed, regarding complex event detection.

Such data-driven actions must be aided with computational-intelligence. One of the most efficient ways to retrieve information or knowledge, from either a data base or streaming on-line data, is to apply a pattern recognition tool, which is often a machine learning algorithm. Therefore, it has become a paradigm in the energy research community, to implement and analyse the performance of different machine learning methods. The following section introduces some of the interesting applications of machine learning in power systems.

## **2.0.2 Applications of Machine Learning in Power System**

### **Power System Steady State Data**

Availability of a large quantity of data opens up new avenues in power system analysis. As data is generated, stored and made available for future references, the context of using machine learning techniques gets more relevance. Machine learning techniques in general offer clustering, classification, regression and pattern finding to effectively manage data and extract useful information. Power system engineers are also often being benefited by applying pattern recognition algorithms such as neural networks, support vector machine, k-means clusters, classification trees and visu-

alization techniques [34, 35]. Numerous literature can be found where effective machine learning based methods have been introduced, analysed and compared for solving power system problems like security assessment, economic feasibility, on-line steady-state analysis, transient stability analysis, load profiling, system control, fault detection, state estimation and more [36]. In a broad sense, the application fields are generalized as problems in visualization, clustering, classification and regression. Data visualization is a key method to understand and express data during operation and control of a system. In reference [37] a geographic data view (GDV) based is proposed to facilitate rapid decision-making to solve problems such as line overload. The method displays a wide range of power system information with multiple attributes, along with the geographic location data. An interactive visualization technique that is based on a 3d virtual environment platform is proposed in reference [38]. The underlying objective of the work is to establish a bridge between multiple complex networks such as those in a power system, as well as the economics associated with the power system, using interactive visualization techniques. However, the objective of this research work is not to propose a new visualization technique, but rather to modify the existing pattern recognition techniques to facilitate dynamic power system operation and control. Thus, understanding the benefits of the visualization techniques is thought to be important, but not the key focus of this work. This work would consider the standard visualization methods to express, analyse and compare results.

In a general notion, data mining with the help of machine learning refers to a process where pattern recognition gets the key attention. Out of different pattern recognition tools, clustering is one of the most popular ones. A multitude of articles has addressed clustering techniques as a component to solve power system problems. In reference [39] a clustering technique is used to improve performances of the security assessment for daily operation and planning. Unsupervised clustering is used in reference [40] to solve dynamic security assessment problems. Reference [41] introduces a method where clustering and optimization based control design is used to address FACTS device

and PMU based, wide-area, oscillation damping issues. Some other interesting works that use clustering techniques are cable layout design for wind farms [42], load pocket identification [43] and load forecasting [44]. The clustering process represents unsupervised learning, which can be used as a means for data pre-processing. In this thesis other types of clustering techniques have also been used for data pre-processing.

The classification techniques can be considered as the supervised counter part of the unsupervised clustering method. The most common methods for classifying data are decision-tree based methods, support vector machine, k-nearest neighbour, etc. Classification techniques are often utilized in power system calculations, such as in secured economic dispatching [45], locating series compensators [46], classifying voltage disturbances [47], wind power forecasting [48] etc. Once it becomes obvious that, for an increasing data volume, classical expert opinion-based control and operation schemes are becoming obsolete every day, techniques like machine learning should get more attention. The machine learning in power system operations and control can be observed in two different ways. First one is to evaluate the performance of the machine learning routine by solving the individual problem. Here, based on historical data, an operation and or control scheme will be chosen, which will be compared with traditional approaches. This method does not require dealing with the dynamic characteristics of the system in real-time. Rather it can be a post-mortem analysis, where the only critical evaluation will be carried out by understanding the performance of the method or algorithm for a typical case. The second approach will be to consider the stream of dynamic data, analyse them in real-time and make on-line decisions for one or multiple events. This method requires not only to understand the dynamic nature of power system data, but also the dynamic nature of the machine learning methods proposed. It is important to compare the performance of these methods with the traditional machine learning schemes.

The following section introduces literature, that deals with different methods to do dynamic data analysis with machine learning algorithms in power systems.

## **Power System Dynamic Data**

Dynamic data analysis methods can be implemented in different fields. Power systems could be a potential candidate where dynamic data analysis can prove to be highly effective. Research shows that dynamic methods are frequently being used for power system data analysis. This section will discuss some of those works.

Dynamic data management is key to dynamic data analysis. So far, dynamic data analysis has frequently been observed in energy management systems. In reference [49] big data analytics has been introduced for dynamic energy management purposes for a Smart Grid (SG). SG allows consumers and micro energy producers to play an active role by enabling bidirectional power as well as the data flow. Dynamic energy management (DEM), therefore, depends on the instantaneous quantity of load and renewable energy generators. In a large network, data analysis becomes quite relevant, while considering DEM in play. It is because, DEM requires power flow optimization, system monitoring, planning and real-time operation. In short, DEM has a heavy dependency on complex multi-variate statistics. Such dependency demands advanced data analysis methods and machine learning to discover meaningful patterns from that data stream. But, machine learning for SG is a complex process of cascaded tasks, such as dimension reduction, classification, forecasting and distribution of resources to minimize processing time. To attain that goal, an effective near real-time monitoring and forecasting system is needed. In reference [49] a forecast model, that deploys pattern recognition in a parallelized computation platform, is proposed to solve this issue. The proposed method is shown in Figure-2.1. The key point observed here, is the addition of a feature extraction layer that can facilitate pattern recognition for solving the problems of forecasting.

One criticism of the systems mentioned above is the curse of dimensionality. Large-scale data analysis may not perform efficiently as the system grows bigger. Therefore, other data analysis methods are also sought out from time to time. One probable method to deal with the ever-growing electricity industry is to distribute resources. This distribution can also be implemented while con-

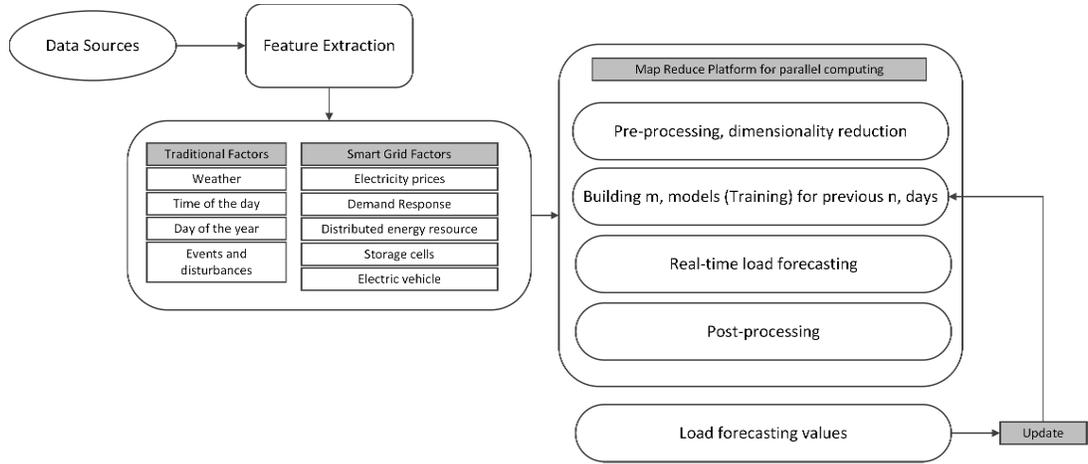


Figure 2.1: Proposed forecasting model for smart grid [1]

sidering machine learning based methods. In reference [50], the authors explored distributed approaches for analysing and processing data. The idea there is to increase flexibility in local control. The article assumes that, shortly along with the traditional energy companies, certain households would also supply energy. Multiple companies would then be interested to understand the distributed energy generation and supply patterns, and a household or a company may subsequently restrict data centralization. To deal with such phenomenon, the authors used multivariate linear regression and distributed rank aggregation techniques. The model is node based, yet topology independent that considers each node generates homogeneous data tuples given by  $X_i^m$  where  $m$  is the number of features. Multivariate regression was deployed asynchronously, while the approximate linear coefficients were calculated locally to get the overall global regression model.

Reference [50] has also deployed a distributed ranking application on power consumption data. They considered a scenario where the price changes based on the load variation in a smart grid platform. The pattern of power consumption can be ranked locally, and by aggregating the local

rankings, a global rank can be established. As for a large system, centralized systems may not be effective enough as a node-based distribution system that can communicate with its neighbours and become more efficient. The global aggregation is inspired by the Condorcet criterion that declares a candidate as the winner that has majority votes against all probable candidates. In reference [50], the time periods that were considered as candidates and the period that had the highest energy consumption was ranked at the top, in the global aggregate ranking.

In reference [51], a large collection of signal processing schemes, as well as statistical models, have been reviewed for solving, forecasting and optimizing problems. The summary is presented in multiple tables that give a brief idea of how signal processing schemes are deployed on each set of problems.

In reference [52], a pattern discovery (PD) based fuzzy classification system is proposed for power system dynamic security assessment. The novel method at first improves performances of the traditional pattern discovery schemes by adding a method called centroid deviation analysis. Then the improved PD system is deployed, and the results are used in a fuzzy logic-based classification method security index of a given power system. The pattern discovery process discussed here is developed on residual analysis and recursive partitioning. The PD system relies on some basic conceptions such as events  $E_i$ , volumes  $V$ , observed frequency  $n_{oi}$  and virtual frequency  $n_{vi}$ . The residual analysis is carried out to understand how event frequency departs from a uniform statistical distribution. The events are further categorized into significant, insignificant and negative significant groups. Then a recursive partitioning is used based on four predefined conditions to trace out significant events. After that, a well-defined centroid deviation analysis is carried out to improve the accuracy of the pattern discovery. Pattern discovery can detect multiple patterns for a given operation point (OP). Based on those patterns the system must predict whether the power system is in the secure or insecure state. At this stage, a fuzzy classification is used to classify the system states either as 'secure' or 'insecure.' The article implements this PD-based fuzzy classifier

initially using 6000 operating points in an IEEE-50 machine system. Finally, it concludes with a noticeable improvement in the pattern discovery scheme for identifying the power system transient disturbance assessment problem.

Reference [53] discusses an IDS (intrusion detection system) that classifies system behaviour into three categories; specific disturbances, normal operations, and cyber-attacks. It utilizes the knowledge obtained from the sequences of an event called 'common paths.' Time-stamped sensor data have been chosen to find such common paths. A state  $S = \{TS, f_1, \dots, f_n\}$  is proposed to have features  $f$  and a normalized time stamp  $TS$ . The path mentioned earlier can be denoted with  $P$  where the  $i$ -th path  $P_i$  has a number of states  $\{S_1, S_2, \dots, S_N\}$ . The sequence is a subset of the path.  $G$  is the set of all paths considering an event  $Q$  has taken place.  $Q$  can be defined based on some support thresholds. A sequence may have multiple paths and based on the support threshold criteria a common path for any event can be set. Once several common paths have been identified a classification is carried out to compare an event to those common paths. If there is more than one candidate, common paths become maximal and the event is categorized as unknown. For testing the method, a distance-protection scheme in a three-bus two-line transmission system has been deployed. Different scenarios such as a single line to ground fault, fault replay attack, open relay problem, command injection attack, normal system operations have been deployed to test the methodology. The result showed around 94% accuracy of the tested scenarios.

A Kohonen neural network has been implemented for dynamic security assessment and classification based on transient stability issues in reference [54]. It has used steady-state pre-contingency operating variables to figure out whether the system would shortly sustain a line fault or outage. Correlation between pre-contingency operating variables and critical clearing time has been the key to train the neural network. The algorithm successfully classified 90% of the test events.

Dynamic security assessment has also been addressed in reference [20]. It proposed a growing hierarchical self-organizing map (GHSOM) technique via a supervised artificial neural network. As

the dynamic security assessment for a large scale power system is a colossal computational task in reference [20], the authors reduced the complexity by dividing the problem into smaller tasks and then by training dedicated neural networks to classify each of the tasks. Here PNSI (project network with stochastic interconnects), which is a three-layer supervised neural network, has been used. The array of PNSI took pre-fault state vectors as inputs, whose features have already been selected in parallel, to estimate the post-fault state vector for each contingency. Predefined state vectors for both pre and post-fault scenarios have been used to train the networks where the hidden layer units have been determined adaptively. The outputs of the network were trainable projection matrix. The output is then fed to GHSOM trained to classify the post-fault vectors. The method projected three clusters, namely safe, critical and unsafe operation points of an IEEE-9 bus model, having 14 lines and three generators. The proposed method showed accuracy and high-speed execution for a large-scale power system.

In reference [55], a decision tree (DT) has been proposed to enhance the dynamic security of a power system. The idea is to determine the transient stability related security regions and their boundaries. Rules developed by the decision tree algorithm, sensitivity analysis, and optimal power flow are the key driving criteria in this work. DT provided security margins for the ongoing operating conditions and based on that either generation rescheduling or load shedding has been implemented to enhance security measures. The knowledge base of the security regions has been enhanced to facilitate accurate and robust decision-making. The method has been tested on an 'Energy Power System' with two specific  $N - 2$  contingencies. The overall method showed significant promise. Reference [56] has also dealt with dynamic security assessment, focusing transient stability, using a neuro-fuzzy based approach.

In reference [2], hierarchical clustering is applied as a means of post-disturbance analysis to figure out the dynamic signatures of the power system. The system response is obtained by the Monte Carlo simulation. The machine learning approach taken here is an off-line operation to classify dy-

dynamic data on-line. The hierarchical clustering uses an agglomerative strategy to merge clusters of one level to the next one. It uses a linkage criterion to link sets of objects into the hierarchical tree. This method is applied to the data obtained for every contingency. By suitably cutting the trees, a group is prepared to represent the patterns in generators that vary contingency to contingency. The number of clusters in each loop is not specified, rather they have been calculated automatically. The generator groups are developed based on their rotor angle characteristics for each contingency. The cut-off criteria of the tree are predefined by the rotor angle stability margin. If the rotor angle crosses a certain value, then the system is identified as unstable, otherwise stable. The members of each cluster can always stay within a common distance ( $360^\circ$ ) by this method. This approach only considered the rotor angle at the last instance which may affect taking an optimal solution. The article also proposed a second approach that considers the complete simulation of the swing curves of the rotor angle of the generators. Euclidean distance between two rotor angle vectors is considered to define the stable contingencies. The process is summarized in Figure-2.2. A key observation can be made that, a simplified measurement of maximum Euclidean distance can help linearizing data for a group-wise pattern recognition scheme.

The method is then applied on a 16-machine, 68-bus reduced order equivalent model of the 'New England Test System' and 'New York Power System' (NETS & NYPS). Both the methods have proven their effectiveness on grouping the generators in many contingencies. The article concluded that these methods are suitable to label data for identifying potential dynamic signatures based on these groups.

In reference [57], a method to carefully select data and define them in a simple linear regression is proposed to do short-term load forecasting. The foundation of this article lies on simplification of the complexities faced by other popular methods such as exponential smoothing, autoregressive integrated moving average, neural networks, fuzzy systems, etc. For example, ARIMA models are popular but linear. The selection of order in an ARIMA model is difficult to apply, which can be

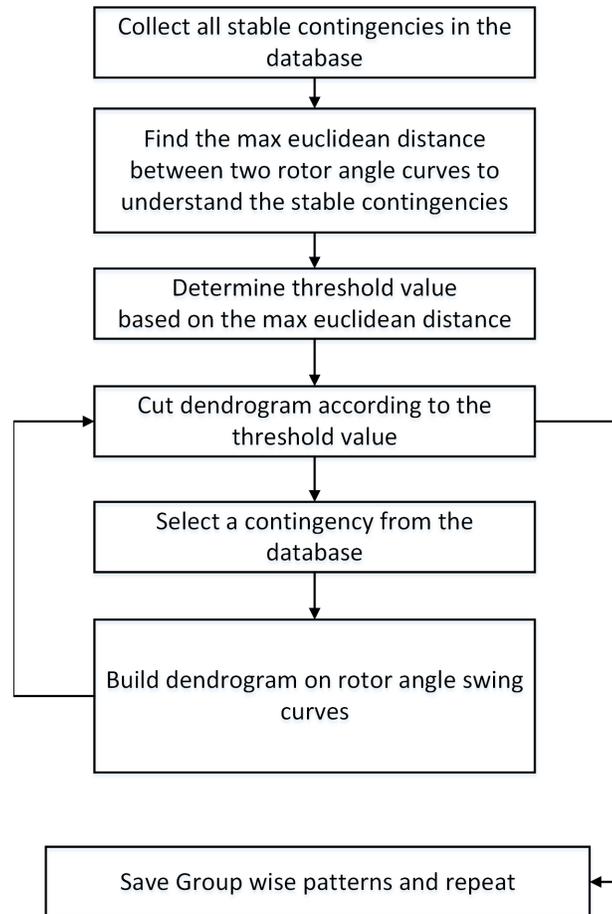


Figure 2.2: Proposed second approach [2]

considered as an obstacle. Here data processing is carried out to define the patterns of the seasonal cycle which helps by eliminating the nonstationary and filtering trend and cycles longer than the daily one. The method identifies two patterns namely input  $x$  and output  $y$  patterns. To find out the input patterns the first vector mean of the load, up to a particular point, is subtracted from that data point and then divided by the length of resulting the vector. This process provides a normalized

vector of input patterns. In a similar process the output patterns are obtained, which is the vector of future load patterns, but instead of using the mean and length of the unknown forecasted values, the known mean and length used to find out input patterns are used. It is done so that by using a simple transformation the loads can be figured out from the output patterns, as the mean value and length is known, and the  $y$  patterns are forecasted. This pattern-based linear regression can be summarized as the following;

1. Mapping original values into patterns
2. Selection of K-nearest neighbors of the query pattern where proximity refers to the same day of the week.
3. Construction of the linear regression model
4. Determination of the forecasted value from the mapping
5. Decoding forecasted values to get original load value

The proposed model is then compared with the popular models for load forecasting such as ARIMA, exponential smoothing, multi-layer perceptron and nonparametric regression model. The result showed mixed performance where the proposed method was proven to be as effective as the others, if not better. In this thesis, a similar concept has been explored in Chapter-8.

Reference [58] talks about an operation optimization decision support system that implements technologies such as data warehouse and data mining. A fuzzy association algorithm is used here to find out associative rules in quantitative data. The fuzzy rules enable a smooth transition between the member and non-member. The overall decision support structure is shown in Figure-2.3

Reference [59], discusses different aspects of the sources of the big data in the smart grid and provides an outline of how to use segmentation to reduce it. It mainly works with data obtained from digital fault recorder during faults. The first step of data reduction is segmentation. Out of

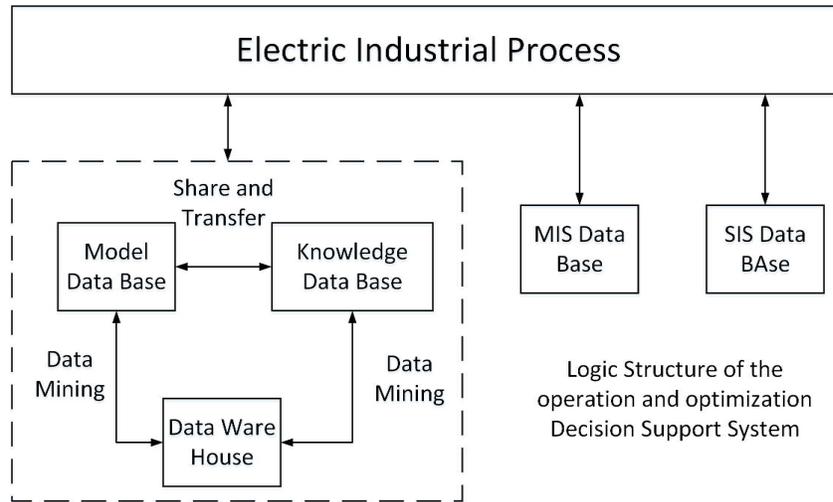


Figure 2.3: DW and DM Based Decision Support System

the different segmentation techniques this work advocates for the wavelet-transform based method. The purpose of segmentation is only to focus on the fault section. The idea is to save our work with a segment that only holds the interesting portion of the data.

Power system analysis often poses a challenge to work with narrow security and stability margins, with high expected reliability. In reference [60], this challenge is addressed by focusing on accurate dynamic modelling, updating parametric values within shorter intervals, applying filtering mechanisms and at the same time discussing computational challenges for achieving all these goals.

[61] proposes a fractal theory and reconstructed phase space-based time series data mining for power system disturbance analysis. The fractal dimension is associated with a set in its space that defines the thickness of the set by using an index. Fractal dimension is a capacity dimension also known as a box dimension. If  $F$  is a limitary nonempty subset and  $N(\epsilon)$  represents the minimum number of the minimum set that has a diameter of  $\epsilon$  to cover  $F$  then fractal dimension ( $D_B$ ) can be expressed;

$$D_B = \lim_{\epsilon \rightarrow 0} \frac{\ln N(\epsilon)}{\ln \frac{1}{\epsilon}} \quad (2.2)$$

$I(\epsilon)$  is information equation which can replace  $N(\epsilon)$  in the above equation, where  $I(\epsilon)$  is equal to  $\sum_{i=1}^{N(\epsilon)} -P(\epsilon, i) \ln (P(\epsilon, i))$ .  $P(\epsilon, i)$  is the probability of the element in the set that belongs to a subset of  $i$ . Fractal dimensions can also be called correlation dimension;

$$D_B = \lim_{\epsilon \rightarrow 0} \frac{\ln C(\epsilon)}{\ln \frac{1}{\epsilon}} \quad (2.3)$$

where  $C(\epsilon)$  is the correlation function that depends on  $H$ , Heaviside function. The reconstructed phase space theory deals with a single variable time series. It at first introduces a delay factor  $\tau$  a then divide the series into a  $m - dimension$  phase space where  $m$  is an embedded dimension.

$$X_i = [x_i, x_{i+\tau}, \dots x_{i+(m+1)\tau}]^T \quad (2.4)$$

where  $i = N - (m - 1)\tau$ . This work has found out the fractal number of eight out of ten power disturbances. Also, the authors here found out the trajectories of those disturbances to compare them to a normal sinusoidal signal. Both the methods proposed here yielded interesting results.

In reference [62] two-time series fitting models have been proposed to evaluate wind energy data for power systems. The first series can simulate auto-correlation, seasonal property and diurnal distribution of wind energy. The second proposed series can only simulate the auto-correlation of wind speed.

Reference [63] also discusses data mining techniques on time series data and claims that data mining methods can outperform traditional statistical analysis based methods. The field of application chosen here is electricity market analysis. Authors have explained how different methods worked in time series analysis and presented a review on those. The methods discussed here are linear methods such as AR, MA, ARMA, ARIMA, ARCH, etc. Nonlinear forecasting techniques

such as neural networks, support vector machine, nearest neighbour have also been analysed. Rule-based decision-making, and wavelet techniques have also been covered. Along with those, some of the ensemble methods have also been mentioned.

Dynamic regression and transfer function models have been discussed in reference [64], for 24-hour electricity price forecasting. In a competitive electricity market, environment time series data exhibits the following characteristics;

1. High Frequency
2. Inconsistent mean and variance
3. Multiple seasonal attributes
4. Calendar effect
5. High degree of volatility
6. Unusual pricing phenomenon during high demand

The outline schemes utilized in both the methods are;

1. Step-1: A model is prepared based on an underlying hypothesis.
2. Step-2: The parameters of that model are estimated.
3. Step-3: If the aforementioned hypothesis is validated then continue from step three to four.  
Or otherwise start again from the initial step.
4. Use the model for forecasting.

The dynamic regression model establishes a relation between the price at hour  $t$  to the previous prices at hour  $t - 1$ ,  $t - 2$ ,  $t - 3$  as well as previous demands at hour  $t - 1$ ,  $t - 2$ ... etc. This is

done to produce a model that has uncorrelated errors. Initially a model is chosen to work with. This model can be expressed as;

$$p_t = c + w^d(B)d^t + w_p(B)p_t + \epsilon_t \quad (2.5)$$

where  $p_t$  is price at time  $t$ ,  $c$  is a constant and  $d_t$  is the demand at that time.  $w^p(B)$  are the polynomial functions of the backshift operator  $B : B^l z_t = z_{t-l}$  and  $\epsilon_t$  is the error at time  $t$ . At the initial state the terms are chosen randomly from a normal distribution having zero means and a constant variance of  $\sigma^2$ . This approach heavily depends on the selection of appropriate parameters. The parameter selection process is an iterative one which is validated using series residuals.

The transfer function model deals with a serially correlated error. The generalized model can be defined using the following equation;

$$p_t = c + w^d(B)d_t + N_t \quad (2.6)$$

Here the new term  $N_t$  is a disturbance term that follows an ARMA model  $N_t = \frac{\theta(B)}{\phi(B)}\epsilon_t$ .  $\theta(B)$  and  $\phi(B)$  are the polynomial functions. Finally when the models have been prepared a refinement method is implemented for fine tuning of the models.

Based on the correlation analysis of the multi-dimensional time series data a fault prediction method is proposed in reference [65].

A power load clustering model with time series data is used in reference [66], to design a controller that ensures the stability of a SMIB (Single Machine Infinite Bus) system. The stabilization is achieved by using two schemes, linear quadratic regulator based optimal controller and robust fuzzy controller. However, [67] introduces the concept of local linear model and simultaneous confidence band on a nonlinear solar power generation data.

In reference [68], a novel technique that identifies coherent clusters of synchronous genera-

tors in a multi-machine scenario has been discussed and evaluated. The choice of algorithm is agglomerative hierarchical clustering. The technique clusters synchronous generators based on the coherency measures obtained from time domain response data, collected after a disturbance. The strength of this work lies in addressing as many events that lead towards oscillation as possible. The algorithm is claimed to be fast and accurate. The article, however, does not discuss much of the dynamic data preparation methods for the algorithm. The sampling rate and frequency of updating the matrices are not the focus of the work, thus it becomes imperative to do a query on dynamic data preparation methods that can enable implementation of such clustering techniques in on-line operation and control applications.

The above discussions led this work to choose a path where initially a simpler method is proposed and evaluated regarding the dynamicity of a power system data. To achieve that goal the performance of a popular pattern recognition algorithm namely the artificial neural network is discussed under different contexts such as forecasting data and predicting controlling parameters. The underlying objective of this approach is to observe the performance of the algorithm, while addressing a dynamic set of data generated by a power system. Based on the performance of the artificial neural network, later stages of this research work are designed, where the key focus would be to implement dynamic characteristics in a data mining method to perform tasks on an on-line basis.

The following section discusses the basic structure of some of the well-known machine learning platforms.

### **2.0.3 Existing Methods for Dynamic Data Analysis**

This section introduces some of the previous works both on the dynamic data mining techniques as well as dynamic mining methods applied on power system data. With the growth of data volume traditional data analytic methods are losing efficiency day by day. Incorporating new information into already discovered knowledge should be the modern approach. Thus, arises the concept of

dynamic data mining, which is a relatively new concept, that analyses the impact of data mining applications on a dynamically changing data set. Unlike a single deterministic run, dynamic data mining techniques run continuously to explore more and more rule spaces. Thus the problem of over generalization can be resolved [69, 70]. The following sections will briefly discuss some of the interesting research work, relevant to dynamic data analysis, carried out in this field namely dynamic item-set counting engine, neighbourhood rough sets, fuzzy recurrent neural network, dynamic cluster tree implementation and adaptive neuro fuzzy inference.

### **DICE Method for Dynamic Data Mining (DDM)**

Traditional data mining techniques often must heavily depend on pruning to minimize the space of rules. However not all the applications allow pruning to be effective either because of restrictions, or because pruning cannot reduce much in those cases. The association mining with a large data set often produces potential rules in exponential quantities. Dynamic data mining (DDM) can be of use in those instances, as it runs continuously. With DDM exploration rule spaces become more comprehensible than a single deterministic run [69]. The DDM method presented in reference [69] relies on user defined weights and is guided by heuristics called heavy edged property. Heavy edged property is an assumption on the underlying data and weight function. It relies on the assumption that heavy hyper edges are likely to have heavy neighbours. The concept of hyper edge is slightly modified than that of edges. An edge is a set of two nodes or vertices while a hyper edge is a collection of vertices.

Dynamic data mining (DDM) introduces some architectural changes to the traditional data mining process. Instead of having a closed end, finite process or complete set of items, DDM works with a continuous process. The process does counting in intermediate stages, thus waiting till the end is no more required. The user defined weights can be continuously adjusted on the fly. Thus, they are useful to dynamically adjust the responses obtained from DDM.

This DDM uses a method called dynamic item set counting (DICE). The key aspects of DICE are:

1. The DICE keeps track of its current position in the given data set. This position after wrapping the data does not reset itself rather it increments.
2. For each item set the DICE keeps track of counting of that item set with the position mentioned previously. Once the data set is completed it halts and completes counting the item set.
3. For higher efficiency in the process of counting the item sets are kept in a hash tree.
4. New items are added to the hash tree with its position in data set and a count of zero from that particular position.
5. If required items are deleted with their memory to facilitate adding newer more significant items.
6. For unknown yet important item sets the DICE item set manager adds counter and resets its position as zero in the data set so that the collective effort of new previously unknown items and old known items can be observed from the very beginning to develop effective rule spaces.

There are different types of weight functions available. Let's define  $P(s)$  as the support of an item set  $S$ . Where support of a function is the set of points or values for which the function yields non-zero output. Mathematically for any function  $f(x)$  this is defined as;

$$support(f) = \{x \in X | f(x) \neq 0\} \quad (2.7)$$

The weight functions then can be defined as  $\text{LogInterest}(S)$ ,  $\text{SetInterest}(S)$  and  $\text{SetConviction}(S)$ .

$$\text{LogInterest}(S) = \frac{\log\left(\frac{P(S)}{\prod_{i \in S} P(\{i\})}\right)}{|S|} \quad (2.8)$$

$$\text{SetInterest}(S) = \max_{i \in S} \frac{P(S)}{P(\{i\})P(S - \{i\})} \quad (2.9)$$

$$\text{SetConviction}(S) = \max_{i \in S} \frac{P(S - \{i\})(1 - P(\{i\}))}{P(S - \{i\}) - P(S)} \quad (2.10)$$

Where the support of all the immediate subsets must be known before the calculation of weight is started. Thus, a well normalized log-interset function is a preferable choice. In brief, log-interset function can be expressed as the total number of occurrences of an item-set divided by the expected quantity of that item-set. This ratio is then normalized on a log scale. To make an item-set interesting, a support threshold can be set. The key to implementing the DICE method is to use effective sampling of both the data and item-sets. This method is interesting to mine a large data set, where it is nearly impossible to explore all the possible spaces of rules [69].

### Neighbourhood Rough Sets for Dynamic Data Mining

The fundamental concept of rough set theory is to approximate the classifications of knowledge for an item-set of interest. The knowledge being a set the rough set theory estimates its upper and lower spaces. The lower spaces are definitely some of the elements of the interest subset where the upper spaces are probable elements of that subset. The approximation space  $\kappa = (U, R)$  has two elements  $U$  called universe and  $R \subseteq U \times U$  is an equivalent relation on  $U$ . The  $\kappa$  is called approximation space.  $R$  divides  $U$  into separate subsets. The subsets are represented as a quotient set denoted by  $U/R$ . Any element  $x$  if belongs to  $R$  - th equivalence class the class including  $x$  can be denoted as  $[x]_R$ .

The approximation space  $\kappa = (U, R)$  is understood through an information system  $S = (U, A, V, f)$ ; where,  $U = x_1, x_2, \dots, x_n$  is a non-empty finite set of objects.  $A = a_1, a_2, \dots, a_m$  is a non-empty finite set of attribute or features.  $V$  is the domain of attributes and  $f : U \times A \rightarrow V$ . Any typical information function  $f(x, a) \in V_a$  for every  $x \in U, a \in A$ ;  $f(x_i, a_j)$  represents the value of object  $x_i$  on the attribute  $a_j$ . To understand the properties and characteristics of rough set some definitions are worth mentioning.

Definition1: if  $B \subseteq C$  be subset of attributes and  $x \in U$ . The neighbourhood  $\delta_B(x)$  of  $x$  in the feature space  $B$  is defined as

$$\delta_B(x) = \{y \in U \mid \Delta_B(x, y) \leq \delta\} \quad (2.11)$$

Here  $\Delta$  is the standard distance function. Satisfy all the necessary conditions related to any two points on a two dimensional coordinate system.

1. The distance between any two points  $x$  and  $y$  is either equal or more than zero. It is zero when both of them are the same points.
2. The distance between point one to point two and the distance between point two and point one are equal.
3. The distance between any two point would be either less or equal to the sum of distances between a third point and the two points in hand.

In pattern recognition three distance function are often used. They can be expressed by a single equation;

$$\Delta_P(x, y) = \left( \sum_{j=1}^m |f(x, a_j) - f(y, a_j)|^P \right)^{\frac{1}{P}} \quad (2.12)$$

where,  $x$  and  $y$  are two objects in an  $m$ -dimensional space  $A = a_1, a_2, a_3, \dots, a_m$ .  $f(x, a_j)$  is the value of  $x$  in the attribute number  $a_j$ . Equation-2 is a Manhattan distance when  $P = 1$ , Euclidean distance when  $P = 2$  and a Chebychev distance if  $P = \infty$

Definition 2: If  $\kappa_N = (U, N)$ , where  $U$  is a set of sample and  $N$  is the neighbourhood relation to  $U$ ,  $\delta(x)|x \in U$  is member of that neighbourhood then  $\kappa_N = (U, N)$  is called the neighbourhood approximation space.

Definition 3: For the  $\kappa_N = (U, N)$  where  $\exists X \subseteq U$  the upper and lower approximations are as two subsets defined with the following equations. Here the lower approximations are definitely the part of  $X$  where the upper approximation might be shared by other than  $X$ .

$$\underline{N}(X) = \{x \in U | \delta(x) \subseteq X\} \quad (2.13)$$

$$\overline{N}(X) = \{x \in U | \delta(x) \cap X \neq \emptyset\} \quad (2.14)$$

The boundary of  $X$  is represented by;

$$BN(X) = \overline{N}(X) - \underline{N}(X) \quad (2.15)$$

The degree of roughness of the set  $X$  is characterized by the boundary region. For example, boundaries for decision attributes should be as little as possible. Otherwise the uncertainty level will increase in a decision process [70]. If the attributes generate a neighbourhood relation over the universe then the information system called a neighbourhood relation over the universe. It is defined as  $NIS = (U, A, V, f)$ . Where  $f$  is defined using  $f : U \times A \rightarrow R$ . A neighbourhood infoemation system turns into a decision system  $NDT = (U, C \cup D, V, f)$  if  $A = C \cup D$  where  $C$  is condition attribute and  $D$  is decision attribute.

Definition 4: In the neighbourhood decision table if  $U/D = D_1, D_2, \dots, D_N$ ,  $\delta(x)$  is the neigh-

bourhood information units including  $x$  itself, and generated by the attribute set  $B \subseteq C$  then the lower and upper approximations of the decision  $D$  with respect to attributes  $B$  are defined as

$$\underline{N}_B(D) = \cup_{i=1}^N \underline{N}_B(D_i) \quad (2.16)$$

$$\overline{N}_B(D) = \cup_{i=1}^N \overline{N}_B(D_i) \quad (2.17)$$

where,  $\underline{N}_B(D_i) = \{x \in U | \delta_B(x) \subseteq D_i\}$  and  $\overline{N}_B(D_i) = \{x \in U | \delta_B(x) \cap D_i \neq \emptyset\}$ . The decision boundary region of  $D$  with respect to attributes  $B$  is defined as

$$BN(D) = \overline{N}_B(D) - \underline{N}_B(D) \quad (2.18)$$

[70]

### Fuzzy Recurrent Neural Network for rules extraction

For dynamic data mining a method using 'Fuzzy Recurrent Neural Network (RFNN)' is proposed in reference [3]. The proposed structure can be observed in Figure-2.4. Where  $x_j^l(t)$  is the  $j$ -th fuzzy input to the  $i$ -th neuron of layer 'l' at time step 't'.  $y_i^l(t)$  is the corresponding output signal at the same time step  $t$ .  $y_j^l(t-1)$  is the activation of the neuron 'j' at time step of '(t-1)'. The box elements are memory cells that store the values of activations of neurons at previous time step. Lets define the other parameters such as  $w_{ij}$  as the connection-weight between neuron 'i' and neuron 'j',  $\theta_i$  is the fuzzy bias of neuron 'i' and  $\vartheta_{ij}$  is the recurrent connection-weight from neuron 'i' to 'j' in the same layer. The activation function 'F' for a total input to any neuron 's' is calculated as;

$$F(S) = \frac{s}{1+s} \quad (2.19)$$

So, for the neuron 'i' at layer 'l' the equation stands;

$$y_i^l(t) = \frac{\theta_i^l + \sum_j x_j^l(t)w_{ij}^l + \sum_j y_j^l(t-1)\vartheta_{ij}^l}{1 + |\theta_i^l + \sum_j x_j^l(t)w_{ij}^l + \sum_j y_j^l(t-1)\vartheta_{ij}^l|} \quad (2.20)$$

The fuzzy signals and connections weights are generally fuzzy set of numbers that can be represented using  $L_0 \dots L_{n-1}$  and  $R_{n-1} \dots R_0$ .

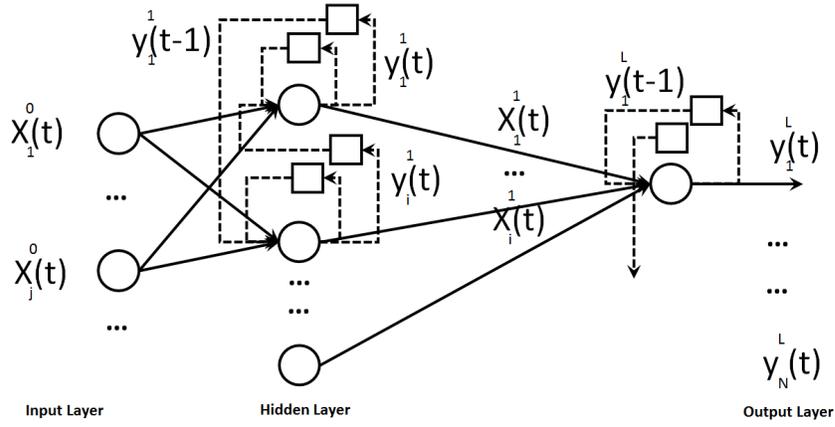


Figure 2.4: Structure of RNN [3, 4]

For the learning of FRNN the article has used genetic algorithm (GA). Here the selection of best genome is carried out using a fitness function. The fitness is calculated from the FRNN error performance indices. The FRNN error performance index is calculated by the following formulation;

$$E_{tot} = \sum_p \sum_i D(y_{pi}, y_{pi}^{des}) \quad (2.21)$$

where 'p' is the learning data entry. 'D' is the distance measure between two fuzzy sets desired and computed. In reference [3], the distance has been calculated using Hamming distance method. Once the total error performance has been evaluated the fitness for any corresponding genome of the GA algorithm is calculated as;

$$f = \frac{1}{1 + E_{tot}} \quad (2.22)$$

### Dynamic Clustering Methods: ClusterTree

In reference [71] a novel ClusterTree has been proposed for dynamic data that embeds a time stamp in data, which can later be used for pattern recognition. This method allows the ClusterTree to remain in the most updated form always. A clusterTree represents clusters of a data set in a hierarchical fashion. Like other clustering methods it also has internal and leaf nodes. The internal nodes assign pointers towards sub-clusters and leaf nodes assign pointers to data points. For individual clusters the ClusterTree calculates several parameters namely: number of data points, volume of the minimum bounding sphere  $S$ , centroid  $c$ . The different centroids are calculated as;

$$c_i = \frac{\sum_{j=1}^N o_{ji}}{N}, 1 \leq i \leq d \quad (2.23)$$

where  $N$  stands for total number of data points and  $o_{ji}$  is the  $i$ -th value of data point  $o_j$  in the cluster. In this way each of the clusters can be presented inside a hyper sphere  $S$ . Any two hyper-spheres may share a common data region and some hyper-spheres may not contain any data points, in a given time. The cluster density can be thus represented by the following equation;

$$\text{Density - of - cluster}_c = \frac{\text{numberofpointsin}C}{\text{volumeof}S} \quad (2.24)$$

$$\text{Density - of - cluster}_c = \frac{\text{numberofpointsin}C}{\frac{2\pi^d/2_{r,d}}{d\Gamma(\frac{d}{2})}} \quad (2.25)$$

The  $\Gamma(x)$  or Gamma function is defined as;

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \exp^{-1} dt \quad (2.26)$$

where,  $\Gamma(x+1) = x\Gamma(x)$  and  $\Gamma(1) = 1$ . If the density of a cluster is lower than a preselected threshold value, then the cluster will be divided into smaller sub clusters. There are three approaches to associate time information in a cluster tree;

1. Add time information as an additional dimension.
2. Use queuing mechanism to process time
3. Use a separate  $B^+$  tree like structure only to handle time information

In reference [71], the authors used an individual  $B^+$  tree-similar structure to handle the time information. This approach can cater to both the time-relevant and time-irrelevant queries. Time irrelevant queries can be done using the original clusterTree and time-relevant queries can be done using the modified clusterTree and  $B^+$  or here  $B^t$  tree structure to get the final result. Internal and leaf nodes are the two kinds of nodes of the clusterTree. An internal node can be defined as; Node: identifier,  $\gamma$  number of entries,  $ot$  the oldest time when data have been inserted,  $nt$  is the newest time in the order,  $(entry - 1, entry - 2, \dots, entry_\gamma)$ ,  $(m_{node}$  minimum number of entries  $\leq \gamma \leq M_{node}$  the maximum number of entries) where the entries can be defined as  $Entry_i$ : ( $SC_i$  pointer to the  $i$ -th subcluster,  $BS_i$  bounding sphere for the  $i$ -th subcluster,  $SN_i$  number of data point in the  $i$ -th subcluster).

Similarly the leaf nodes can be defined as; Node: leaf node identifier,  $\gamma$  number of entries,  $ot$  the oldest time when data have been inserted,  $nt$  is the newest time in the order,  $(entry - 1, entry - 2, \dots, entry_\gamma)$ ,  $(m_{leaf}$  minimum number of entries  $\leq \gamma \leq M_{leaf}$  the maximum number of entries) where the entries can be defined as  $Entry_i$ : ( $ADR_i$  address of the data point residing in the secondary storage,  $T_i$  time information when the data point is inserted,  $L_i$  link to the time information point of the  $B^t$  tree). The  $B^t$  tree is similar to the  $B^+$  tree that incorporates all the time data during the insertion of a data point. The modified characteristics of the  $B^t$  tree are:

1. No minimum number of entries are required

2. In leaf nodes an extra field is assigned to trace the time data point back to the ClusterTree data point.

The construction of clusterTree includes the construction of  $B^t$  tree in parallel. The clusterTree is dynamically processed via the three following steps: insertion, query, deletion. Insertion allows three types of data points: Cluster points that fall under the defined boundary of a cluster, close by points that fall very near to cluster boundary and random points that either are not bounded by any cluster or do not have any neighbouring cluster points. Queries are of two types. The first kind is time relevant and the second kind is time irrelevant. Time irrelevant queries are associated to position of a data point within a cluster and time relevant queries are the insertion time of that data point. Deletion ensures that the data base manager should be able to delete either obsolete data or a data obtained in a given time frame. Deletion also enables the system to be adjusted by itself.

## **ANFIS**

In reference [72] authors have used Adaptive Neuro Fuzzy Inference System (ANFIS) to model dynamic loads whereas in reference [1] ANFIS has been used to predict water level in a reservoir. Reference [5] has implemented ANFIS to predict dam inflow based on a long term weather forecast. The ANFIS uses a neural network to learn and a fuzzy reasoning system to map input into output. Overall it acts as a feedforward network. Here the verbal power of fuzzy system and numerical strength of neural network are combined. Though time consuming, the ANFIS system is quite strong in learning and classifying data. For simple representation, a fuzzy inference system is often considered as having dual inputs ' $x$ ' and ' $y$ ' that result in an output of ' $z$ '. For a first order Sugeno fuzzy model a typical if-then rule can be established as the following;

Rule 1: If  $x = a_1$  and  $y = b_1$  then output  $z_1 = p_1 \times x + q_1 \times y + r_1$

Rule 2: If  $x = a_2$  and  $y = b_2$  then output  $z_2 = p_2 \times x + q_2 \times y + r_2$  The terms  $p_i$ ,  $q_i$  and  $r_i$  are the linear parameters of the consequent 'THEN' part of the first order fuzzy inferencing model. The

five layered ANFIS model is shown in Figure-2.5.

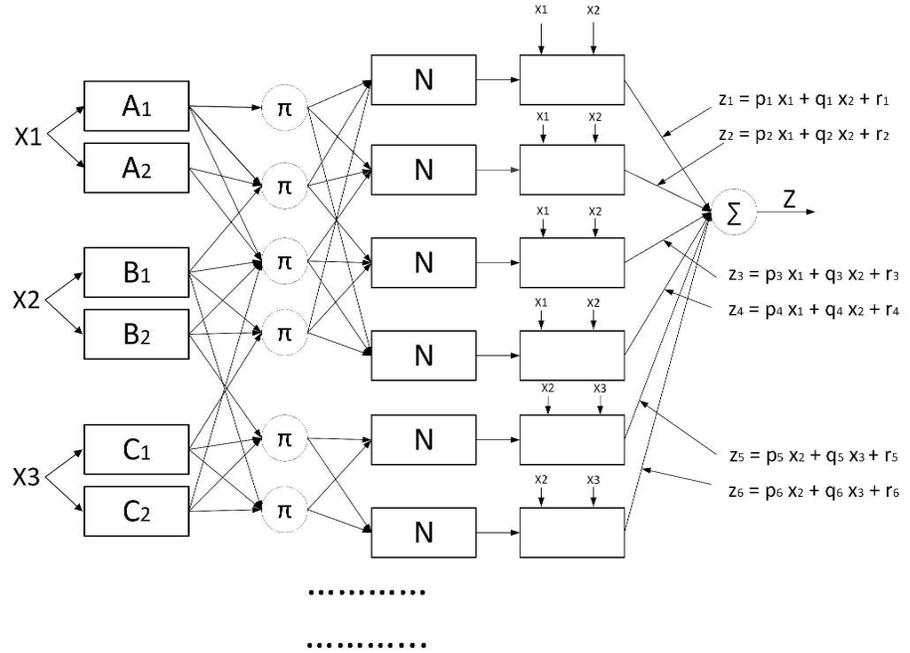


Figure 2.5: Five-layer Two-input Four-Rule fuzzy model (Sugeno) [1]

The first layer is the input node. Membership functions have been used here to generate membership grades at each node that corresponds to the appropriate fuzzy sets.

$$O_{1,i} = \mu_{A_i}(X) \quad \text{for } i = 1, 2 \quad (2.27)$$

$$O_{2,i} = \mu_{B_{i-2}}(Y) \quad \text{for } i = 3, 4 \quad (2.28)$$

where  $X$  and  $Y$  are the crisp input to node  $i$ ,  $\mu_A$  and  $\mu_B$  are the bell-shaped membership functions.  $A$  and  $B$  are the characteristics label defined by the appropriate membership functions. The bell-shaped membership functions can be denoted as;

$$\mu_{A_i} = \frac{1}{1 + \left(\left|\frac{x-c_i}{a_i}\right|\right)^{2b_i}} \quad (2.29)$$

$$\mu_{B_{i-2}} = \frac{1}{1 + \left(\left|\frac{y-c_i}{a_i}\right|\right)^{2b_i}} \quad (2.30)$$

Here,  $a_1$ ,  $b_i$  and  $c_i$  are the parameter set of the membership function that result in changing the shapes of the membership. Parameters in this layer are called the 'Premise parameters'. A tentative example of fuzzy reasoning is shown in Figure-2.6.

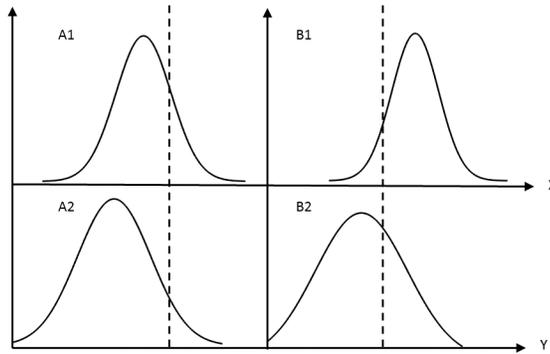


Figure 2.6: Fuzzy Reasoning [5]

The second layer is called the rule nodes. Here the 'AND' operation is implemented to get a single output obtained from the preceding conditions of that rule, which is also called firing strength. Firing strength refers to the range (degree) of the antecedents to be satisfied for that fuzzy rule and also it accordingly shapes up the output. The output  $O_{2,k}$  is actually the products of the corresponding degrees of the previous layer, means the input nodes.

$$O_{2,k} = w_k = \mu_{A_i}(X) \times \mu_{B_j}(Y) \quad \text{where, } k = 1, \dots, 4; \quad i = 1, 2 \quad j = 1, 2 \quad (2.31)$$

The third layer is called the average nodes. The objective of this layer is to calculate a ratio. The ratio is in between two firing strengths, firing strength of the  $i$ -th rule and sum of the firing strengths of all the rules. Once such an operation is carried out a normalized output ( $O_{3,i}$ ) is obtained.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{\sum_{k=1}^4 w_k} \quad (2.32)$$

The fourth layer is called the consequent nodes. It calculates the contribution of each rule towards the final output or layer five. The output of the layer four is denoted by  $O_{4,i}$ .

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i \times x + q_i \times y + r_i), \quad i = 1, \dots, 4 \quad (2.33)$$

Here  $\bar{w}_i$  is the output from the previous layer for the  $i$ -th rule and  $f_i$  is the linear function of the Sugeno fuzzy model.

Finally, the fifth layer is called output nodes. This is a singular node that combines or adds up the contributions from layer four. This is also a defuzzification layer that translates the fuzzy results of individual rule into a crisp output value  $O_5$ .

$$O_5 = \sum_{i=1}^4 \bar{w}_i f_i = \frac{\sum_{i=1}^4 w_i \times f_i}{\sum_{i=1}^4 w_i} \quad (2.34)$$

These layers are trained based on the supervised learning schemes. By supervised learning, it means that the network parameters are adjusted iteratively based on training data set inputs and corresponding outputs. To update these parameters, ANFIS uses both the gradient descent and

least-squares methods. The gradient descent is applied to estimate the nonlinear parameters of the premise  $(a_i, b_i, c_i)$  and the least squares method is applied to estimate the linear parameters  $(p_i, q_i, r_i)$ . Finally, by adjusting the overall quadratic cost function the parameters are fixed. In Chapter-6 a case study is carried out where ANFIS has been used as the chosen machine learning algorithm.

### **Ensemble of Bagged Decision Trees**

The goal of implementing ensemble methods, is to converge towards a solution, produced from a diverse yet individually accurate set of algorithms. The concept, behind this overall process, is to eliminate classification errors as much as possible by obtaining votes from those individual entity (algorithm). By voting, the most popular class is selected. One of the prominent algorithms in the field of ensemble learning is ‘Bagged Decision Trees’ [10, 107, 125]. The term ‘Bagging’ refers to ‘Bootstrap-Aggregation’, where a set of decision trees can grow deep, without considering over-fitting. Then from those deeply grown trees one decision class is voted out, thus the task of aggregation is performed. An individual tree is trained from a subset of the training data set. The subsets are prepared using the replacement technique, which means that each set can have some common data available in one or more other sets. This replacement technique allows different combinations of training sets that in turns can have more samples than the original training set. For example, if  $S_T = x_1, x_2, x_3$  is a training data set then, by using replacement technique more than two training subsets can be prepared  $sub_1 = x_1, x_2$ ,  $sub_2 = x_2, x_3$  and  $sub_3 = x_1, x_3$ . This method has been frequently used in this research, and therefore, is further discussed in the subsequent chapters.

### **Time Series Data Analysis**

Crude time series analysis comes as a natural ability to a human being, but it is considered a complex problem for computers to solve. The common aspects of time series analysis can be divided

into parts, such as representation techniques, distance measures and indexing methods. In recent years, time series analysis has matured from an interest of pure methodology development towards modern problem solving such as economic forecasting, intrusion detection, gene expression etc. One of the key challenges of time series data mining is dimension. The data mining tasks in time series are a set of objectives. Namely these are querying, clustering, classification, segmentation, prediction, anomaly detection, motif discovery. Query by content in a data space can be obtained in two possible ways. Firstly, by e-range query that searches data within a given range. In a two-dimensional space the range can be a circle. The second method would be a k-nearest neighbour search that returns several points ('k') adjacent to the query point. The clustering is the method of finding natural groups in a set of data points. The objective is to find out most homogeneous groups distinct from each other despite having different size. On the other hand, the classification task puts a label on a series of data. The difference between classification and clustering is that classification labels are predefined. Once a new data point is entered, the active set is based on the predefined labels and the new data is categorized into different classes. Nearest neighbour, support vector machine, decision forests are some of the prominent classifiers. The segmentation technique tries to retain the key features of the signal, while reducing its dimension. Piecewise linear approximation is one of the most popular segmentation techniques. Prediction finds out the next few data points in a time series which has not yet been observed in a data set. Prediction is one of the most discussed applications in time series analysis. The anomaly detection figures out abnormalities in a data set by comparing it to the most frequent and prominent features, or in other words the normal characteristics of the set. The motif discovery, another predictive analysis, discovers subsequent phenomenon in a longer time series. The time series mining tasks have some implementation components to deal with. These are essential elements to retrieve patterns from a dataset and often more extensive than the mining algorithms themselves. Preprocessing, representation, similarity measure, indexing fall under that category [73] [74].

# Chapter 3

## Stochastic Data Model

This section introduces the methods applied to generate stochastic data. This data represents energy generation from the wind power plant, solar plant and the noncritical loads. These methods have been implemented in the Matlab-Simulink platform.

### 3.1 Preparation of the Data Models

The first condition of any effective machine learning routine is to have sufficient data [75]. This work is not any different from the others. The underlying pretext is that without a significant amount of data, analysing the implications of dynamic machine learning methods on a power system would be inconclusive. Power system data can be obtained using several possible ways. Here two approaches are taken to collect or generate data. For the initial stages, where different methodologies will be developed and assessed, a Simulink based power system model is used. The purpose of the Simulink model is to generate different types of data under multiple contingencies. The contingencies can be a discrete phenomenon, or multiple contingencies can occur simultaneously. Based on these contingencies, events dynamic power system parameters is collected as viable data for further analysis. The initial data collection approach is not to collect any model specific data, rather to

collect event specific data, so that the developed methodology can be implemented on other systems too. Initially, a simple two area-based power system is chosen. Then for the later stages of this research, where more patterns are to be sought out, larger models are used. Another modelling approach is to use simpler parameters, such as magnitudes, in the initial stages of this research work. For complex decision-making processes, complex parameters are introduced, such as complex power, voltage magnitudes as well as angles, etc. The third and final data collection approach is to collect real-life data from the New Zealand national grid.

Along with generating the data it must be prepared for the machine learning platform. Preparation of data is a prerequisite before the algorithm can be deployed. Based on different types of events and data, different types of data preparation processes are offered in this experiment.

### **Wind Energy Model**

While preparing the wind turbine model for this work, certain conditions have been taken into account such as; the model only considers wind speed as input variable, the velocity of wind is uniformly distributed on the surface of each blade, the air density is constant during each calculation period, the neural network model only predicts the active power generated by the wind turbines. The fundamental equation to calculate the active power from wind velocity is;

$$P = C_{total} \frac{1}{2} \rho A v^3 \quad (3.1)$$

where,  $P$  is active power output,  $C_{total}$  is overall efficiency of the wind turbine,  $\rho$  is air density,  $A$  is swept area and  $v$  is the wind velocity. The energy model used in this thesis is a 9-MW wind farm, rather than a single turbine, and is driven by wind velocity as input information. In some of the earlier research, the wind velocity has been considered as a function of temperature [76]. In this thesis a fixed location has been selected and the wind velocity has been observed. For the training purpose the data has been taken from [6]. The plant shown in Figure-3.3 is a similar example.

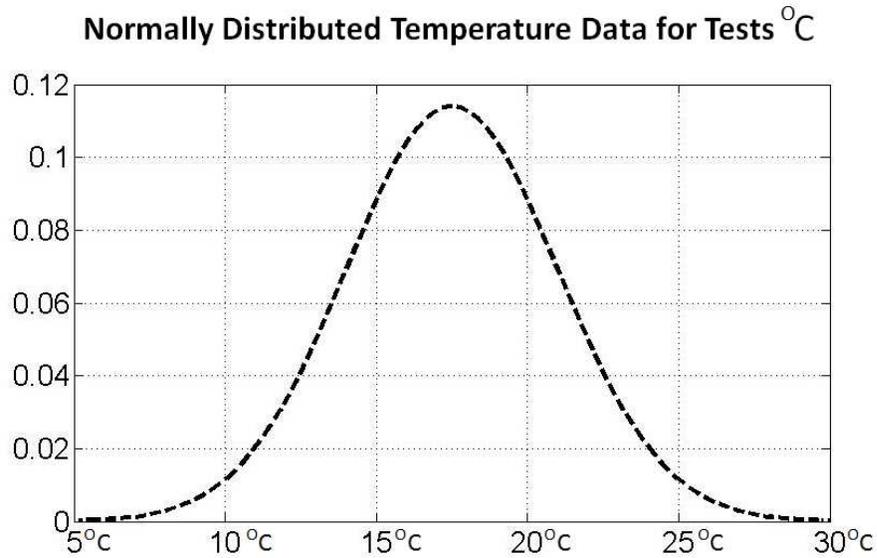


Figure 3.1: Normal Probability Distribution of Temperature (Degree Celsius)

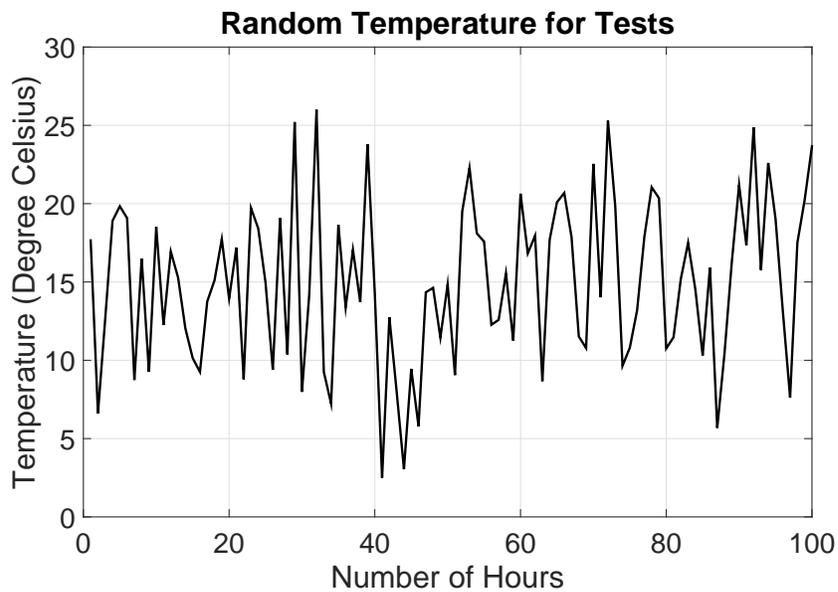


Figure 3.2: Random temperature values (Degree Celsius)

### Solar Energy Model (Photovoltaic)

To simulate the characteristics of solar power generation, a simplified model is used in this work.

The irradiance based photovoltaic cell model of a solar cell is as following;



Figure 3.3: Wind Power Plant

$$I = I_{ph} - I_s \left[ e^{\frac{V+IR_s}{N_1 V_t}} - 1 \right] - I_{s2} \left[ e^{\frac{V+IR_s}{N_2 V_t}} - 1 \right] - \frac{V + IR_s}{R_p} \quad (3.2)$$

where  $I_{ph}$  is the solar-induced current that can be further explained by  $I_{ph} = I_{ph0} \frac{I_r}{I_{r0}}$   $I_r$  is the irradiance in  $W/m^2$ ,  $I_{ph0}$  is the solar current obtained for irradiance  $I_{r0}$ ;  $I_s$  and  $I_{s2}$  are the saturation currents of the Diode-1 and Diode-2 inside;  $N_1$  and  $N_2$  are the quality factors diodes;  $V_t = \frac{kT}{q}$  is the thermal voltage, (k is Boltzmann constant, T is device temperature in Kelvin);  $R_s$  and  $R_p$  are the series and parallel resistances; V is the terminal voltage [77, 78]. Once the PV-model is prepared a solar radiation data model based on air temperature is used for prediction [79,80]. For predicting the solar radiance from temperature data using a neural network, the data model shown in Figure-3.1 and Figure-3.2 has been chosen. To train the neural network, a specific location has been selected with a latitude of  $25.7^\circ N$  and longitude of  $89.3^\circ E$  [6]. Two different locations have been picked to observe sufficient variation in wind as well as solar data so that a robust neural network can be prepared for multi scenario cases.



Figure 3.4: Solar Plant

### **Electrical Load Model**

The application of a neural network to predict electrical load has recently received much attention. Electrical load forecasting is not an easily achievable task. The task becomes even more critical when the economic losses for miscalculations are considered. Though the neural network has yet to demonstrate a high degree of accuracy to convince the sceptics, it holds immense potential [81]. The electrical load has a complex nonlinear relationship with several factors such as temperature, days of the week and time of the day [82]. This thesis adopts the method to predict electrical load based on daily temperature. Figure-3.1 and Figure-3.2 refer to the normal distribution of the temperature model and the randomly chosen set of values of temperature for prediction. The temperature chosen is normally distributed by a probability distribution function having a mean value of  $17.5^{\circ}\text{C}$  and standard deviation of  $7.5^{\circ}\text{C}$  inspired by the work [83]. The mean and standard deviation has been picked to mimic the temperature data for a nearby location mentioned in Figure-

3.4. For training the system a random day in winter (24-hours) has been selected. Then the average hourly load has been plotted against average hourly temperature.

Once the electrical load model chosen, a third-degree polynomial curve fitting has been deployed to find the functional relationship between the electric load data and temperature. The curve fitting performed is a least squares approach to minimize error [84].

$$P_{Load} = f(T) = a_0 + a_1T + a_2T^2 + a_3T^3 \quad (3.3)$$

where  $a$  = multiplying constants,  $T$  = temperature in degree Celsius. With a goodness of fit by  $R^2$  0.9077, the fitting resulted in  $a_0, a_1, a_2, a_3 = 2.385e^8, -1.136e^7, 8.413e^5, -1.747e^4$ . Based on this polynomial fit, the load presumed from the test temperature data is shown in Figure-3.5. The load curve looks random because in this experiment the temperature data for testing is chosen randomly, which is shown in Figure-3.2.

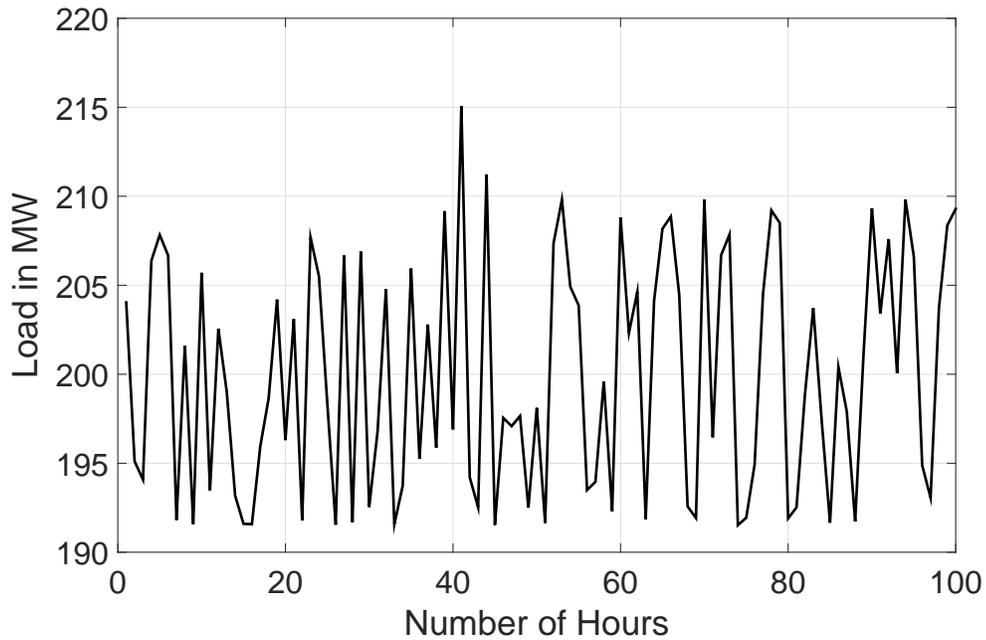


Figure 3.5: Calculated Electrical Load

## 3.2 Machine Learning for Power System Stability: Traditional Approach

After the initial data models have been prepared, a neural network has been trained and tested. For testing the ANN's performance a two-area power system with a hydroelectric power plant, a diesel generator plant and a fixed load have been proposed. The chosen load model is a three-phase RLC load. The value is obtained from the 24 hours ahead estimation, using a nonlinear autoregressive network with exogenous inputs discussed in detail in the later section.

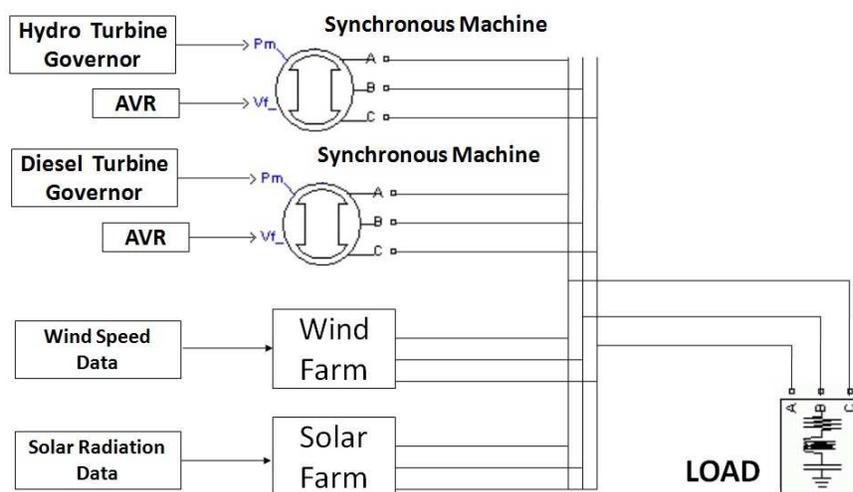


Figure 3.6: Two Area AGC Based System

The hydro turbine governor takes three fundamental inputs to manipulate the output mechanical power. Namely, they are electric power output, the angular speed of the rotor and reference values. On the other hand, the diesel governor works by taking the rotor angular speed and a predefined reference as input data. The hydro power plant has been chosen to have a nominal power of 190 MW and the diesel plant of 25 MW with both synchronized at  $60s^{-1}$  in the Matlab-Simulink platform. Figure-3.7 and Figure-3.8 show the transfer functions designed for the turbine governors for the

purpose of automatic generation control.

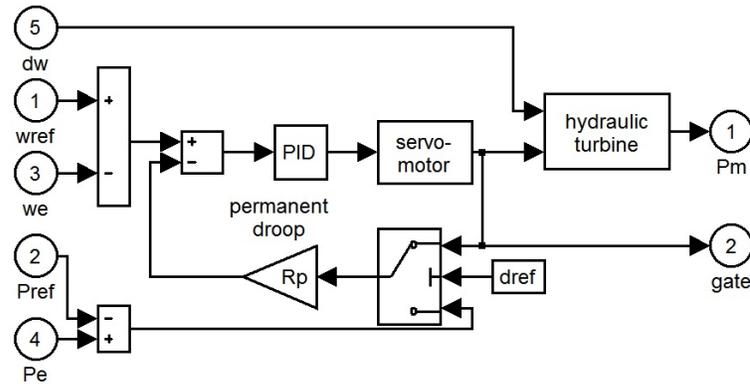


Figure 3.7: Hydro Turbine Governor

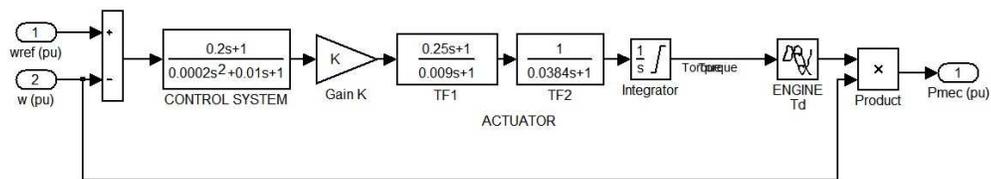


Figure 3.8: Diesel Turbine Governor

After system modelling the neural network has been trained in several steps, hence, this work claims to be a multi-level application of a neural network. The first set of training is carried out with both the hydro and diesel governors. The structure follows a feedforward network with one hundred hidden neurons. To facilitate prediction of both the hydro and diesel power plant the randomly chosen electrical loads are more than the maximum production capability of the hydropower plant. In the training process an array of load scenarios has been chosen and based on those individual loads, turbine governor data has been fed to the neural network for training, testing and validating. The data division is made 70%, 15%, and 15% respectively. The training process is stopped either after one thousand iterations, six consecutive validations with the lowest possible mean squared error or when the mean squared error (MSE) is equal or less than  $1e^{-5}$ . In this experiment for both

cases, the training of a neural network has been stopped after six consecutive validations. Figure-3.9 and Figure-3.10 are showing the mean squared error (MSE) for both the training networks. Identifying the proper network for this thesis has been an iterative process. Once the best MSE is obtained, then the networks have been chosen for the next steps.

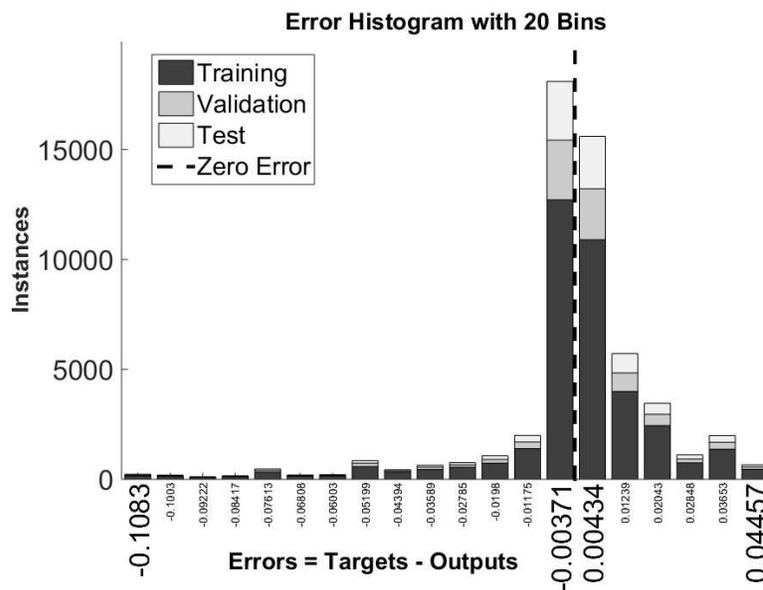


Figure 3.9: MSE of ANN training for hydro power plant

For training the neural network for predicting wind and solar power as well as the electrical load at twenty-four hours basis, a NARX (nonlinear autoregressive with external input) model of a neural network has been used which is shown in Figure-3.11 [82]. The training model uses daily average temperature as input and average daily powers (wind, solar, load) as outputs. Below Figure-3.12, and Figure-3.13 show the 24 hours ahead prediction of the average hourly wind speed and average hourly solar radiance. The actual data has been collected from the areas shown in Figure-3.3 and Figure-3.4 using the software mentioned in reference [6]. The solar radiance data has been collected only at the hours when sunlight is available. The NARX model uses 50 hidden neurons with  $d = 24$  delays, referring to a twenty-four hour ahead forecasting in this study.

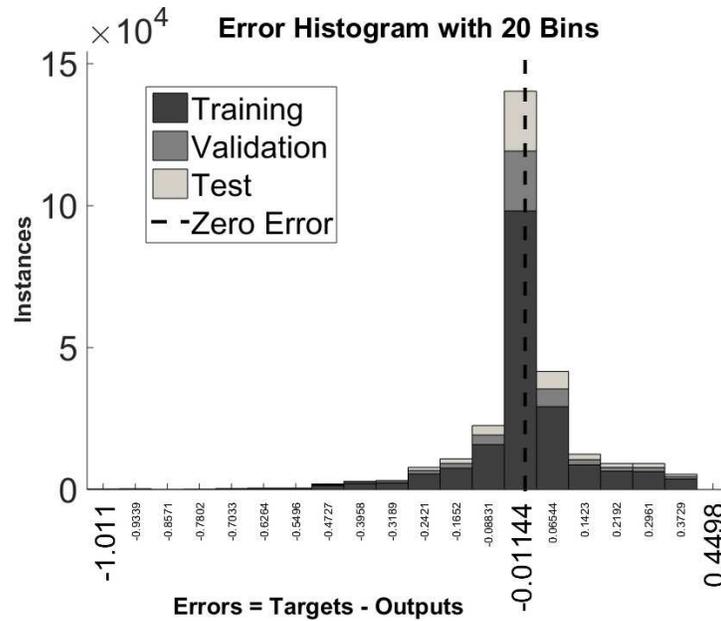


Figure 3.10: MSE of ANN training for diesel power plant

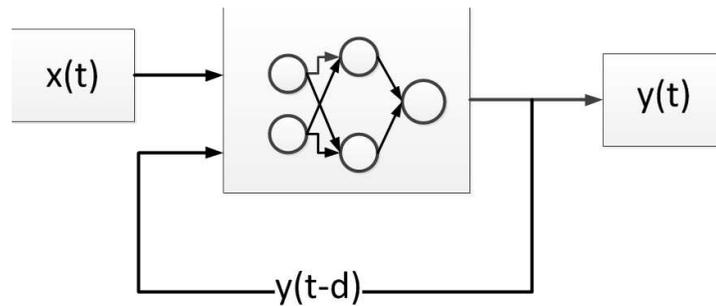


Figure 3.11: NARX network

Through a similar process, the 24 hours ahead prediction of the electrical load has been carried out. The idea of this study is to predict the turbine governor parameters of the hydro and diesel power plant for a certain load level, predicted 24 hours after deducting the predicted generations from distributed generations by wind and solar power plants. Figure-3.14 is the prediction model of the electrical load with a 24 hour ahead basis. Out of the 48 data points of predicted load, to

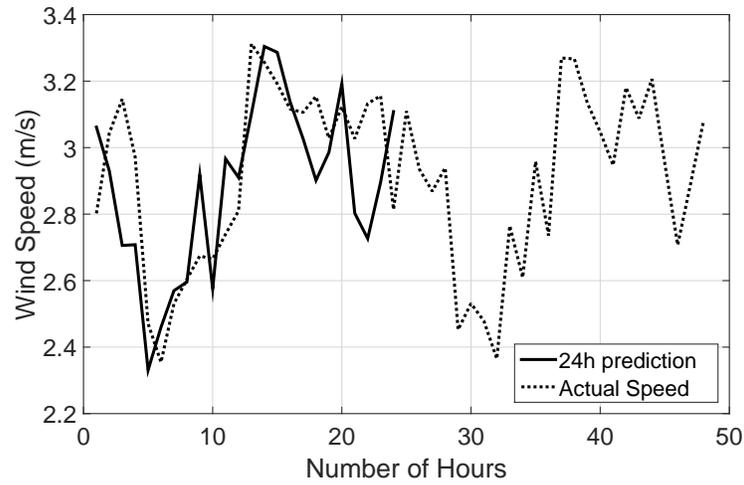


Figure 3.12: Wind speed prediction [6]

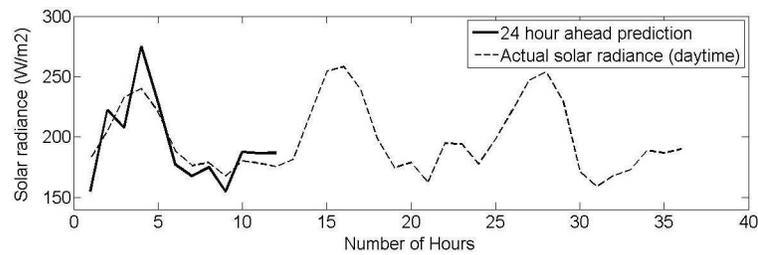


Figure 3.13: Solar radiance prediction [6]

match the average temperature profile of the training data the load area has been chosen to have around 210 MW. Which suits, the parameter values of the generators in this study. The load value has been specified after comparing it to a generation plant mentioned in reference [84] to maintain the practical aspect of this study. The load prediction has been carried out in two ways. Along with the NARX model of a neural network, a fitting model has also been applied to the prediction of electrical loads based on ambient temperature. The reason for applying a second prediction scheme is to compare the outcome of different methods for predicting load. Figure-3.15 is showing the fitting model. The comparison between a neural network fitting model and NARX model compels the authors to select the NARX modelling approach over a neural-fitting approach for such a time

series analysis.

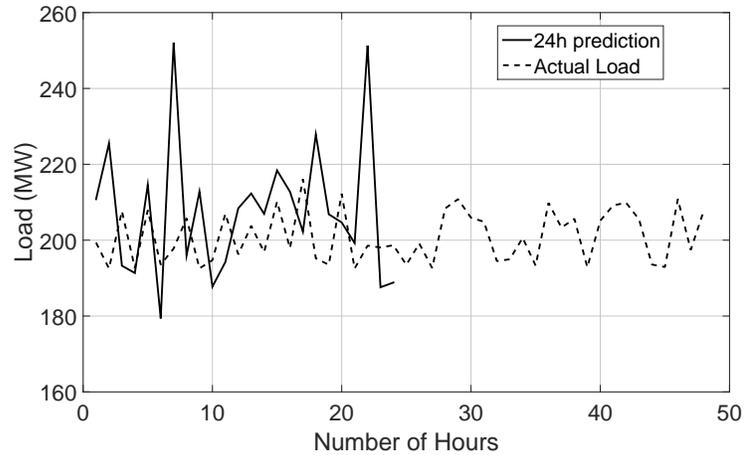


Figure 3.14: Predicted load 24 hour ahead (Fitting Model)

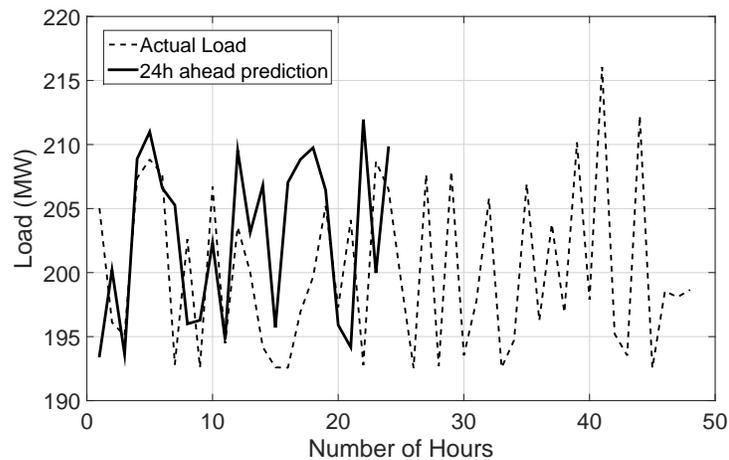


Figure 3.15: Predicted load 24 hour ahead (NARX Model)

### 3.3 Understanding the Limitations of Traditional Approaches

Once the predictive models, as well as 24 hours ahead predictions have been established, the mechanical power inputs to the governors are predicted. In an autonomous solution the equation that will be continuously followed is  $P_{synch\_generators} = P_{load} - P_{wind+solar}$ . To reduce complexity here,

only one instance has been mentioned, where the total generation required from the hydro and diesel plant after considering the distributed generations is 207.5 MW. This indicates that the synchronous generators have to produce 207.5 MW to maintain load balance. The predictions of input mechanical power for both the generators are shown in Figure-3.16, Figure-3.17 and Figure-3.18.

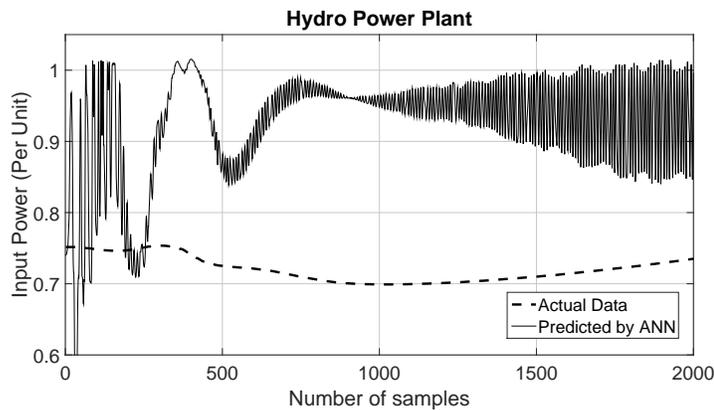


Figure 3.16: Prediction of hydro turbine governor (Transient Part)

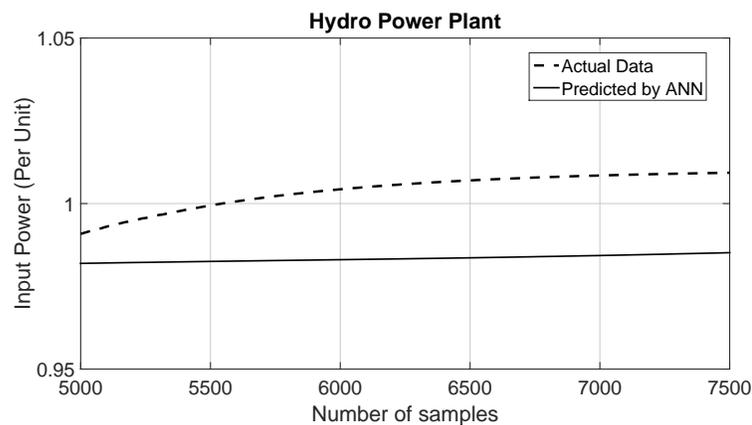


Figure 3.17: Prediction of hydro turbine governor (Steady Part)

From Figure-3.16, Figure-3.17 and Figure-3.18 it is quite evidential that the performance of the neural network during transient phases of the analysis is not up to an overwhelming standard. In the later halves or the stable factions, the ANN model shows sufficient accuracy. Various reasons can be attributed to such inadequacy of ANN for dynamic analysis. The most important reasons would

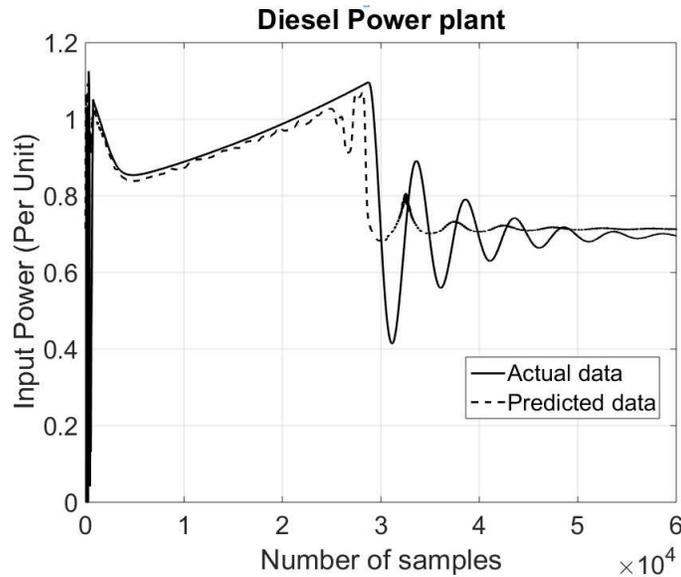


Figure 3.18: Prediction of diesel turbine governor

be the lack of training data as the transient stages are very short lived. Another reason might be that different training of a neural network starts from different initial data points. As the training process stopped after six validations with the lowest possible mean squared error, other training possibilities were thus ignored. Besides, the training data contains both the transient and steady-state data points which should put an impact on how the neural network chooses its weights. Though the responses shown in Figure-3.16, Figure-3.17 and Figure-3.18 are impressive, more tuning could have been implemented as spikes and irregularities are spotted in the transient parts. Hence further modifications have been proposed.

The modified approach divides the data into dynamic and steady state factions and train two neural network submodels [84]. Each model has been tested with either the dynamic data or steady state data. Any phasor that contains oscillation is considered 'dynamic,' and a phasor without any oscillation is 'steady.' With such an approach the dynamic responses improved a lot. The

stable responses remained almost like the previous method. Some of the spikes in the output for the dynamic parts have not been properly reduced. So further steps were carried out. A moving average was imposed on the output obtained from the dynamic responses. However, applying moving average on the output data introduced a small delay in the signal, so a further signal enhancement was done. This enhancement included matching the pick values by shifting the response curves on time domain. The results are shown in Figure-3.19 and Figure-3.20. However, this overall attempt is still rudimentary and not efficient.

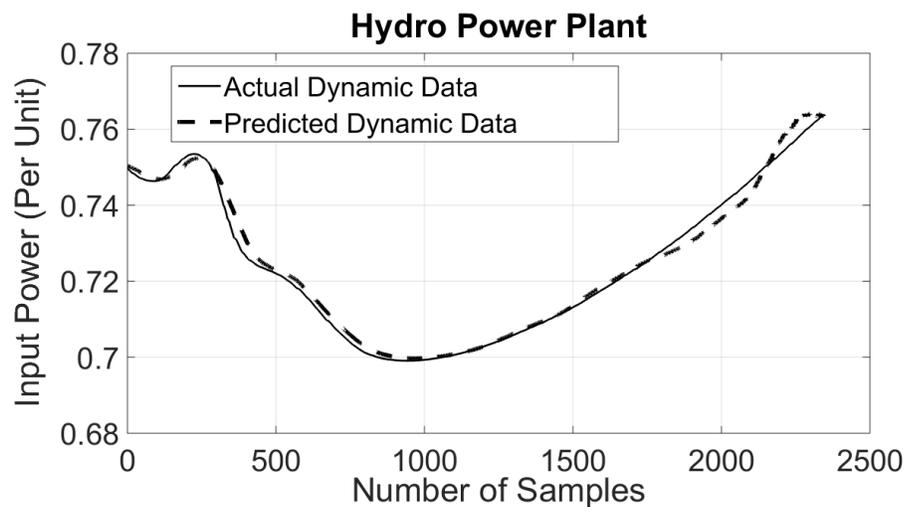


Figure 3.19: Prediction of hydro turbine governor (with sub networks)

Selecting a neural network for solving a nonlinear problem has its demerits too. The application of a neural network in predicting a pattern is already field-proven, however, applying ANN to predict future power generation, loads, as well as control parameters, are not pragmatic without higher degrees of accuracy. Besides, the performance of the neural network varied due to its probabilistic nature. So, a real-time controller of a turbine governor with an ANN might not be suitable during transient periods. Thus, it can be concluded that a better framework is required for dynamic data. The following chapter discusses the development of an algorithm, that can be implemented in an

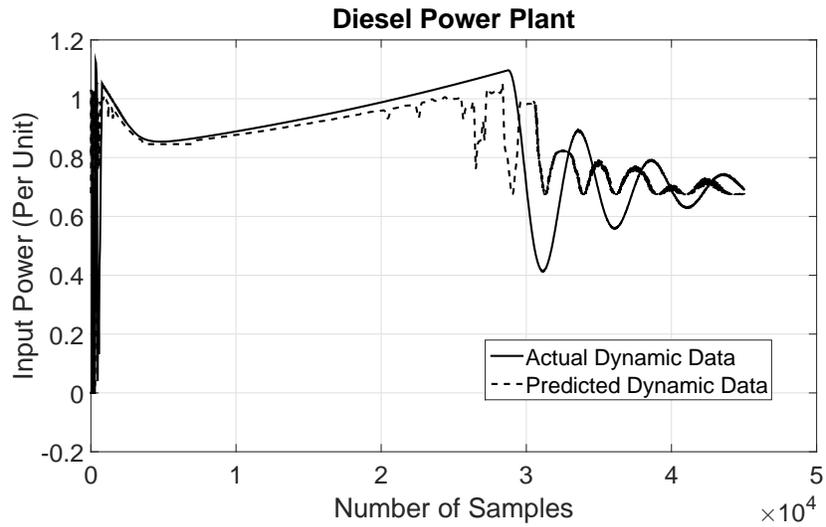


Figure 3.20: Prediction of diesel turbine governor (with sub networks)

online platform for power system post-fault restoration purposes. The idea behind this experiment is to propose a method for dynamic data analysis. The study is conducted on a use case basis. Therefore, a self-healing microgrid has been considered as the subject for data generation.

# Chapter 4

## Machine Learning for Supervised Hierarchical Platform

This chapter addresses the challenges observed in the previous chapter. Alternate means, such as, changing the data analysis method and selecting a suitable algorithm for the data types used, have also been discussed here. The key observation made in this chapter is that a hierarchical platform for the data analysis can be a relevant candidate for a self-healing algorithm. This chapter has been published in the following conference:

This chapter has been published in ISGT-Asia, 2016. ( M. A. Karim, J. Currie and T. T. Lie, “A distributed machine learning approach for the secondary voltage control of an Islanded microgrid,” 2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia), Melbourne, VIC, 2016, pp. 611-616. doi: 10.1109/ISGT-Asia.2016.7796454 )

### 4.1 Post Fault Restoration

The concept of a smart grid emerged to improve reliability and sustainability through intelligent means. The idea is to offer better strategies for the operation that the conventional electric power

system often struggles to realize. Distributed energy generation, deregulated energy market, government incentives and demand response are some of the key attributes that demand an enhanced control architecture in a modern power system. Recent studies show that the microgrids can be considered as the building blocks of a smart grid due to the distributed and stochastic characteristics that allows bi-directional energy flow [85]. A microgrid is a smaller, distributed and somewhat independent entity in a power system that mostly provides energy to a remote distribution network. It usually contains distributed energy generation facilities such as wind power generation, solar power generation, gas turbine based plants, diesel generator based plants, fuel cells, energy storage devices, etc. A microgrid strives to mitigate local demand from its inventory. If required, it can borrow or provide power to the grid [86]. Microgrids can also be operated on a stand-alone basis. Usually, a stand-alone microgrid appears out of a fault, or it is built in a remote location or island.

The voltage and frequency in a microgrid are controlled by synchronous generators. Due to the stochastic or uncertain nature of solar and wind power generation, power electronic interfaces such as inverters and converters are often used in a microgrid. These uncertainties can influence innovative strategies for droop control [87]. As electricity production must be balanced with the consumption, the random nature of the distributed generators often poses threats against maintaining this balance. This characteristic sometimes requires the microgrids to deploy energy storage mechanisms [88]. On the other hand, operation of a microgrid without any energy storage device has also been considered due its cost-effectiveness and high-power quality under normal operation scenarios. However, operating a microgrid without energy storage poses many challenges [89]. In reference [90], an energy storage less control scheme, that utilizes pulse width modulation technique for maximum power tracking from PV-array and fuzzy logic-based diesel generator speed control, is proposed. Such a mechanism is good enough for the PV-diesel based hybrid system. However, the diesel generator used in the article had a maximum capacity higher than the peak load predicted to provide full energy support at night. This planning strategy would not be sufficient for

a wind-PV based isolated microgrid, that uses synchronous generators, having lower capacity than the peak demand.

Microgrids with sufficient distributed energy generation capacity may use low capacity synchronous generators [91], but in those microgrids, during an emergency, optimized load shedding may be a necessity [92]. A typical microgrid often does not accommodate the concept of having a base load, as large variations may be expected due to the frequent use of inductive loads [93]. Due to this reason, optimization techniques are often discussed in the literature to manage the controllable resources [94–96]. Commands based on those optimization techniques can be carried out from a control centre. However, the continuous bi-directional communication could be an expensive venture [97]. To mitigate these issues, research such as [98] proposed decentralized control schemes at every subsystem node, but how the uncertainty of the distributed energy resources influences the control schemes in each of the sections has not been adequately discussed, which is one of the keys focuses in any decentralized microgrid management system [10].

Decentralized control schemes such as droop controller are one of the most popular schemes. This scheme, time to time may suffer from the presence of the small signal instability issues [18]. This study proposes a distributed machine learning algorithm architecture, established on the synchronous generators to provide a secondary control scheme under different scenarios. The purpose of the secondary control is to maintain the desired voltage level throughout a microgrid by identifying and mitigating an imminent voltage instability problem. The designed microgrid would be in a stand-alone mode without having access to an energy storage device. The machine learning algorithms would be trained based on the historical energy production data and the variable demand data.

### **4.1.1 Microgrid Model**

This chapter follows the lumped load on a common transmission grid approach shown in reference [96, 98, 99]. However, this model differs in dividing the system load into two parts: critical and controllable loads [100]. The critical loads are comparable to the base load of a system that remains unchanged throughout a certain period, while the controllable loads are unpredictable in nature and sudden changes can be observed at any given time. The system has two thermal power plants run by synchronous generators: a wind farm, and a solar plant. The overall demand in the system is higher than the total capacity of the synchronous generators. This means that the wind power plant and the solar plant play a significant role in mitigating the demand. The wind power plant deploys an induction generator-based variable speed wind turbine. While the solar power plant is designed as a current source with a diode, multiple resistances (internal and leakage current) followed by a voltage source converter (VSC) [101]. Voltage source converters with appropriate controllers can ensure grid stability, however in this article, a VSC based solar plant is only implemented as a variable energy source. Both the solar plant and the wind power plant are not subject to the secondary voltage and frequency control rather, from the secondary controller's point of view, these plants serve as a data source for the proposed binary classifier using a decision tree based machine learning algorithm.

## **4.2 Proposed Secondary Control**

The purpose of the primary control, which is a part of the hierarchical control scheme in a power system, is mainly to stabilize the system frequency and voltage. The primary control in a standalone microgrid that does not have any storage units which may cause frequency deviation even if the system remains in a steady state condition. The secondary control acts on the deviations observed after the primary control has been implemented and tries to balance the system. The secondary

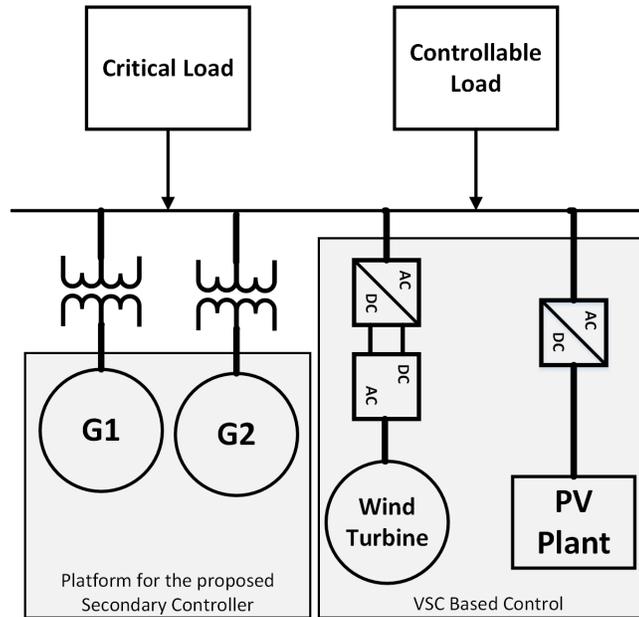


Figure 4.1: The Microgrid Model a Modified Approach from [7]

control has a slower dynamics than the primary control in a power system. The idea behind this is to introduce a decoupling mechanism between these two [7]. Although the usual practice is to implement the secondary control centrally in this study, a distributed supervised secondary control scheme is proposed. The initial feedback data to trigger the secondary control is taken from the transmission line voltage and the frequency of the individual machine. The secondary control is applied after the machine initialization has been taken care of and the system has reached a steady state condition. The proposed secondary controller acts as a restorative measure over the primary controller. This study follows a similar approach for the conventional droop characteristics principle discussed in reference [7] for the primary frequency and voltage control. This principle observes that the droop control coefficients for active and reactive power  $D_P$  and  $D_Q$  should be determined based on the nominal ratings of the individual machine. By calculating  $D_{Q_i}Q_{N_i}$  and  $D_{P_i}P_{N_i}$  and subtracting those from the no-load output voltage and angular frequency, the reference

voltage and angular frequency can be fine-tuned during standalone operations of the microgrid. The overall distribution should satisfy the condition  $D_{P_i}P_{N_i} = \Delta\omega_{max}$  and  $D_{Q_i}Q_{N_i} = \Delta E_{max}$ . Where ‘N’ is the number of distributed energy generators which in this work is two and ‘P’ and ‘Q’ are the generated active and reactive power,  $\Delta\omega_{max}$  and  $\Delta E_{max}$  are the maximum allowable angular frequency and voltage deviations. After that, the dynamic response can be observed through linearizing the active and reactive power equations to find out their small signal models.

$$\Delta P(s) = \frac{G}{s + D_P G} \Delta \omega^*(s) \quad (4.1)$$

$$\Delta Q(s) = \frac{H}{1 + D_Q H} \Delta E^*(s) \quad (4.2)$$

Here,  $G$  and  $H$  depend on the nominal terminal voltage and load angle of the generator and the transmission line voltage where power is delivered. Through the deviation in angular frequency, the primary controller controls the active power reference to the synchronous generators that in turn controls the active power generation and through the deviation in terminal voltage the controller regulates reactive power generation from the synchronous generators. This study implements the secondary control scheme once the threshold of the  $\Delta\omega_{max}$  and  $\Delta E_{max}$  has been breached under a steady state operating condition. Figure-4.2 shows the model of the secondary control system. The secondary control is considered as a part of the hierarchical control of the microgrid [7, 87, 102]. The key element in this control scheme is represented via a switch. The switch is used to move the incoming signal for active power and voltage references between two modes; normal mode and the contingency mode. In the normal mode the incoming reference signals are multiplied by a value of **1** and in the contingency mode the value is modified by the proposed algorithm and multiplied with reference signal as a discrete gain. The switch routes these two signals when needed. In this study, the secondary control is designed based on the distributed machine learning algorithms trained under different voltage stability conditions. The following section discusses the preparation

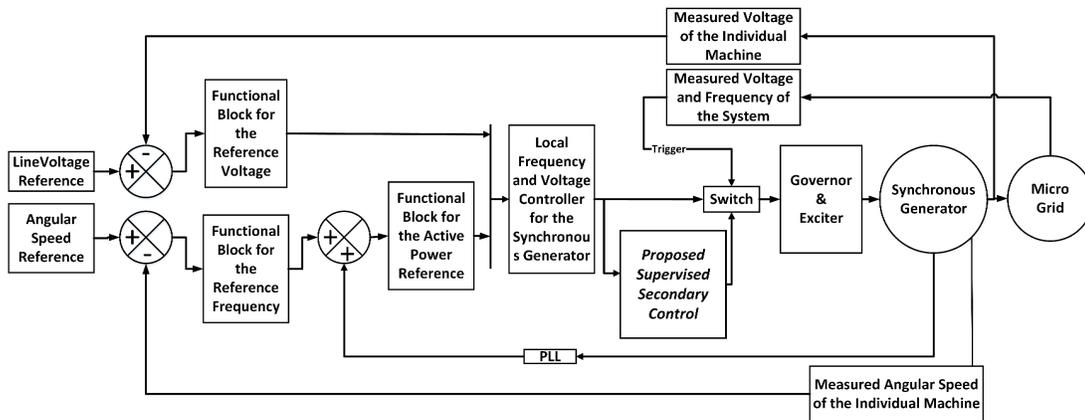


Figure 4.2: Secondary Control Scheme

of the control algorithm. The main objective of designing the algorithm is to maintain the voltage quality at the transmission line [7, 102].

#### 4.2.1 Data Preparation

The proposed machine learning based secondary control scheme is applied to the synchronous machines, based on the amount of power generation available from the wind and solar power plant, as well as the controllable residential loads. The algorithm works in a cascaded form. The initial part of the algorithm is a binary classification prepared with four different attributes and one target attribute. The available wind power and solar power, controllable loads and the sudden changes in load are the input attributes. The target attribute is the binary class of the system either being stable or unstable.

For establishing such a relation in this study, the data set has been prepared out of fifty randomly selected scenarios chosen at a very close to voltage instability point to intentionally reproduce voltage instability at times. Thus, the operating point of the system is chosen near the lowest voltage margin on the system PV-curve [8] shown in Figure-4.3.

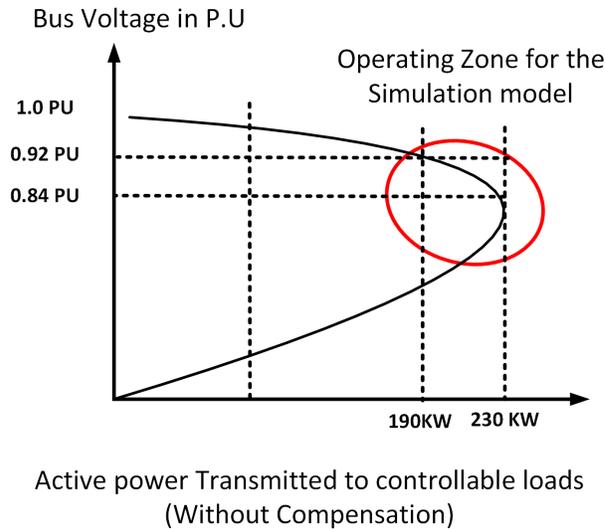


Figure 4.3: Operating Near the Voltage Instability Point (Without Compensation) [8]

As the system has been forced to operate near its stability margin, the probability of losing the system stability increases with a sudden change in load. The simulation model ensures a significant change in load to introduce voltage instability in the system. The sudden change in load breaches the active power margin, that introduces power swing in the system, specially between the two synchronous generators, due to the impact of primary controller in both the locations. To observe the instability of the system, fifty random incidents have been prepared. It is shown in Table-4.1.

The transmission line voltage of these fifty incidents is then observed. The voltage characteristics are distinguished based on its difference  $\Delta V$  between any two points  $N$  and  $N - 1$ . The characteristics of the transmission line voltage are monitored using a four-second window. Within this four second window if the voltage fluctuation crosses the allowable fluctuation-threshold  $\frac{\Delta V_{max}}{\Delta T}$  more than a certain predefined number  $M$ , then the voltage fluctuation is considered as persisting and the individual scenario is labelled as unstable. Otherwise, the voltage of the system is considered as stable. A similar approach is also proposed for the rotor angular frequency at the individual synchronous generator.

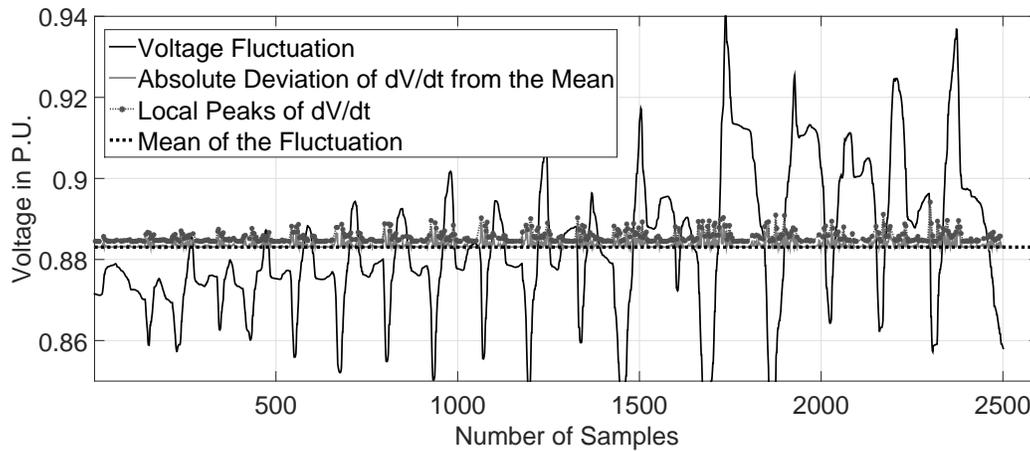


Figure 4.4: Fluctuation During Voltage Instability

## 4.2.2 Preparing the Distributed Algorithm

Once the model data has been prepared, an unsupervised K-means cluster has been prepared. The purpose of using an unsupervised cluster is to narrow down the target data before preparing another machine learning algorithm, the neural network. Unsupervised clustering would make sure that the data falls under one category at a time. Five predefined categories have been chosen for the clusters. The purpose of using the neural network is to do parameter fitting. In this case, the parameters are the references for rotor speeds and the references for field voltages of the two synchronous generators. For a better prediction accuracy, in this study separate neural networks are proposed for five different clusters [56]. The clustering data is shown in the Table-4.1. The clustering process is shown in Figure-4.5.

The second step of the algorithm is to prepare a classification routine, using an ensemble of bagged decision trees. The purpose of this classifier is to establish a binary class of stability. Some of the fifty incidents are stable, and some are unstable. The bagged decision tree classifier is trained with these incidents for future predictions. Once a new set of data arrives, the binary classifier is invoked to predict the imminent outcome of this incident during a sudden and significant load

Table 4.1: Data Table With Clusters (Units in watts)

Wind Power	Solar Power	Residential Load	Load Change	Cluster
200000	200000	203000	28000	4
187500	200000	204000	50000	4
225000	100000	198000	50000	5
187500	150000	198000	47000	1
...	...	...	...	...
187500	100000	204000	47000	5
225000	100000	191000	45000	5
200000	50000	192000	42000	3
187500	100000	201000	46000	5

change. If the classifier predicts probable instability, based on the previously prepared cluster value of that particular incident a suitable neural network is selected. The purpose of the neural network in the individual cluster is to propose necessary modifications in the active and reactive power control that is currently under operation through a primary controller. The modifications are applied based on the scenarios if the systems were stable with a similar amount of remaining residential load. It means the neural network under each cluster is prepared based on the final value of the residential load ('post-event') considering that the sudden change has occurred. The system is simulated with the 'Post-event' load scenario maintaining a steady state condition. The output of the neural network predicts the appropriate value of the reference rotor speed  $\omega_{SG}$  as well as reference field voltage  $V_{fld}$  of the individual synchronous generator (SG) at the steady state condition. The purpose of the neural network in this study is to calibrate the primary control system for those above fifty predefined conditions. The overall architecture of the algorithm is shown in Figure-4.6.

### 4.2.3 Results and Discussions

The test microgrid shown in Figure-4.1 has been used to apply the fifty different scenarios close to the system instability points. Then the decision tree was applied and tested on the system data

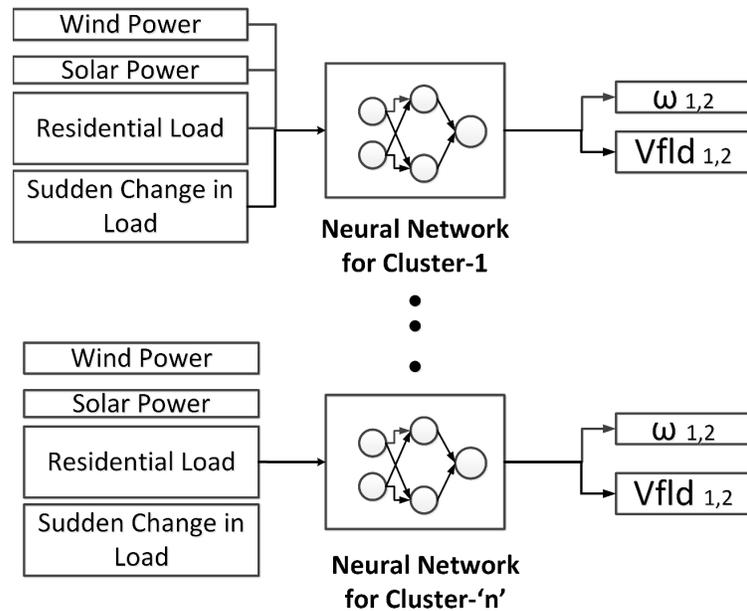


Figure 4.5: K-means Cluster based Neural Networks [9]

and a binary classifier was prepared. The result obtained from the classifier is shown in Figure-4.7. The tree classifies two scenarios; ‘Y’ represents unstable scenarios and ‘N’ represents the stable scenarios. The “ResLoad” stands for residential load.

In a parallel working stream, the data is also fed to an unsupervised clustering algorithm and five different data clusters have been prepared. The centroids of each cluster have been stored for labelling the newly arrived data. The data labelling is based on the previously prepared clusters. The centroids of each cluster are given in the Table-4.2.

Table 4.2: Centroids of the Clusters

Cluster	Wind(kW)	Solar(kW)	Load(kW)	Changes(kW)
1	191	150	198	42
2	190	200	202	36
3	204	50	200	39
4	200	100	202	44
5	198	0	199	44

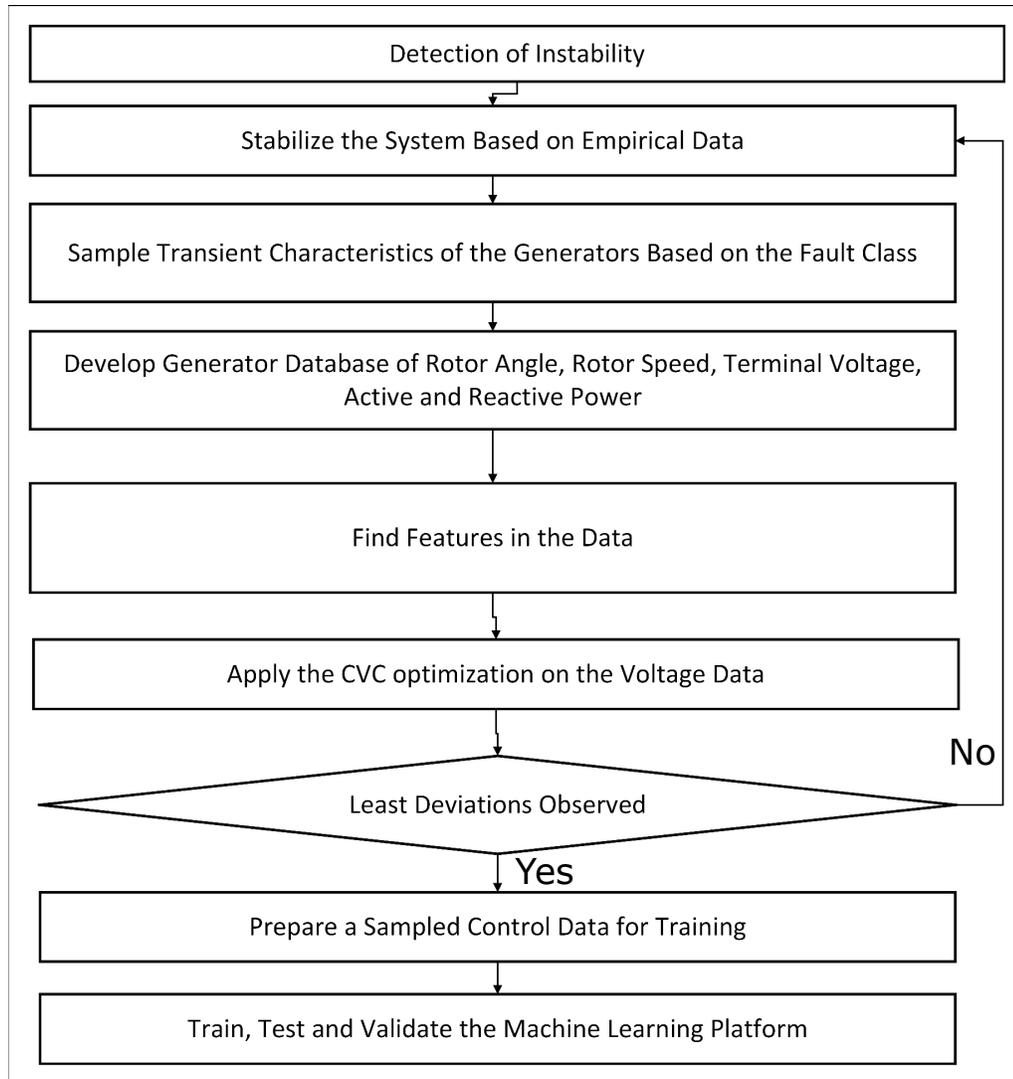


Figure 4.6: Work-Flow of the Proposed Algorithm

When a new system data has been collected, the decision tree-based classifier predicts a probable voltage instability in the system. Then based on the proposed clustering scheme an appropriate neural network was invoked and prepared for the imminent threat. If the system identifies voltage instability the secondary control scheme forces active and reactive power control on each of the

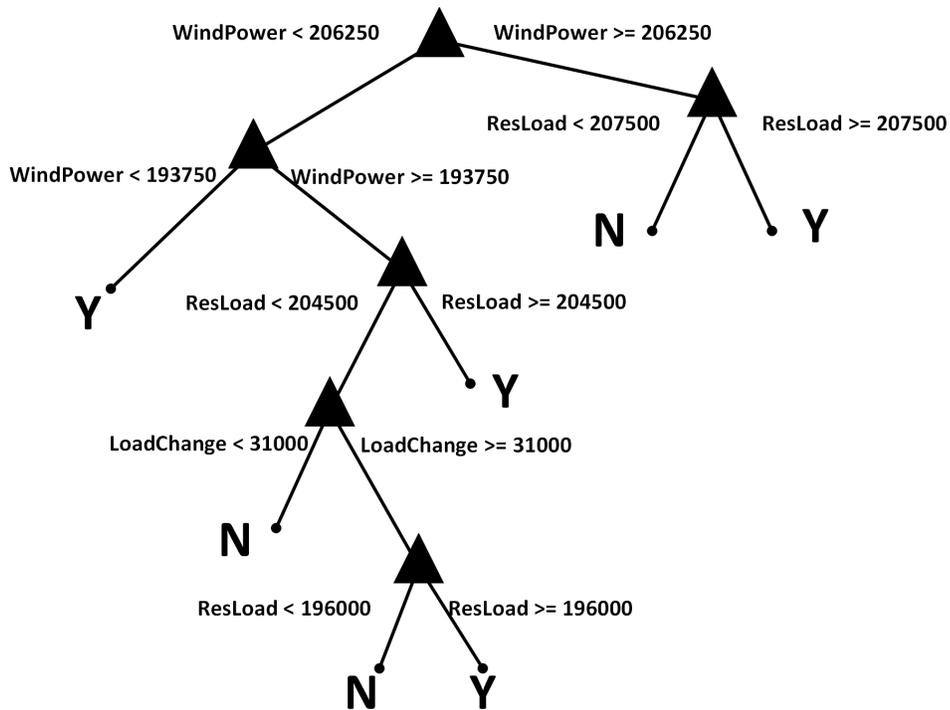


Figure 4.7: Binary Classification of the System Stability (Units in watts)

generators based on the predicted outcomes of the neural networks. The result of the two neural networks are shown in Table-4.3.

Finally, Figure-4.8 shows the performance of the proposed algorithm in detecting and subsiding the voltage instability with the data cluster number '4', and Figure-4.9 shows the performance with data cluster number '3'.

Figure-4.8 and Figure-4.9 have been divided into multiple segments for better visualization. The circled lines on the left show the point where the system observes a sudden unexpected change in load. As the initial control scheme is the linearized primary droop control, the immediate impact is a sudden rise in voltage. The voltage rise is followed by voltage instability that the primary controller cannot regulate due to the mismatch in active power in the system. This point is shown using

Table 4.3: Predictions Using Neural Networks

Active Power Reference: Generator-1				
Cluster	Actual	Neural Net	Error %	
4	0.8013	0.848	5.828	
3	0.87	0.849	2.41	
Field Voltage Reference: Generator-1				
Cluster	Actual	Neural Net	Error %	
4	1.00	1.054	5.4	
3	1.00	1.0075	0.75	
Active Power Reference: Generator-2				
Cluster	Actual	Neural Net	Error %	
4	0.8014	0.86	7.31	
3	0.87	0.8709	0.103	
Field Voltage Reference: Generator-2				
Cluster	Actual	Neural Net	Error %	
4	1.00	1.065	6.5	
3	1.00	1.007	0.7	

the 2<sup>nd</sup> line with cross marks. The oscillation starts to rise and crosses the permissible maximum  $\Delta V$  for the transmission line. At this moment, when the instability is spotted, the system forces the secondary control mechanism to take over the synchronous generator. This transition is marked using the squares. After the secondary control is applied the transmission voltage becomes momentarily stable and reaches towards a new equilibrium point. In Table-4.3 the performance of neural network in predicting reference values are shown. The table shows that the prediction has some errors in it. The observation also show that the predictability of the proposed method works better with the data cluster-3 than the data cluster-4. The results are coherent with Figure-4.8 and Figure-4.9, in explaining why a better response is observed in Figure-4.9. The reason for such differences can be traced back to the bin count of the clusters. In these fifty cases, cluster-3 has occurred fourteen times while the cluster four has occurred nine times. Such a difference could influence the training process of the neural network. Despite having these small errors, the proposed secondary controller damps the oscillation effectively. However, due to that minor prediction error, after the system becomes stable the transmission voltage tries to reach towards a new dynamic stability. This

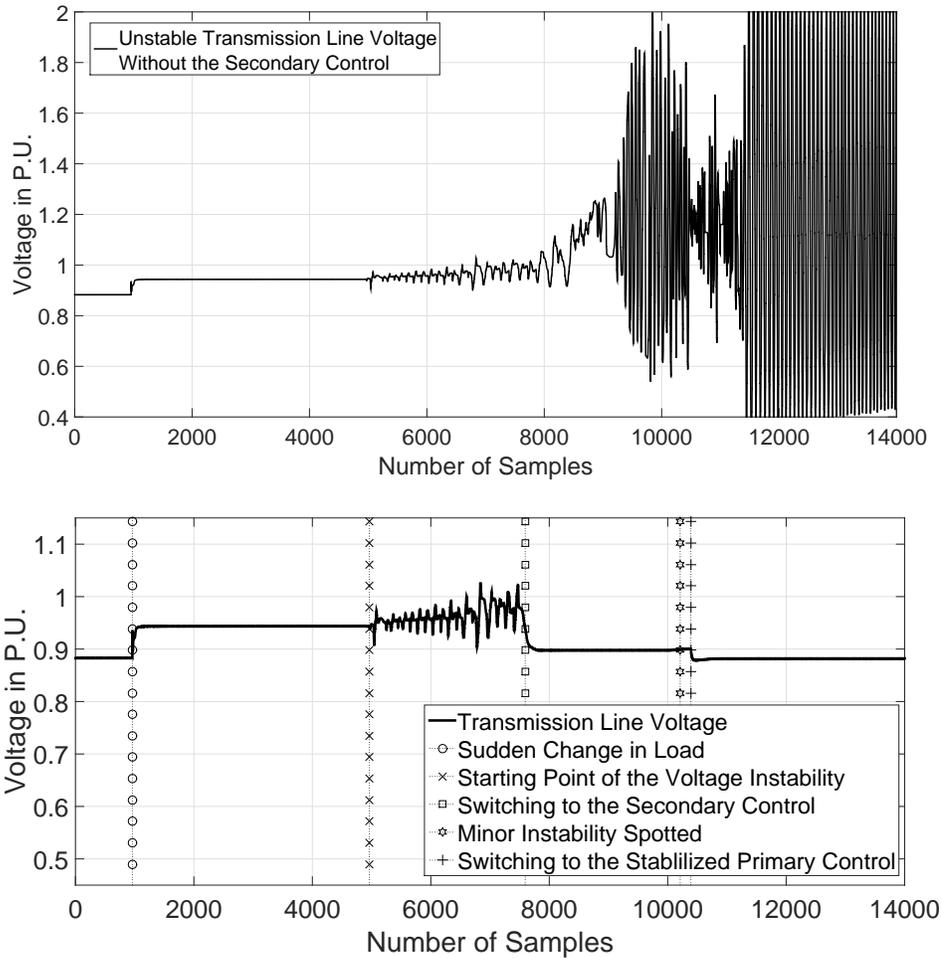


Figure 4.8: Performance of the Proposed Secondary Control Scheme (Cluster-4)

phenomenon is shown using the 4<sup>th</sup> line with pentagrams. At that point of new dynamic stability, the system is switched back to its primary control mechanism. This switching is shown using the 5<sup>th</sup> line with plus marks. The system remains stable after that point. The secondary controls applied in both the machines do not use any communication channel between them. The centralized control scheme for supervising the performance of the secondary control is only used when secondary control is incapable of maintaining stability and the centralized control would either shed additional load or switch to the second synchronous generator, having lower energy ratings off.

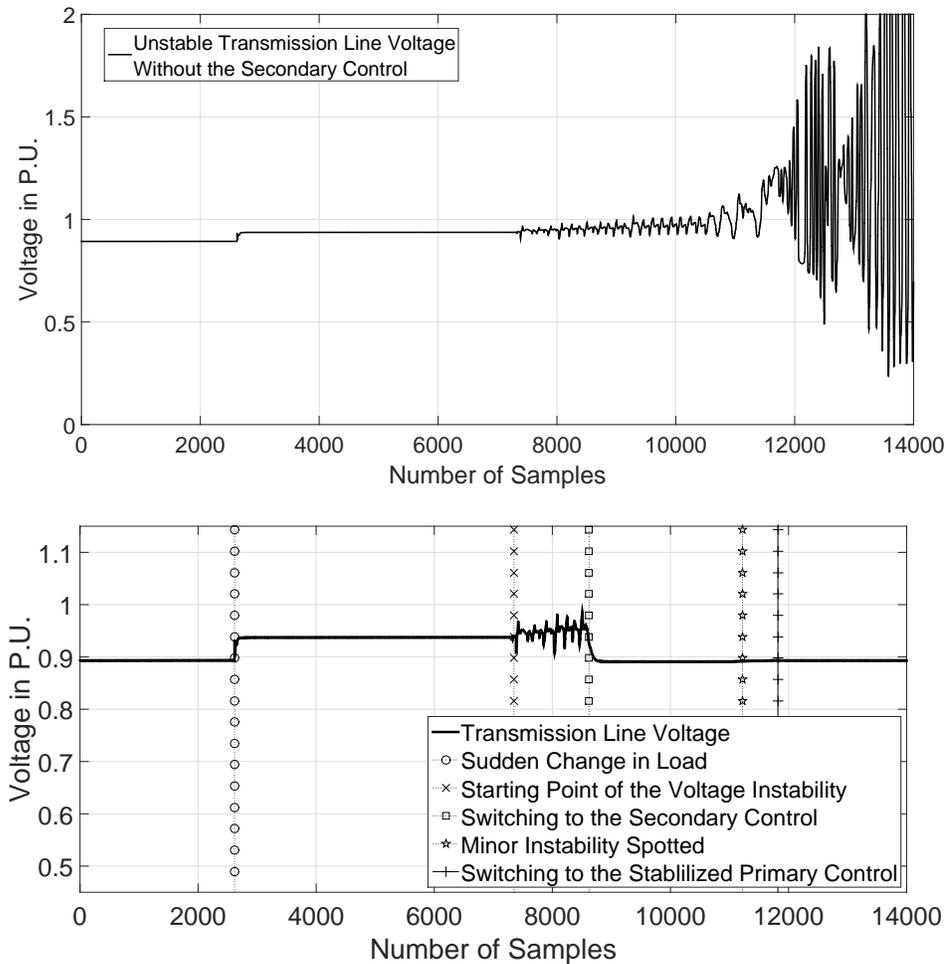


Figure 4.9: Performance of the Proposed Secondary Control Scheme (Cluster-3)

#### 4.2.4 Analyzing the Limitations

This study implements a cascaded method of applying machine learning algorithms in a distributed architecture, to provide secondary supervisory control on an isolated microgrid. The results obtained through this study are significant. However, there is a need to discuss the impact of misclassification on the system operation. It is also important to address the modification required for the distributed control system. On the other hand, the performance of the classifier, as well as the neural network fitting model, heavily depends on the number of use cases. In this study, only fifty

scenarios have been used near the region of stability margin in the PV-curve. The accuracy of both the binary classifier and neural network fitting models will improve if many scenarios are incorporated. Another critical observation is that the proposed model is quite simple. Therefore, only a limited number of complex scenarios could be generated. In a smart microgrid, the number and types of generators may vary. In those cases, the method mentioned above may not be proven as effective as it has performed so far. The following chapter analyses these scenarios and addresses the limitations by introducing a simplified feature selection method.

# Chapter 5

## Simplified Feature Selection for Dynamic Data

This chapter provides a solution for the problem of complex event detection. In multi-area power systems, data preparation plays a vital role. In this chapter, feature selection is used to address some of those challenges observed in the data preparation stage. This chapter has been published in the IEEE Transactions on Power Systems, 2018.

( M. Al Karim, J. Currie and T. T. Lie, “Dynamic Event Detection Using a Distributed Feature Selection based Machine Learning Approach in a Self Healing Microgrid,” in IEEE Transactions on Power Systems, vol. PP, no. 99, pp. 1-1. doi: 10.1109/TPWRS.2018.2812768 )

### 5.1 Addressing the Limitations with Feature Selection

The advent of Phasor Measurement Units (PMU) has improved Supervisory Control and Data Acquisition (SCADA) system in recent years. It has also enabled new possibilities in the field of protection and stability analysis, that eventually lead towards self-healing [97, 103]. Through a framework, autonomous systems can be deployed in an interconnected network to maintain system

reliability [18]. However, an end-to-end system configuration, which can deploy self-healing mechanisms, may require multiple layers of decision-making processes. These can be represented as a hierarchical, three-tier structure. The bottom layer is placed near each measurement device close to the physical system. The middle layer is intelligence at the substation level and the top layer is the centralized command and control operations undertaken by engineers [14]. The deployment of an immediate control scheme in the bottom layer is considered a well-established strategy to address and solve contingencies locally, without any intervention from the central station. The sole purpose of such an action is to reduce costs and increase speed of operation. The goals of reducing costs and resources have motivated recent authors to explore local decision-making methods [8–10, 104]. In reference [10] a novel sectionalized self-healing approach was introduced, featuring a sectionalized grid system to mitigate energy distribution problems. The objective set by [10] is to guarantee supply and demand balance in each subsection, either by adjusting the power output from the dispatchable sources or by shedding loads. The article proposed a rolling-horizon optimization method in order to schedule the outputs of the dispatchable distributed generators (DGs). The proposed method has significant merit; however, the article assumes that the self-healing process does not give rise to any dynamic stability issues, which often is not the case [100]. On the contrary, power system stability relies heavily on the dynamic behaviour of the generators. In reference [10] the authors argue that system instability can be assumed by observing the dynamic behaviour of the generators. In other words, the development of an automated algorithm should rely on the idea of low-level machine control.

Recent literature also suggests that local or distributed control is quite dependent on the identification of dynamic events under different contingencies. An appropriate detection of power system events can lead to smart restoration strategies [10, 97, 103, 105]. This study integrates the concepts of both [106] and [10] and introduces a novel algorithm that detects dynamic events from the distributed generator data in a sectionalized microgrid. The dynamic events are associated with

the self-healing process of the microgrid used in this study. The algorithm interprets the dynamic events and decomposes those into user-specified regions to facilitate decision making in the context of restoring an unstable power system. The proposed algorithm has the capacity to detect patterns in the dynamic data and distinguish the data based on the underlying events. Once the underlying event in terms of the affected generator is recognized, the algorithm can take local decisions on each of the generating stations and restore the system after a major fault. The algorithm is installed in each of the generating stations, which are independent of each other, but can make a coherent decision in a post-fault contingency. In the proposed system, two types of non-dispatchable energy sources is introduced; wind and solar. The proposed method is based on a machine learning algorithm which is a modified ensemble of bagged decision trees with an added *Boosting* mechanism. The algorithm uses the generator data collected from the aforementioned grid. For a better prediction, the generator data is augmented using a simplified feature selection process. The purpose of the feature selection process is to prepare the data set for increasing the accuracy of prediction [107–109]. However, some of the major concerns regarding feature selection are high sensitivity to tuning and added redundancy in the algorithmic steps [110]. For these reasons, additional algorithms for improving relevance and removing redundancy go hand in hand with any feature selection process [107]. Solving such challenges would require more resources and would be a time-consuming process to enable data preparation; although this is very important it is unnecessary to address it in a near-real-time application . This study considers an alternative approach, avoiding feature selection in real time by implementing a pre-processed set of input features [109]. The relevance and effectiveness of these features are discussed in the later sections. To develop, train and test the proposed system a two-area-based microgrid was prepared in Matlab-Simulink. The overall contribution of this study can be summarized as:

1. Developing a feature selection-based machine learning algorithm that detects dynamic events in a stand-alone and self-healing microgrid.

2. Exploring the potential of feature-based augmented datasets, to enhance the predictability of a multiclass classifier algorithm for dynamic time series data.

A similar test system as introduced in the earlier section is used. However, Like [106] the system is designed to have both a normal operating mode and a self-healing mode. The self-healing mode consists of series compensators (SCs) in both areas near the synchronous generators. The post-fault contingency is considered as a first swing stability problem, which can be damped using a linear continuous control. From each distributed controller's point of view the grid is a single machine infinite bus system; thus, the active and reactive power transmitted through the transmission line can be modelled as  $P_i = \frac{V^2 \sin \delta}{(1-k)X}$  and  $Q_i = \frac{2V^2}{X} \frac{k}{(1-k)^2} (1 - \cos \delta)$ . Here  $i$  is the distributed generator,  $X$  is the line inductance and  $k = \frac{X_c}{X}$ ;  $X_c$  = series capacitance. Once the proposed algorithm detects an instability it operates the SCs and damps the oscillation. For simplicity the compensators are considered to have two modes; *Bypass Mode* for damping the oscillations in a post-fault scenario and *Blocking Mode* for normal operations. Once a post-fault rotor angle instability is identified, the SC closest to the most affected generator is switched to *Bypass Mode*, to suppress the oscillation and stabilize the system. The SC is then returned to *Blocking Mode* after the system has been restored [111].

## 5.2 Proposed Methodology

The overall process is shown in Figure-5.1. It starts with a data preparation stage where a Monte Carlo-based simulation is carried out with four different fault locations close to the wind power plant, solar plant, hydro and diesel power plants. The process is followed by the preparation of a dynamic response database. The database contains the pre-fault, during fault and post-fault dynamic responses of each generator. The feature extraction stage follows, to extract some predefined features and augment the dataset for training machine learning algorithms placed near each of the

generators. Once the training process is over the algorithm is evaluated with a new set of randomly chosen datasets. This phase has three outputs. The first output is invoked when classification error is observed while training. This link adjusts the feature space. The second output ensures the system is properly trained and completes the training process. The third output contributes in the actual decision-making phase on right.

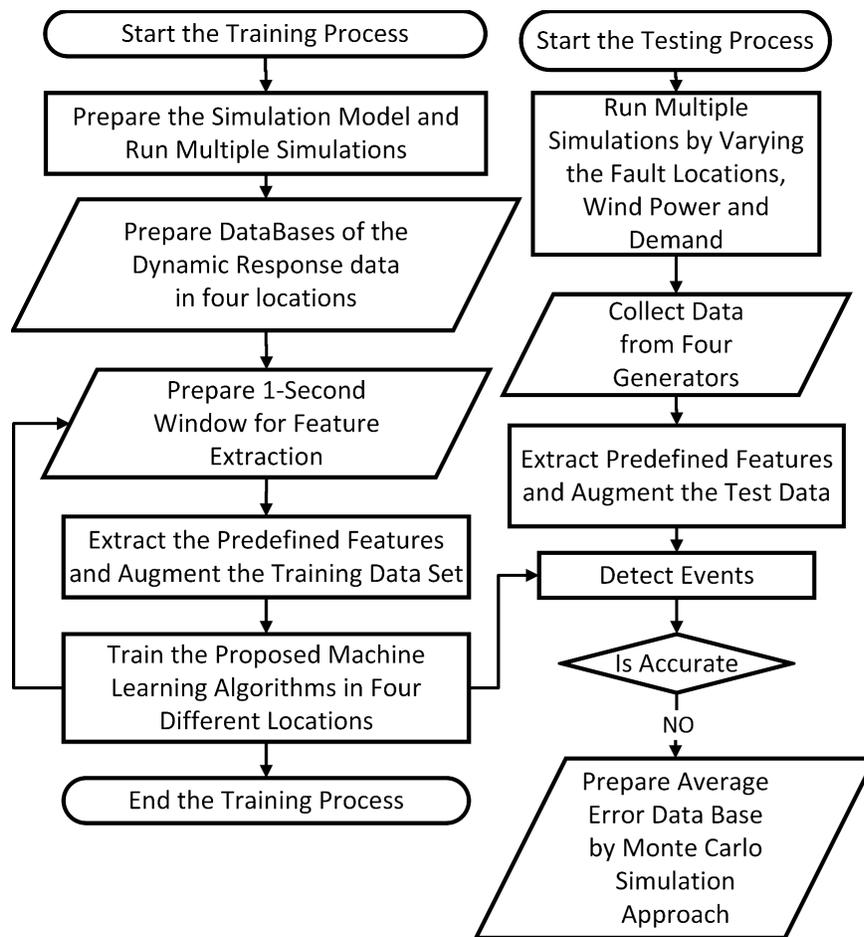


Figure 5.1: Workflow of The Proposed Algorithm

### 5.2.1 Data Preparation

The data preparation stage does the data selection and the feature selection shown in Figure-5.1. In order to develop the system dynamic response database from the generators, **250** simulations were carried out with randomly selected wind generation and demand for each of the four fault locations as shown in Figure-5.2 resulting in a total of **1000** different scenarios. For example, the random scenarios relevant to the wind power plant were created with a mean wind speed of  $10m/s$ , and a standard deviation of  $1m/s$ . This study only considers scenarios where rotor angle instability emerges after a major fault.

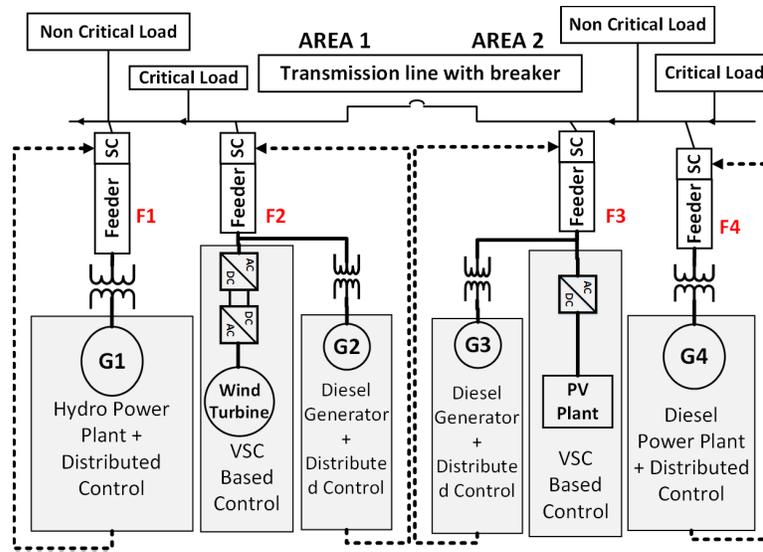


Figure 5.2: The Microgrid Model

=

The feature selection mechanism adopts the concept presented in reference [112]. The method implies that it should be agnostic to use machine learning for modelling; it should account for a streaming dataset; and it should be scalable to support many events. However, unlike [112], this study implements a fixed feature set to eliminate dependency on any wrapper (different models

for different subsets) or filter methodology for feature selection. As the system analysis is largely based on post-fault data, non-linear mapping functions, which are mostly required for load and price signals, can be avoided [110]. Thus, a predetermined feature set can be considered. As the time constants of MGs are relatively smaller than those of a high voltage system, generator parameters were carefully considered to select the appropriate features [113]. To produce the generator data a conventional differential algebraic equation-based model was used [114].

This study carried out multiple simulations based on the above principles by randomly varying the wind power and demand while considering four probable fault locations. The overall process of damping the rotor angle instability is mentioned in reference [8]. However, unlike [8], a supervised control scheme was not implemented; rather the study relied on primary control, secondary control and appropriate series compensation. For simplicity, only cases where oscillation was completely damped were considered as valid candidates for training the machine learning system. Analysis of the undamped scenarios was regarded as out of scope and not undertaken.

To establish a relationship between the generator model and the constant impedance load model, classical energy functions were used [109];

$$M_i \frac{d^2 \delta_i}{dt^2} + D_i \frac{d\delta_i}{dt} = P_i - \sum_{j=1, j \neq i}^m (C_{ij} \sin \delta_{ij} + D_{ij} \cos \delta_{ij}) \quad (5.1)$$

where,  $m$  is the number of synchronous generators.  $\delta_i$  is the rotor angle of the  $i$ -th generator.  $C_{ij}$  and  $D_{ij}$  are the function of transfer conductance and susceptance of the reduced network. The per unit inertia constant of each of these generators is  $H = \frac{1}{2}M\omega$ , where  $\omega$  is the synchronous speed.

The dynamic data produced from the generators was selected based on a multi-band power system stabilizer (PSS) model for the hydro power plant and a traditional PSS model for the diesel power plant. The parameters are rotor speed  $\omega$ , rotor angle deviation  $d\delta$ , reactive power generated  $Q_E$  and terminal voltage  $E_G$  [115, 116].

Parameters are represented in per unit . Those generators and consumers not close to the vicinity of the fault are kept constant during the period when a transition between two events is taking place.

## 5.2.2 System Dynamic Responses

The proposed algorithm implements a simplified feature selection technique for data preparation and applies multiple machine learning algorithms in order to detect underlying events from the time series data. The method is initialized by preparing a database of the system dynamic responses. Four attributes namely: generator rotor speed  $\omega$ , rotor angle deviation  $d\delta$ , reactive power generated  $Q_E$  and terminal voltage  $E_g$  were chosen for preparing the database [10, 117, 118]. The rotor angle deviation was measured against the center of inertia (COI) angle;

$$\delta_{COI} = \frac{\sum_{i=1}^n H_i \delta_i}{\sum_{i=1}^n H_i} \quad (5.2)$$

$\delta_i$  is the rotor angle and  $H_i$  is the inertia constant of the **i-th** generator.

The system dynamic response is characterized by different events taking place in the proposed two-area-based system. The events are divided into nine different categories. The categories or events are based on the transient, sub-transient and steady state behaviour of the pre-fault, during fault and post-fault system states. Two events are associated with the pre-fault state, one event identifies the fault, three events are associated with the post-fault state with classification errors and the remaining three events are associated with the post-fault state without classification errors. For each of these categories or events, generator data was collected from different power plants and stored as a matrix of time series databases. Considering  $t_n$  is the length of time and  $m$  is the number of generators, the attributes of the classification algorithm therefore become  $m$  data points ranging from  $(t_n, E_1, Q_1, \omega_1, d\delta_1)$ ,  $(t_n, E_2, Q_2, \omega_2, d\delta_2)$  to  $(t_n, E_m, Q_m, \omega_m, d\delta_m)$ . The duration of the fault was kept uniform. It was just long enough to introduce a rotor angle instability in the system. However, the length of post-contingent scenarios is not uniform; rather they were chosen

based on the time required to achieve system wide stability. This measure was taken to monitor whether a loss of synchronism appears before initiating the next event.

A power system is often interrupted by uncertain events such as three-phase faults. Once the fault is cleared, the system must be restored and operated at its optimal target configuration. The actual method of system restoration is a well-established research domain that consists of expert systems, mathematical programming and heuristics, as well as soft computing e.g., [105, 119–122]. The primary focus of this study is not to analyse different restoration mechanisms, but rather to accurately detect the underlying events and properly categorize them using an automated system. This study assumes a soft computing-based (machine learning-based) system restoration mechanism as discussed in later sections.

The nine events are briefly explained below. Figure-5.3 and 5.4 show six time-varying dynamic events out of those nine events. Each of these events has different values in the chosen features. These different values help linearize the attributes for the classification algorithm.

1. Starting the generators: In this state the transient and sub-transient phases of the generators during start-up are addressed.
2. Stable operating point after start up: This is the pre-fault stable operating point. Once the system reaches this state the experiment is considered ready for the introduction of faults.
3. Introduction of the fault: A large reactive power mismatch is caused by disconnecting a large inductive load. This state lasts long enough so that a rotor angle instability, and in turn, a high amount of power swing, can be introduced.
4. Fault clearance and post-fault transient state: In this state the fault is cleared and post-fault transience is observed. The SC is kept in bypass mode (damping mode) in the transmission line.
5. Post-fault stable operating point: The steady-state operating point after clearance of the fault.

6. Transition towards the initial stable operating point: Once the post-fault system becomes stable the operating conditions are reverted to the normal pre-fault operating condition. This event represents the transition from the post-fault stable operating point towards the initial restoration states. This state is similar to the sub-transient state of the machine initialization phase.
7. Restoration of the initial stable operating point: The stable operating point equivalent to event-2.
8. Fault clearance and post-fault transient state: In this state the fault is cleared and post-fault transience is observed. The SC is kept in blocking mode (undamped) in the transmission line. It is a post-fault scenario with misclassification.
9. Transition towards the post-fault stable operating point to reach the post-fault stable state or event-5. The SC is kept in blocking mode (undamped) in the transmission line. It is the other post-fault scenario, with misclassification. Event-7 follows both the *event-4* (properly classified decisions) and event-8 (misclassified decisions).

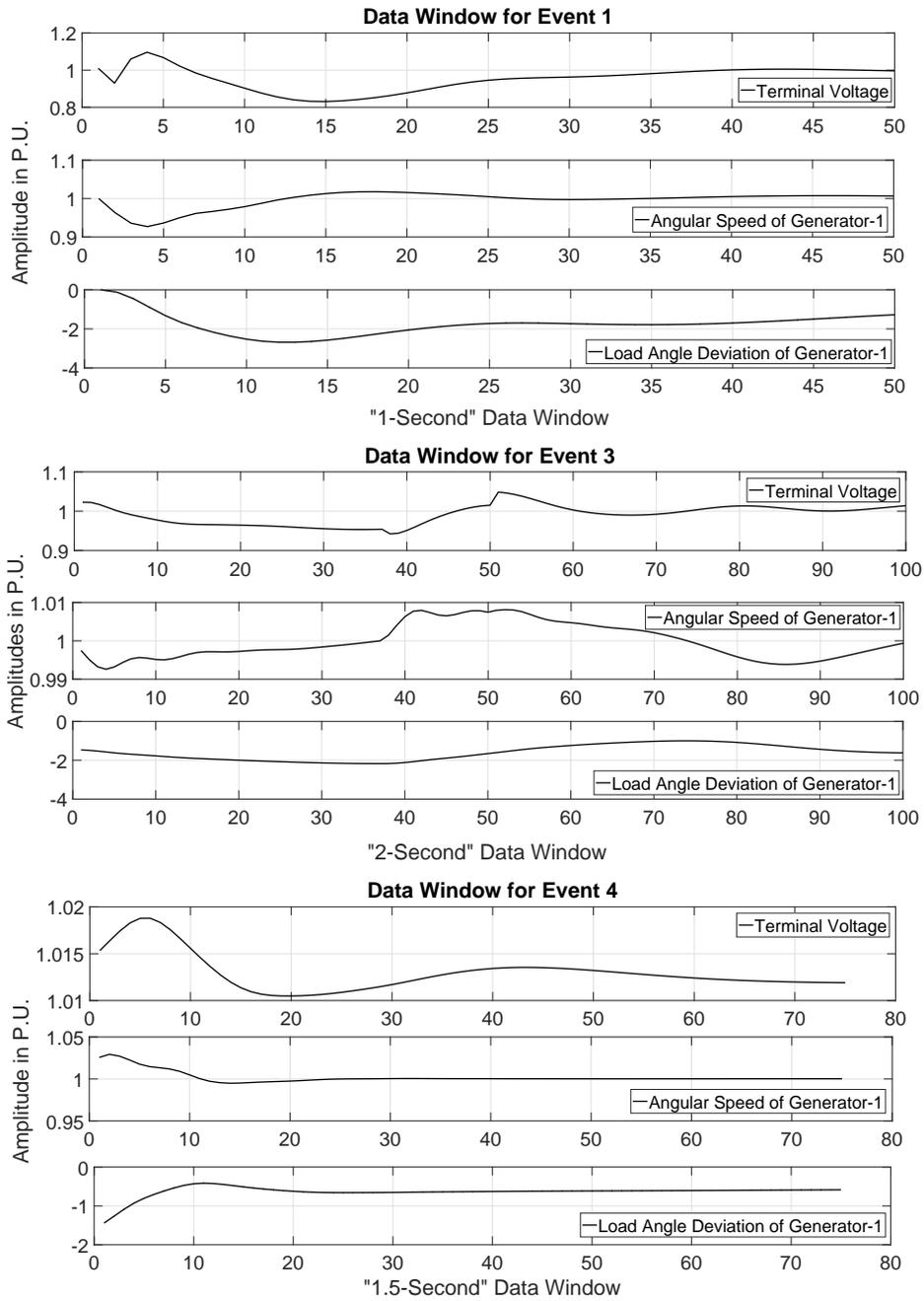


Figure 5.3: Generator Data Under Different Dynamic Events (50 samples/second)

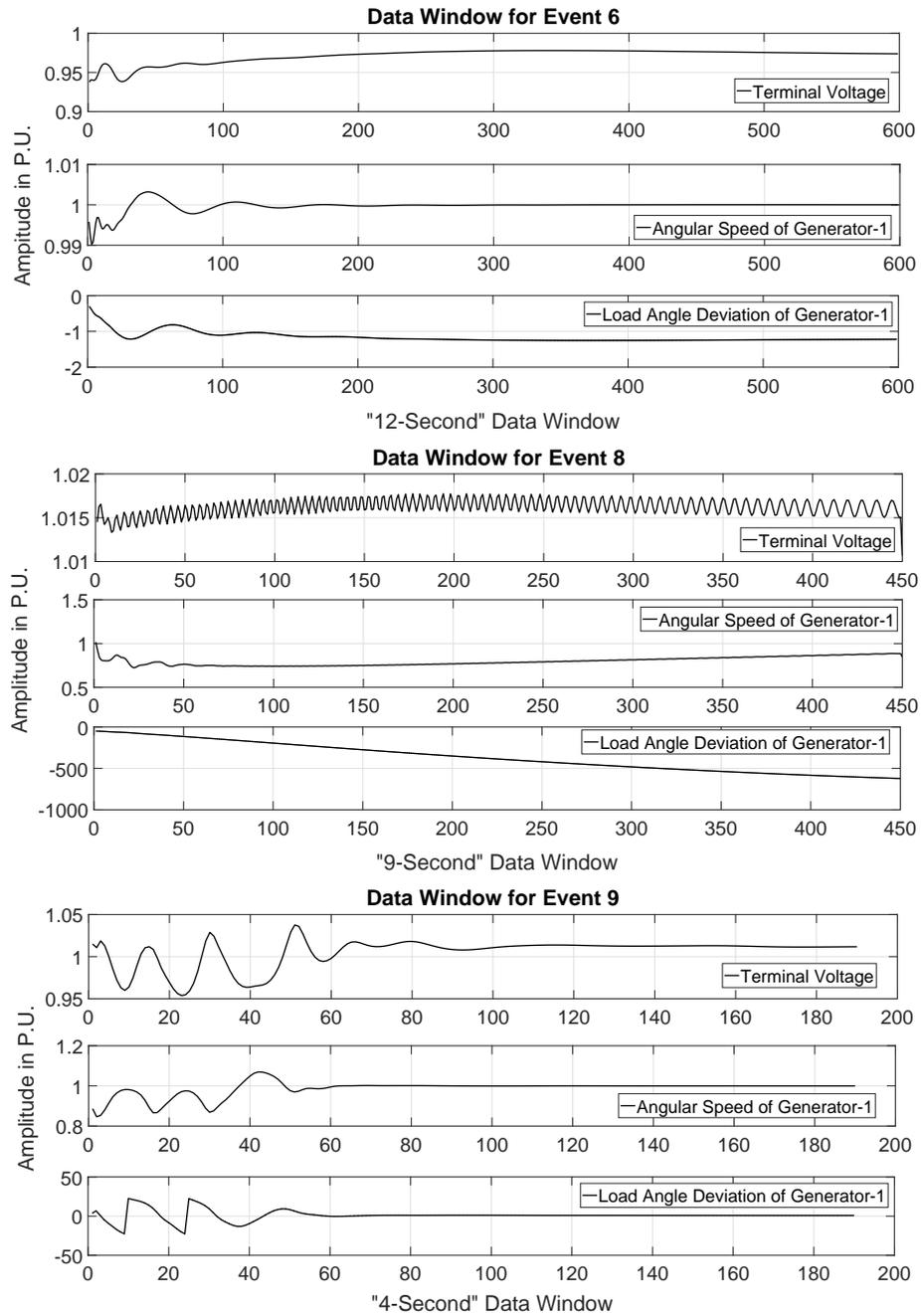


Figure 5.4: Generator Data Under Different Dynamic Events (50 samples/second)

The duration of each of the events may vary depending on the operating conditions of the microgrid. In general, most of the events are longer than one second. During any transition state the later parts become relatively more stable. Thus, identification of that half becomes comparatively less challenging due to the reduced variation in data, and also becomes less effective for identifying features. Therefore, to reduce the computational burden, only the attributes exhibiting certain variation in data, that exceeds a pre-defined threshold, are filtered and sent for feature selection. The rotor speed deviation  $d\omega$  is chosen for preparing this threshold. The crossing of the threshold value is observed for a period of 60ms. In other words there are three consecutive data points  $\Delta\omega_i = \omega_i - \omega_{i-1}; (i \geq 2)$ .

### 5.2.3 Selected Features

Harnessing information from magnitude and frequency of a signal, has been a well-established technique in understanding power quality. Therefore, different signal processing techniques were introduced to understand power quality issues in numerous studies [123]. Scientific analysis in other disciplines such as *Electroencephalography*, signal processing is often implemented to detect events such as analysing the shape of an action potential measured in millivolts [124]. The key argument observed in those studies is that signal shape contains valuable information. This study is also motivated by a similar argument.

After the attribute had been selected, a layer (in the form of attribute columns) of features was added, in order to augment the data to improve predictive capabilities [19, 125]. All the features are normalized to make the overall process generic for different wind power sources and demands. The augmentation is carried out based on three selected engineered features. These features have the potential to understand an underlying event, based on the shape of a time-series data. The features are:

1. Prominence of local maxima

2. The width of the local maxima
3. Available frequencies in the time series data

The characteristics of local maxima in a data set can carry valuable information [126, 127]. Time series data is composed of different frequencies and amplitudes. The characteristics of those amplitude peaks varies based on the event taking place at a certain period [128]. Dynamic phasor data, obtained from an MG, exhibit similar amplitude and frequency variations. This is the primary motivation for selecting the 1<sup>st</sup> and the 2<sup>nd</sup> feature. To select a range in the subject signal, a 1-second window was prepared. The window slides 20ms in each iteration, meaning that after every 20ms one data point is recorded. The justification for selecting a 1-second window is provided later.

As the initial selected feature in the data set, a topographical prominence was chosen. The method followed here is based on the work presented in reference [129]. The 1-second time series window is first normalized. Then the lowest contour line circulating a local maximum point is detected and the height of that point is measured in terms of that contour line. The second engineered feature is the length of the contour lines. It signifies the width of individual peaks. Figure-5.5 shows the method for finding prominent local maxima from a 1-second window in the terminal voltage of the generator in a hydro power plant (generator-1). The process applies Getis and Franklin's (G&F) variation of Ripley's K-function rather than a height threshold. The aim is to determine the number of localizations,  $n$ , in the 1-second data window or the region of interest (ROI). If  $K(r)_j$  is the Ripley's  $\mathbf{j}$ -th K-function and  $\delta_{ij}$  is the Euclidean distance between localization  $\mathbf{i}$  and localization  $\mathbf{j}$ , then;

$$K(r)_j = A \times \frac{\sum_{j=1}^n \delta_{ij}}{n} \quad (5.3)$$

Here,  $\mathbf{A}$  is the area of the ROI and  $\mathbf{r}$  is the radius around  $\frac{\sum_{i=1}^n \delta_{ij}}{n}$ , keeping the localization  $\mathbf{j}$  in

the center. It means  $K(r)_j$  represents the local signal amplitudes in relation to the average signal amplitude over the whole ROI, which in this case was the indicator of the peak prominence factor. The K-function is normalized across its variance and L-function, in preparation for data analysis. The diameter  $2 \times r$  signifies the width of the localization  $\mathbf{j}$ .

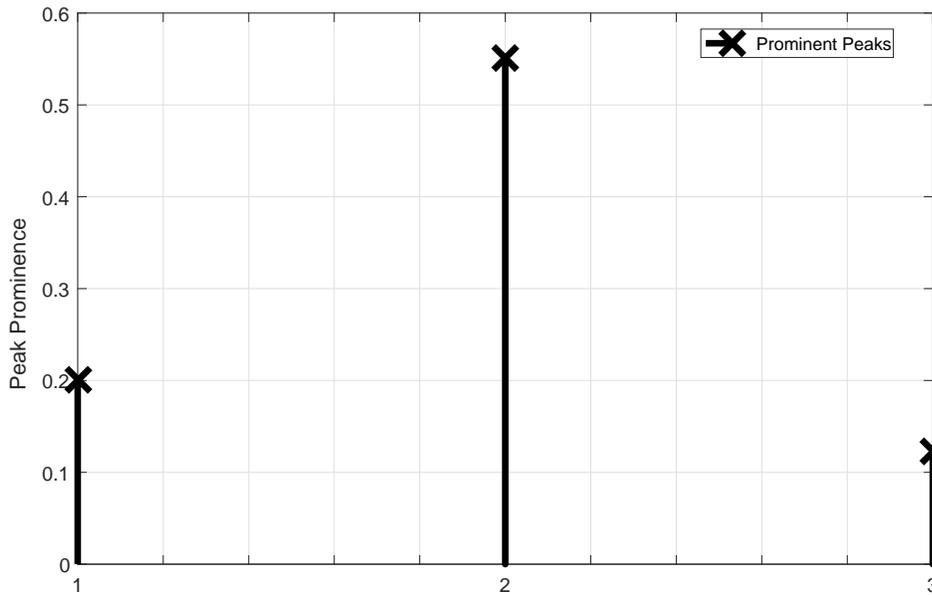
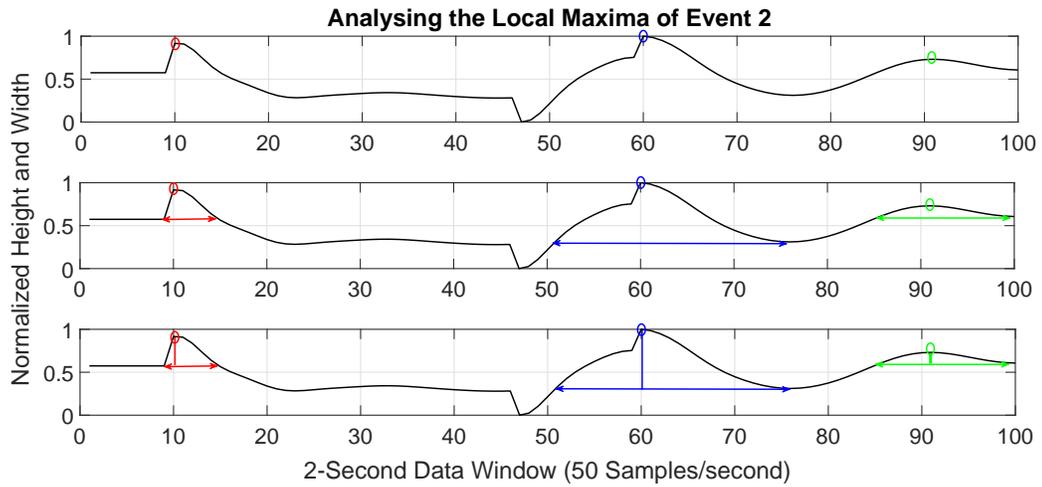
$$L(r)_j = \sqrt{\frac{K(r)_j}{\pi}} \quad (5.4)$$

One of the main ways of understanding power system oscillations is frequency spectrum analysis. Due to the spatio-temporal nature of power system dynamics, frequency spectra can develop valuable features [130], which can be helpful not only for understanding inter-area or intra-area oscillations but also for analysing voltage flicker, as has been done in past research [131–133]. The third chosen feature is therefore the available frequencies in the *1-second* data window. A discrete Fourier transformation (DFT) is used for this feature;

$$x[n] = \sum_{k=0}^{N-1} \frac{1}{N} \tilde{X}[k] e^{-jk(2\pi/N)/n}, K = 0, 1, \dots, N - 1 \quad (5.5)$$

Here,  $\tilde{X}[k]$  is the amplitude and  $x[n]$  is the linear combination of the complex exponentials with that amplitude. DFT has often been used in the field of harmonics and power quality analysis [132, 134, 135]. In this study, a frequency spectrum of the first  $20Hz$  was chosen to analyse the voltage fluctuation.

The commercial availability of phasor measurement units has made it easier to store power system attributes at a rate of 50 samples/second or higher. Here, a sampling rate of 50 samples/second was chosen, and thus, based on the “Nyquist–Shannon sampling theorem” the DFT could be carried out up to a frequency of  $25Hz$ . Therefore, *1-second* is sufficient to capture the frequency spectrum in one event. However, in the proposed microgrid it was observed that any frequency more than  $20Hz$  had little to no contribution in calculating the frequency factor. Therefore, the proposed fre-



Top Three Prominent Peaks of the Above Event (Event-2)

Figure 5.5: Peak Prominence and Width Inside a Predefined Data Window of the Terminal Voltage of Generator-1

quency feature is calculated based on frequencies up to  $20Hz$ . Based on the above argument, as mentioned in earlier sections, a window size of  $1-second$  was chosen. Another motivation for selecting a  $1-second$  window length was to retain enough distinction between different events. In the proposed microgrid, some of the observed events are no more than two seconds long. Therefore,

the frequency features in the transition windows (from unstable towards stable responses), are more likely to be corrupted. The top sub-figure (5.6a) of Figure-5.6 shows a comparison between three events under different window sizes and the bottom sub-figure (5.6b) shows a transition window between two events. From the figures, it is quite comprehensible that if the data window length increases beyond a certain point, the possibility of identifying distinct features decreases.

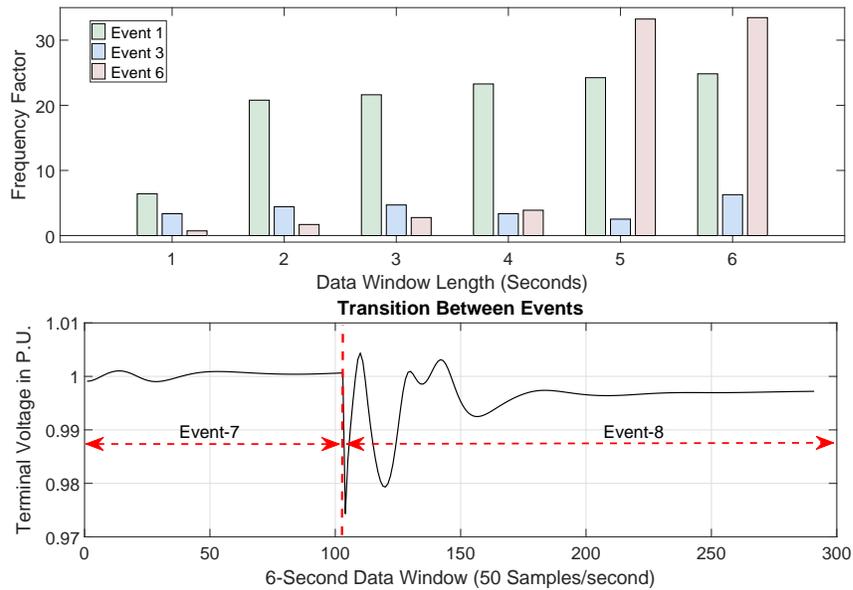


Figure 5.6: (a) Frequency Factor at Different Window Length and (b) Transition Between Two Events

To provide justification for those features, a randomly selected sample data set with three different events was prepared as shown in Figure-5.7. The chosen attribute was the terminal voltage of generator 1. The figure consists of the samples collected from three events *Event1*, *Event3* and *Event9*. Different events, despite having similar overlapping magnitudes, exhibit different characteristics in their signal peaks, widths and available discrete frequencies. This justification is reflected in Table-5.1.

For a better understanding of  $K(r)$  and  $L(r)$  parameters, the aforementioned sample data set as shown in Figure-5.7 was used. To show the calculation process for peak prominence and width, the *Event1* marked with a yellow line was considered. Firstly, from any local minimum  $\mathbf{i}$  to the

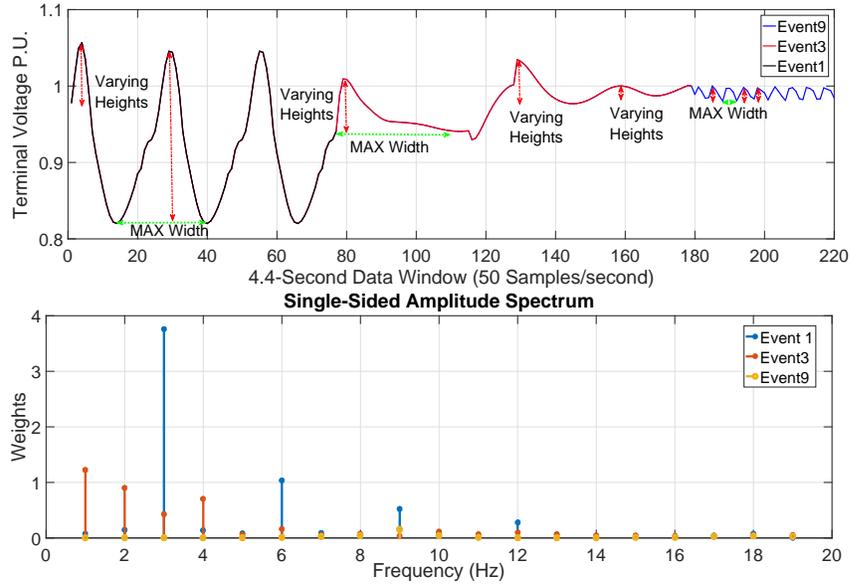


Figure 5.7: Sample Data and Feature Selection

next local minimum  $i+1$  a total of  $n = 40$  samples were selected. The distance **26.002** observed in the  $X$ -axis of the data window, between  $i$  and  $i + 1$  is considered as the diameter of the peak. Area  $A$ , **530.92** is then calculated based on this diameter. The average height  $\frac{\sum_{j=1}^n \delta_j}{n}$  of the  $i$ -th peak is calculated by summing the height of each sample  $0.82, 0.83, \dots 1.044, \dots 0.82$  and dividing it by  $n = 40$ , which is **0.4765**. Once the area and average height are found,  $K(r) = \mathbf{253.90}$  and  $L(r) = \mathbf{8.99}$  are calculated using the above-mentioned formula. Width is calculated from the diameter **26.002**. Table-5.1 was prepared using the above interpretation of the dataset. Figure-5.10 shows this process of including feature selection to aid the classification algorithm. A similar data set to that shown in Figure-5.7 can be used to calculate the frequency features by adding all the weights of the first  $20Hz$  frequencies in one data window. For example, inside one arbitrarily selected data window of *event1*, the twenty weights observed were  $0.072, 0.1446, 3.7611, 0.1407, 0.0843, 1.0338, 0.0901, 0.0365, 0.5217, \dots, 0.072, 0.0060$ . Adding those, a feature value of **6.8565** was obtained. While, The feature value with the first  $25Hz$  frequencies is **6.8567**.

As mentioned earlier, while extracting the features, each data window is sampled with 50 data

points, representing one measurement every  $20ms$ . However, to train a machine learning algorithm to identify the underlying event, each of the windows must be represented as a single row of attributes at any given time. The transition of the data window can thus be considered as a transition of events either from event  $i$  to event  $i$  or from event  $i$  to event  $j$ . This action is carried out in three different ways. For the first feature, a mean value of all the available peak prominences in a data window is selected; for the second feature, the maximum width of any given time frame is chosen; and for the third feature, accumulated weight obtained from the DFT of the first twenty frequencies is considered. Thus, a range of 50 data points is converted into one data point identifying the feature characteristics of that window.

The intention of the feature selection and data augmentation is to provide additional information to the classification algorithm. In this way significant differences can be seen between two time series data sets that have a high degree of similarity. For example parts of *Event-3*, *Event-6* and *Event-9* are shown in Figure-5.8. Along with that, the parts of comparatively more stable events, *Event-2*, *Event-5* and *Event-7* are also shown. The similarity between these parts is very high for proper classification by using the traditional classification methods. Besides, with varying wind power, solar power and demand these similar events overlap each other, which makes it difficult for a machine learning algorithm to gain adequate information.

However, adding the feature space to the original data set brings new information to the classifier algorithm, as shown in Table-5.1. The addition of such information helps to reduce classification error and improves the performance of the classification algorithms. This hypothesis is further analyzed, tested and explained in detail in a later section *Testing the Overall Hypothesis*.

## 5.2.4 Multiclass Classifier

Power system event detection is a multiclass classification problem, where multiclass classifiers are frequently implemented. A multiclass classifier can be an algorithm or a set of algorithms that

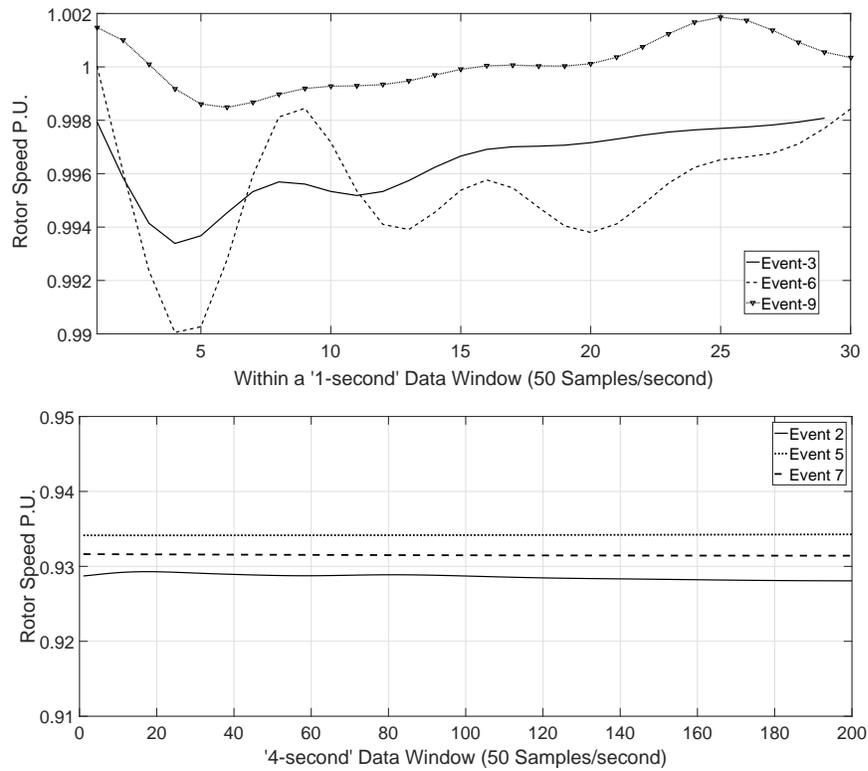


Figure 5.8: Similarity among multiple events

can predict more than two data classes. To solve such problems, researchers often use ensemble decision trees, sometimes also referred to as the random forest. A random forest algorithm uses multiple decision trees and regression techniques, on multiple sub-sets of a training dataset. In this study, once the data preparation stage was over, a multiclass classifier, which is an ensemble of bagged (bootstrap aggregation) decision trees, was trained as shown in Figure-5.10. Bootstrapping divides the training dataset into  $n$ -number of smaller sets with sample replacement, to train  $n$ -number of base classifiers. After training, each of those base classifiers vote for a class. The weights of all the votes are then considered in developing one improved composite model, which significantly increases the accuracy of classification [10, 125].

The key advantage of using bootstrap aggregated or bagged classifier is that it enables a linear combination of the function estimates. Based on the different sets of predictor variables and their

Table 5.1: Features In Three Similar Events

Event	Width Factor	Prominence Factor	Frequency Factor
1	26.002	8.93	6.8565
3	33.47	0.94	4.0348
6	47.83	0.042	0.6782
9	4.59	0.001	0.946
2	56.43	0.00035	0.008
5	77.02	0.00001	0.00197
7	109.52	0.00001	0.00111

estimates, the input data is weighted and re-weighted. For example; a function estimate  $\hat{g}_{ens} = \sum_{k=1}^M c_k \hat{g}_k(\cdot)$  is obtained based on the  $k$ -th re-weighted data with a combined linear estimation coefficient  $c_k$ . This method is helpful in addressing the issue of classification error due to estimation variance and statistical bias, especially in a stochastic scenario [136]. Bagged decision trees are, therefore, not pruned and are longer, which helps the ensemble method to improve accuracy by combining several low-biased sub-models. For example, in reference [133] an ensemble of decision trees was used, on PMU-based post-fault rotor angle data, to predict catastrophe in a wide-area power system. The implementation of ensemble trees improved the accuracy of the prediction by more than 20%, compared to a single decision tree.

However, the accuracy of classification or prediction of such an approach, depends heavily on the number of trees. Figure-5.9 is shown as a proof of concept, where a comparison of mean classification errors with different numbers of trees is made. **100** test cases were used for this proof of concept. The fault was placed near the hydro power plant and the data of *Event-4* was used as the target attribute. Once the use case was prepared, the classification error of the algorithm was evaluated, and the moving average of the error was plotted against the number of trees. The analysis shows that with an increased number of trees, mean classification error decreases. It was also observed that if more than 100 decision trees are chosen for the algorithm, accuracy reaches saturation point.

Therefore, based on the above analyses the ensemble was prepared using 100 fully grown trees

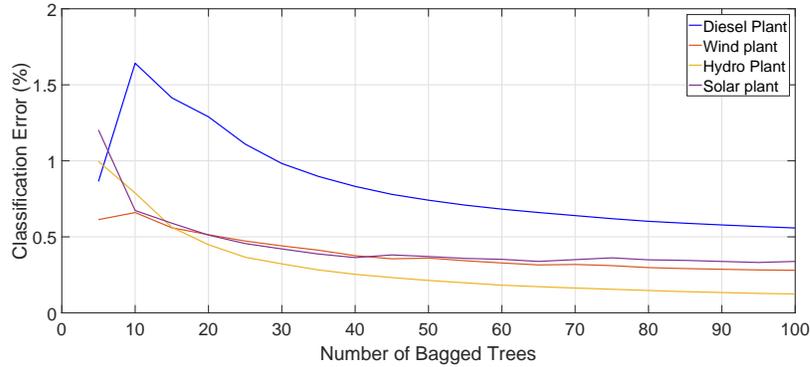


Figure 5.9: Classification Error Vs Number of Trees

by splitting the attribute data into 100 training sets,  $D_1, D_2, \dots, D_{100}$  in order to obtain an improved composite model.  $M_i (1 \leq i \leq 100)$  classifiers vote by predicting a class and the ensemble selects the final class from those votes. The overall technique includes *Bagging* and *Boosting*. The *Boosting* process works in a sequence by considering previously generated classification errors and rectifying those [10, 19, 107, 125]. The classifier is implemented on each of the four synchronous generator data sets. Different contingencies have different durations and thus different dynamic signatures. However, some degree of similarity is present in the dynamic events presented in this study, which affects the process of identifying classes. That is why, unlike [10] and [125], in this study a new scheme of multiple ensemble trees was used to significantly improve the scalability of decision trees.

In this method, using the four attributes of rotor speed  $\omega$ ; rotor angle deviation  $d\delta$ ; reactive power generated  $Q_E$ ; and terminal voltage  $E_g$ , and three features for each of the attributes, three different ensembles of bagged decision trees are trained. Each ensemble then is used to predict the events and an array of predicted events is prepared. The data is then analysed using a simple statistical mode operation and the most frequent prediction in each instance is considered as the final predicted event. This approach can reduce error while predicting the underlying event. The idea of using the statistical mode is to strengthen the estimation process by incorporating different

dynamic responses observed in a synchronous generator. The results are shown in the following section. Figure-5.10 shows the overall process of preparing the ensemble tree [11]. Hypothesis testing, with a limited dataset, is discussed in the following section in order to test the proposed method.

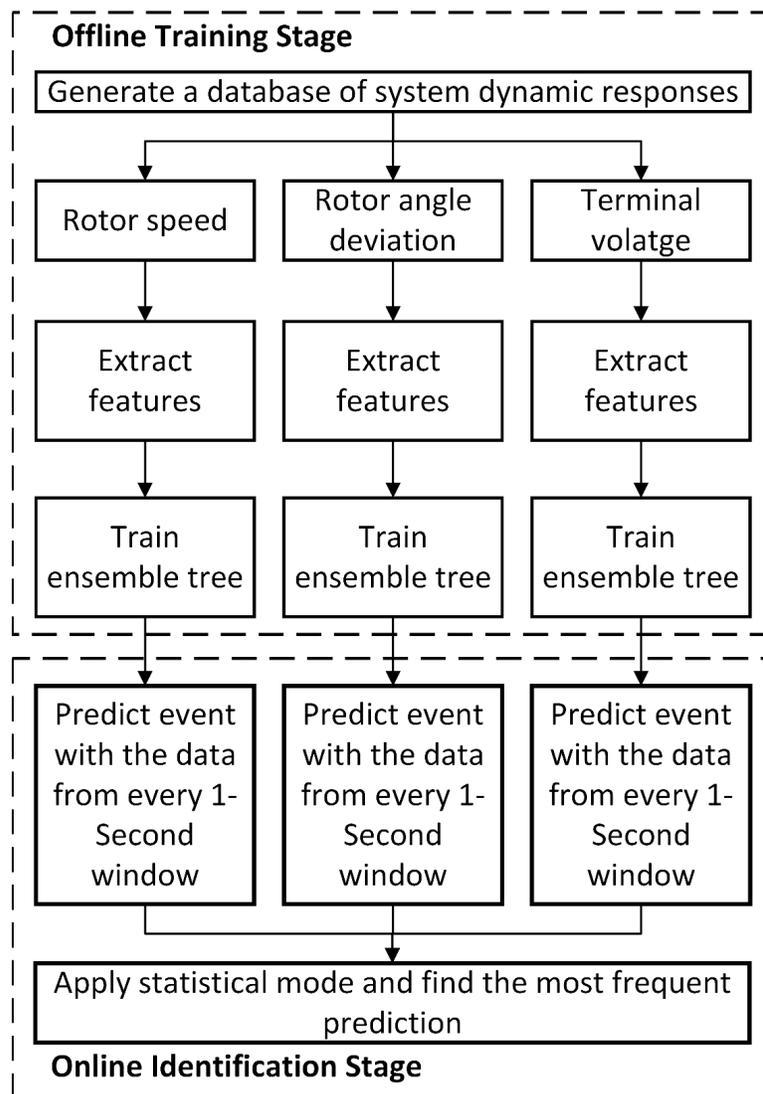


Figure 5.10: Preparation of the Multi-class Ensemble Trees [10, 11]

### 5.2.5 Testing the Overall Hypothesis

The overall hypothesis in this study is that different data types, with their inherent features, can significantly differentiate these events. The proposed algorithm suffers most during the transition periods from dynamic state to steady state, because the features become less distinct. However, if different data types are used to prepare the features, the accuracy of the ensemble method increases. The idea of using a statistical mode is to filter out the classification errors in one data type by using the results obtained from the other data types.

For example, rotor speed, terminal voltage and reactive power are dynamic in nature and each of them produces similarities between different events at certain durations. However, all of these are not similar at the same data points. Thus, if the classifier, using rotor speed, introduces a classification error, the other classifiers can mitigate that because of their own features. Figure-5.7 shows the data sample chosen from the terminal voltage of one of the diesel generators for testing the hypothesis.

The hypothesis also advocates for feature extraction to improve the performance of the classification algorithm. The proposed  $m \times 3$  feature matrix introduced in the earlier section is considered individually irrelevant. The decision tree algorithm recursively tests the attributes and partitions the dataset  $D_t$  containing scalar values into two or more child nodes  $D_{t1}, D_{t2}$ . Here  $t$  is used to represent a continuous data set representing one window. Each child node  $D_{ti}$  is defined by a range query such as  $threshold_i \leq D_{ti} \leq threshold_{i+1}$ . However, this query range cannot give an appropriate split if the time series data available shares similar values for different classes. Figure-5.7 is such a time series data set. However, adding the predefined distinct features that represent an individual class helps to eliminate this condition. A fictitious example is shown in Figure-5.11 to elaborate the concept. If feature extraction is used the algorithm reaches the end node (representing only one class) with higher accuracy in fewer steps. On the other hand, without the use of features, the algorithm reaches the end node with a larger number of steps. The classification error is also

high without the use of features. In the figure, class shapes have been used as additional information along with the position of each data point.

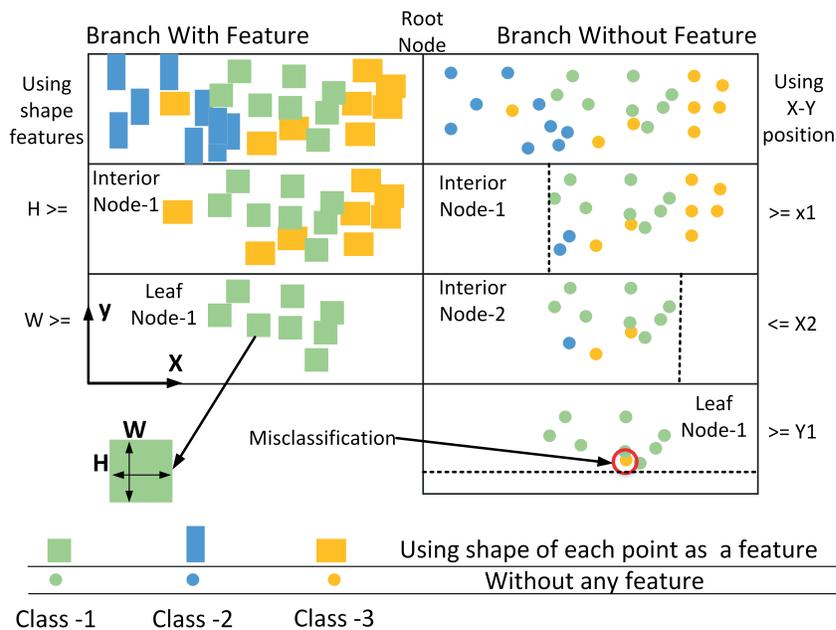


Figure 5.11: Classification With and Without Features (fictitious Data)

To understand the rationale behind the proposed classification approach, a concept called *decision node impurity* is invoked. When the decision tree splits a data set, the node impurity indicates how well the classes are separated. If, in a particular node, all the data points belong to one class without any classification error, then the node impurity is **0** or the node is considered pure. In Figure-5.11 the algorithm with feature augmented data reaches a pure end-node, while without feature augmentation it reaches an impure end-node.

In Figure-5.12 two instances of classification processes with the same data set are presented. The upper tree is based on the time series data only, and the lower tree is based on the augmented features partly observed in Table-5.1. To understand how the tree with feature data is performing better, a determination of goodness of split is carried out.

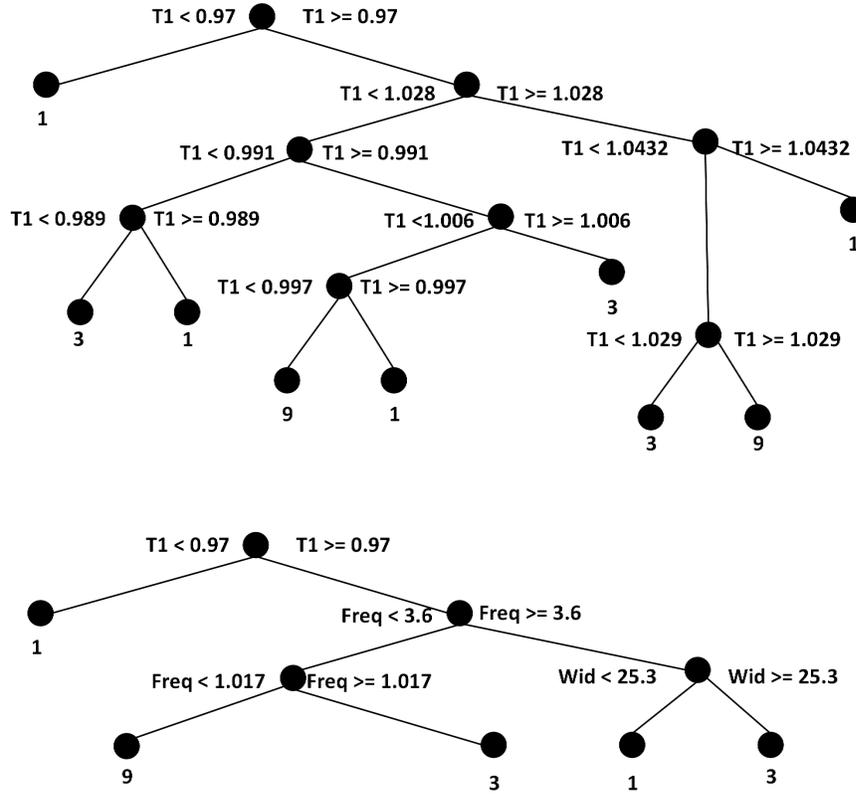


Figure 5.12: Classification Without and With Features

$$\Delta = I_{Parent} - \sum_{j=1}^k \frac{N(v_j)}{N} I(v_j) \quad (5.6)$$

where  $\Delta$  is the information gain;  $I(parent)$  is the impurity of the parent nodes;  $N$  is the number of parent nodes;  $k$  is the number of attribute values; and  $N(v_j)$  is the number of child nodes. The evaluation process measures impurity in order to understand the information gain. The information gain of classifying the first event (*Event1*) was 7.283% for the process without the features and around 26.87% with features. It shows that adding meaningful features reduces impurity. The comparison is done based on the first three interior nodes. The interior nodes are also known as non-leaf nodes, where a data set is divided based on a threshold observed in an attribute. The

impurity was then calculated based the equation shown below for each of the parent and child nodes [136–138]. In the equation,  $N_i$  is the node impurity,  $C_{target}$  is the number of samples representing the target class and  $C_{total}$  is the total number of samples those are available in the child node.

$$N_i = \frac{C_{total} - C_{target}}{C_{total}} \quad (5.7)$$

For example, if in a child node the observed  $C_{target}$  is 50 samples of *event-3* and in the same node  $C_{total}$  is 100 samples of *event-3* and *event-9*, then the node impurity is 50%

### 5.3 Performance of the Feature Augmentation Based System

The proposed method was then examined and compared using a test data base. The test database was prepared by randomly distributing the events. The test set contains  $400 \times 9$  events scattered over the time frame. To test the robustness of the algorithm, a normal random noise was also added to the attribute data. The random noise had a mean of **0** and standard deviation of **0.01** P.U. Table-5.2 shows the accuracy of the algorithm in detecting the dynamic events in each of the four generating stations. Table-5.2 also shows a comparison between the performance of the algorithm and three other traditional machine learning-based methods: a classification and regression tree (CART); an artificial neural network (ANN); and a K-nearest neighbour algorithm. The result shows a significant improvement can be achieved if a feature vector is introduced to the time series data that can add a distinct attribute characterizing each of the underlying events. Classification errors mostly occur during the transition between one event and another event, especially when the system is achieving stability. This is because when a system is near a stable state, features become less prominent and the margin of error increases.

The comparison clearly demonstrates the superiority of a feature selection-based method over traditional classification methods. Some of the classical techniques are very accurate and promis-

Table 5.2: Accuracy for *Event1*, *Event3* and *Event9* in Percent %

	Generator	Proposed	CART	ANN	K-NN
Event-1	1	99.8949	89.4529	< 30	93.2276
	2	99.8603	89.7265	< 30	93.5394
	3	98.9123	89.3287	< 30	90.6113
	4	95.7268	92.4601	< 30	95.4763
Event-3	Generator	Proposed	CART	ANN	K-NN
	1	99.6636	87.3372	< 30	90.0813
	2	99.6765	87.52398	< 30	91.5239
	3	99.6042	88.4577	< 30	92.1306
Event-9	Generator	Proposed	CART	ANN	K-NN
	1	97.1287	81.0249	< 30	89.8023
	2	97.1146	84.2273	< 30	91.2254
	3	97.5543	79.4611	< 30	93.1175
	4	95.2584	73.7433	< 30	94.0028

ing. However, as a stand-alone microgrid is a highly sensitive network, a small classification error in a post-fault scenario can have significant consequences. Therefore, it is desirable to implement an algorithm with improved accuracy.

In Figure-5.13 the overall accuracy in detecting all the **nine** events of the proposed distributed method are determined based on **four** different fault locations. In total **400** simulations were carried out to test the algorithm; **100** simulations for each of the fault locations. In each simulation the wind power and demand were randomly varied within a predefined boundary to develop different test data. The *black*, *blue*, *green* and *red* lines represent the accuracy of the algorithm in the diesel power plant, the solar power plant, the wind plant and the hydro power plant respectively. While the *Dotted ...*, the *Dashed - -*, the *Dot-Dashed -.-* and the *Solid* lines represent faults near the solar power plant, the diesel power plant, the hydro power plant and the wind power plant respectively.

The key application of the proposed method is likely to be in the context of identifying the generator affected most by a fault and making valuable decisions in order to clear the impact and restore the grid.

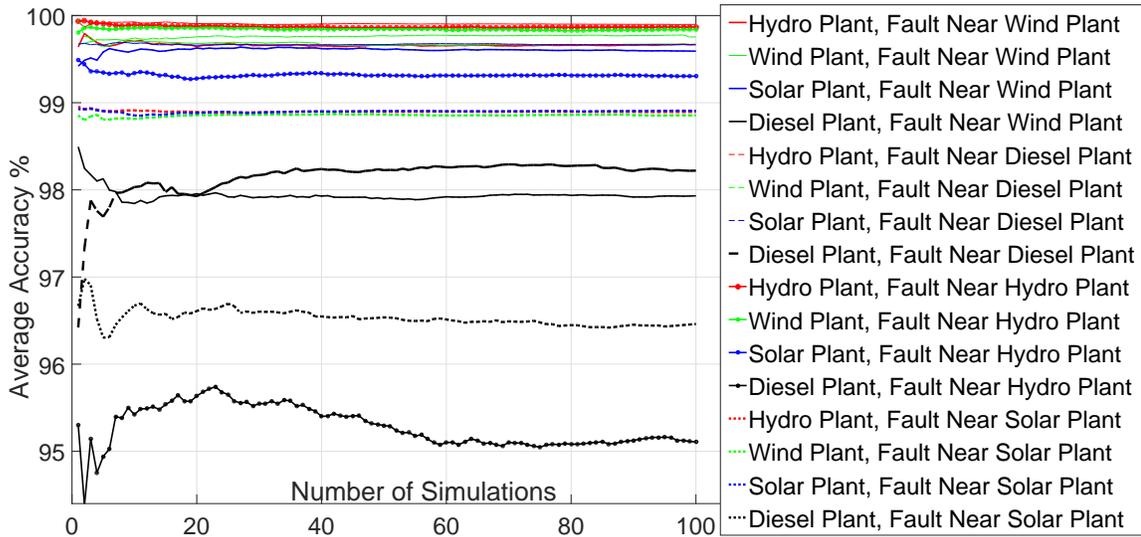


Figure 5.13: Accuracy Chart After 400 Simulations

Based on the determination of the underlying event, the distributed controllers operate the SC located closest to the faulty generator. This means the decision is to identify which SC close to the subject generators should be in bypass mode. The other SCs will remain in the blocking or non-operational mode. Table-5.3 represents all the available features for a *1-second* data window during the fault period. The fault location was altered from close to generator-1 to generator-4 and four scenarios were created. The detection of each underlying event is presented in the table. The algorithms do not share data between themselves. The application shown in column 6 shows whether an SC should be operated in bypass mode or not in order to damp the oscillation. **Yes** means operate the closest SC and **No** means do not operate that SC. Each algorithm can only control the SC located closest to it. Once the algorithm recognizes the transition from **Event3** towards **Event4** the algorithm switches back the SCs into non-operational mode. This switching of the SCs based on the identification of different events is considered an *online* operation.

Figure-5.14 shows a simulated event where a decision error results from a classification error. The decision made with a classification error is then compared with a decision without any classification error. In this event the algorithm placed closer to the Generator-1 fails to identify the location

Table 5.3: Online Decision Making based on the feature data collected from different locations (Observed from different generators)

Observed From	Fault at	Prominence Factors	Width Factors	Frequency Factors	Decision
G1	G4	0.0315	38.878	12.1	No
G1	G3	0.00166	34.72	1.2	No
G1	G2	0.038	37.57	12.4	No
G1	G1	10.64	27.183	49.98	Yes
G2	G4	0.035	21.51	15.2	No
G2	G3	0.00135	22.35	2.7	No
G2	G2	8.66	19.57	37.2	Yes
G2	G1	0.044	21.21	22.3	No
G3	G4	0.038	27.5	8.66	No
G3	G3	7.4	19.88	4.74	Yes
G3	G2	0.002	24.07	5.76	No
G3	G1	0.018	24.3	9.8	No
G4	G4	9.02	26.1	35.46	Yes
G4	G3	0.0002	34.8	9.12	No
G4	G2	0.017	34.75	10.03	No
G4	G1	0.025	35.44	10.82	No

of the fault and does not operate its SC but Generator-4 operates the SC near it, thus affecting the direction of power flow.

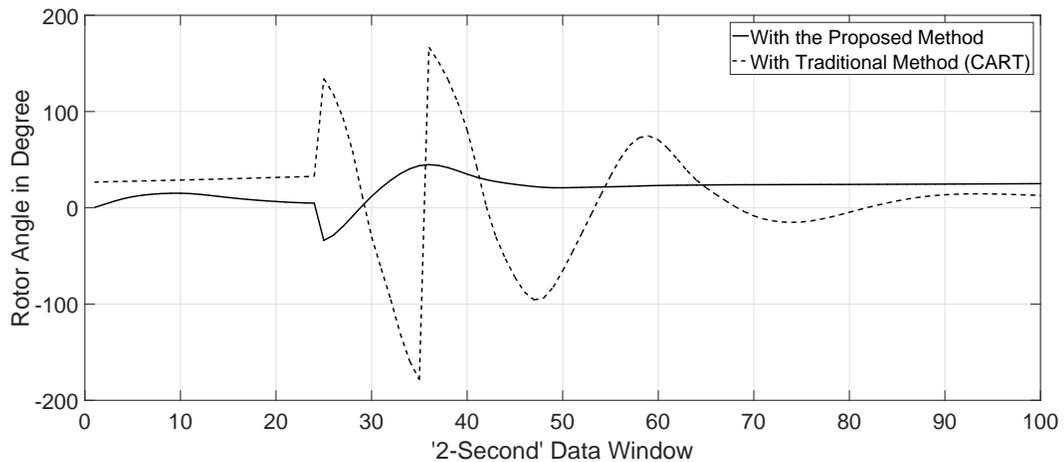


Figure 5.14: Decision making with and without the proposed algorithm

Figure-5.15 shows a post-fault operation based on the decisions made. This shows how the distributed SCs are operated once the fault location is detected near generator-1. Only the SC close to generator-1 is operated in bypass mode and the oscillation is damped to achieve a post-fault stable state. The switching in SC takes place immediately after the fault is cleared.

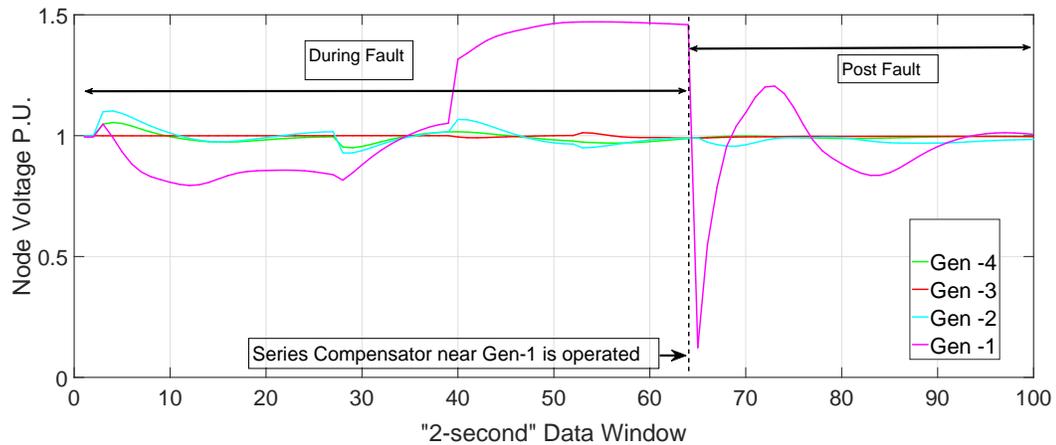


Figure 5.15: Post fault operation of Generator-1 with the Nearby SC in Bypass Mode

## 5.4 Understanding the Limitations

This study develops and demonstrates a novel algorithm to detect power system dynamic events by implementing a less computationally expensive feature selection method. The performance of the algorithm is clearly superior to the traditional classification approaches. Minimal classification errors were found to occur during the transition periods from one event to another. Once the data window falls into only one category the classification becomes close to 100% accurate.

The performance of the algorithm was tested on an offline basis only. Any misclassification due to time lag or missing data was ignored. Another key consideration with the current algorithm is selecting an appropriate data window. Faults are considered to have a fixed length before being cleared. During each of the short data windows, loads as well as non-dispatchable generations are

considered constant. In future studies, dynamicity of the loads, intermittent generation and fault duration will also be addressed.

The detection of events is mostly based on one type of three-phase fault leading to rotor angle instability. More comprehensive analysis needs to be carried out for other types of instability. Furthermore, the decision-making capability can be made more extensive by considering optimal restoration actions for a cost-effective post-fault operation.

Overall, it can be stated that the proposed algorithm has shown promise, despite being less computationally intensive than the traditional methods. However, the concept has been tested only with a smaller system. The method of feature selection must be further explored and tested in a larger network.

# Chapter 6

## Data Augmentation for a Segmented Grid

In this chapter the concept of feature selection is further explored, with a larger power system, having the capacity of being divided into multiple segments under duress. This chapter has been accepted in the conference IEEE-PES, Asia 2017.

### 6.1 Features in Larger Networks

The concept of standalone microgrids for addressing economic and environment-friendly electricity generation is ever pushing the limit of load margin. One of the key objectives in a microgrid is cost-effective operation through the integration of large-scale renewable energy sources. This factor also contributes to voltage instability. To address this challenge, modern grids often put an effort to introduce the self-healing feature through software intelligence. The advent of the phasor measurement unit (PMU) has further strengthened the avenue of self-healing [97, 103]. Through a framework or multiple layers of decision making processes, autonomous systems can be deployed in achieving self-healing mechanisms. An end-to-end system configuration often requires a hierarchical structure; The bottom layer for measurement, the mid layer for deploying intelligence in the substation level and finally the top layer is the centralized command [14]. Deployment of an emer-

gency control scheme in the mid and bottom layer can address and solve some of the contingencies in a distributed order, avoiding any interventions from SCADA which is in the top layer. Such effort helps reduce operational costs and increase speed of operation; thus local decision making for service restoration is getting more attention in recent years [8–10, 106]. After a major fault, service restoration in a sectionalized microgrid demands distinct solutions. In reference [106] a novel self-healing approach has been introduced. It distributes the self-healing process into multiple segments of a fault affected microgrid. The demand balance is maintained either by adjusting the power output or by shedding loads in each of these sections. However, the study considers that no dynamic and transient stability issues would arise while applying the self-healing process, which often is not the case [100]. System instability in a sectionalized microgrid can be understood by observing the dynamic behaviour of the generators both in pre and post-fault conditions [10]. This indicates that to develop an automated service restoration mechanism, while considering dynamic events, understanding the methods behind low level machine control could be resourceful. [139] argues that analytical calculation of the post-fault system instability can give an acute range of post-fault operating points. This method is quite effective in simplifying the post-fault protection schemes but the method does not consider any stochastic element present in the network. Wind power integration, variable demand and fault locations are some of the key elements introducing uncertainties i.e. instability in a multi-machine network. As the line voltages and frequency in a microgrid can be controlled by controlling the synchronous machines in the network, the dynamics in the machine parameters potentially can distinguish different events. Recent literature thus infers that the local or distributed control is quite dependant on the identification of dynamic events under different contingencies [10, 97, 103, 105]. To maintain the post-fault voltage stability, the most common measures are demand response (DR) and involuntary load curtailment (ILC). Under different stochastic scenarios, specially under post-fault contingencies, while the system is subject to transient and dynamic stability issues, the effectiveness of DR is yet to be discussed in larger

contexts [140]. On the other hand, ILC is the least desired avenue. Numerous studies, therefore, turn towards the concept of a corrective voltage control (CVC) scheme, which is based on controlling active power generation from the fast-response generating units and reactive power generation from VAR sources like synchronous generators [23, 140, 141]. In reference [23] the CVC has been formulated as a scenario based multi-objective function. Each scenario reflects the availability of wind power and demand. The multi-objective function minimizes the risks of voltage instability at a lower cost. The overall method, despite being technically sound lacks the ability to analyse the development of scenario-based active and reactive power dispatching scheme from a point of view of a post contingent unstable system. The fast service restoration schemes right after detection of a system instability is therefore not observed in the scenario-based assessments, which is a key factor in any self-healing microgrid. Besides the applied fuzzy satisfying method to select post contingent operating points is not intelligent which can also contribute to delaying the service interruption.

This part of the analysis implements a scenario-based optimized CVC technique similar to [23] from the perspective of controlling a synchronous machine. Instead of analysing cost efficiency the methods address the transient stability issues in the optimization routine, especially in a system which is vulnerable to critical short circuit faults. To address the shortcomings of DR presented in reference [140] a machine learning based system has been proposed that works on the data prepared by the optimization routine as well as a feature extracted from the machine data. The feature selection process is used to prepare the data set for enhancing the decision-making capabilities of the proposed algorithm [108]. The overall objective of the machine learning platform is to identify system instability and based on the stochastic scenario of the wind power and demand, take rapid decisions to stabilize the system, during periods of instability. The algorithm is trained on an offline basis and then tested online using IEEE-39 bus **10**-machine systems with an addition of an offshore wind farm at bus-**36** [142].

## 6.2 System Under Consideration

The proposed microgrid for this analysis as shown in Figure-6.2 is an IEEE-39 bus **10**-machine test system. The system is quite suitable for stability analysis [143]. During a critical contingency, the system can be divided into a sectionalized network, which is necessary to implement distributed control. Individually, these sections are considered as microgrid in this study. For the distributed control, each area has been equipped with multiple synchronous generators, which adopt real power frequency and reactive power voltage droop control. For dispatchable energy generation a wind power plant, based on an induction generator, is considered and connected in bus-**36** at the proximity of Generator-**7** [142]. The system considers two types of loads; critical invariant and non-critical variable load. The synchronous generators are considered to provide for the base load. The capacity of the wind turbine is chosen to cater the variable load. Based on the availability of the wind power and variable load the loading margin is observed either sufficient or insufficient. The variable load is considered lumped in bus **32** [98]. To create different stochastic test scenarios, critical short circuit faults have been introduced. The critical fault introduces rotor angle instability as the generators swing against each other in groups [117].

The rotor angle instability introduces large voltage fluctuation in the transmission lines. The proposed CVC eliminates the rotor angle instability by analysing the dynamic data collected from all the available generators and imposing a distributed supervised secondary control scheme;  $\frac{dE_G}{dt} \rightarrow 0$ ;  $E_G$  = Generator terminal voltage. The supervised secondary machine control, based on the proposed CVC, is a modified and extended approach carried out in reference [8]. The modification is brought in the data analytics platform by considering system sensitivity regarding  $\partial P_M / \partial V_f$  and  $\partial Q_M / \partial V_f$  [144]; where  $P_M, Q_M$  are the machine active and reactive power, while  $V_f$  is the voltage at the fault node. An extension is carried out by introducing a **10**-machine system with multiple fault locations. The system is designed to have both a normal operating mode and self-healing

mode. In the normal operating mode, the grid is resilient enough to stabilize the system after the fault is cleared. On the other hand, the self-healing mode is invoked once a rotor angle instability is observed within a *1-second* window, after the clearing of a critical fault. To identify rotor angle instability after the fault, a threshold value of the rate of change in between  $-5^\circ$  to  $+5^\circ$  has been chosen. The method is shown in Figure-6.1 during a short circuit fault near bus-16 and bus-17.

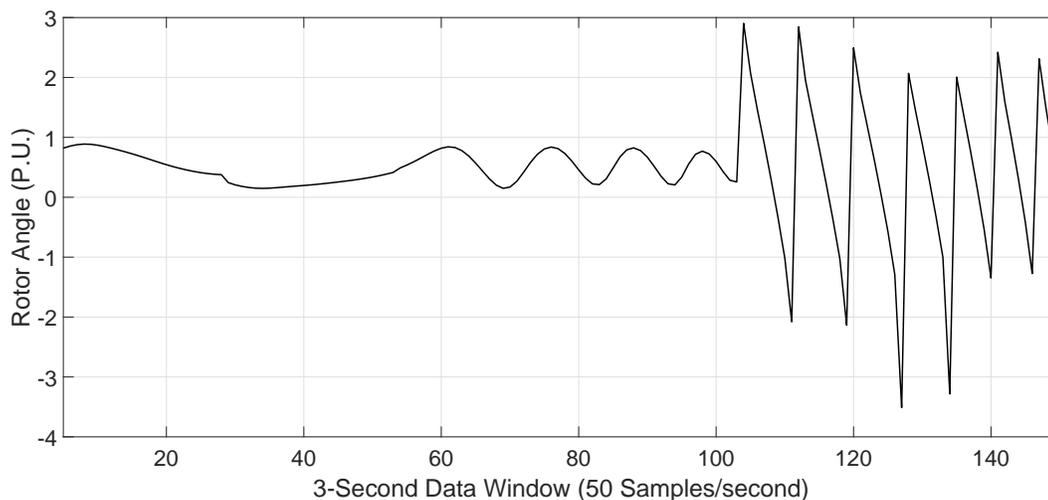


Figure 6.1: High Fluctuation in Rotor Angle due to a Critical Fault

The post-fault contingency is considered as a first swing stability problem which can be damped using a trained linear continuous control (LCC). From each distributed controllers point of view the grid is a single machine infinite bus system; thus, the active and reactive power transmitted through the transmission line can be modelled as  $P_i = \frac{V^2 \sin \delta}{(1-k)X}$  and  $Q_i = \frac{2V^2}{X} \frac{k}{(1-k)^2} (1 - \cos \delta)$ . Here  $i$  stands for the synchronous generators,  $X$  is the line inductance and  $k = \frac{X_c}{X}$ ;  $X_c$  = Series capacitance. Once the proposed algorithm detects an instability, it invokes the LCC and damps the oscillation. The LCC is developed based on the proposed CVC optimization technique, discussed in detail in the later sections.

The model as shown Figure-6.2 serves the purpose of generating time series dynamic data to train, validate and test the proposed CVC based control scheme. The system uses a primary control based on the so called drooping characteristics of frequency and voltage. The reactive load

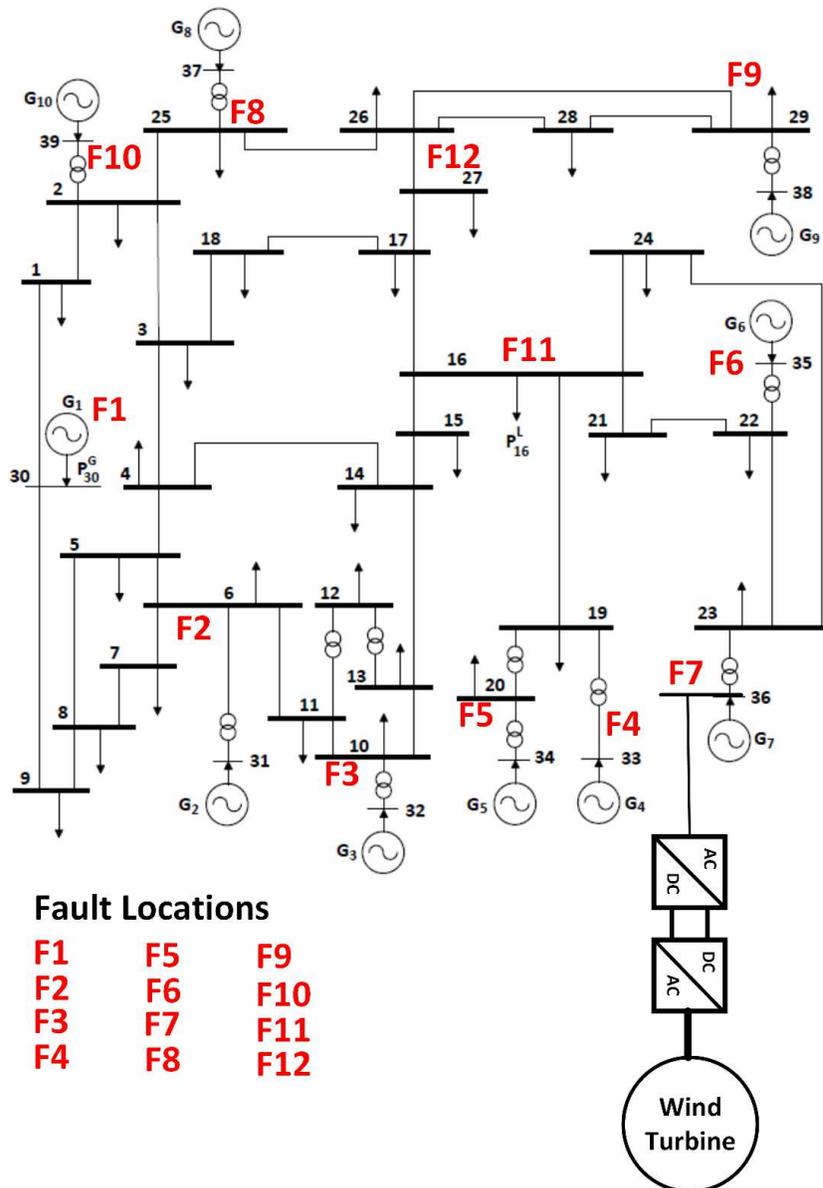


Figure 6.2: The Microgrid Model

sharing capacity of a droop controller largely depends on the feeder impedance. By introducing and clearing critical faults, overall network topology has been temporarily altered. These contingencies

are applied in **Ten** different locations very close to the synchronous generators and **Two** locations in the critical transmission lines that can be disconnected to create a segmented grid. The proposed distributed method identifies the most affected generators by any fault and takes relevant action to mitigate it. The algorithm validates the action (decision) by identifying the *1-second* window following the clearance of the fault. By analysing this window of machine data in terms of CVC optimization, stabilizing actions are categorized and processed. The categorization is prepared based on the post fault transient characteristics of the synchronous machines.

### **6.3 Proposed Secondary Control**

In a hierarchical control chain, the purpose of the primary control is to maintain the system frequency and voltage. The primary control in an isolated grid may cause frequency deviation, after a critical fault, despite maintaining a steady state condition. The secondary control acts on that observed deviation and fine tunes the control parameters with a slower dynamic. The slower dynamic decouples the two control schemes. This study implements a centralized secondary control under normal operating mode and proposes a distributed and supervised secondary control during the self-healing mode. The secondary control is considered as a part of the hierarchical control scheme of the microgrid [7, 8, 102]. The trigger point to activate the self-healing mode is set by creating and observing a stochastic database for the rotor angle and terminal voltage of individual machine.

The secondary control is applied after the machine initialization has been taken care of and the system has reached a steady state condition. The proposed supervised control is a restorative action over the primary controller. The overall method has been shown in Figure-4.2.

The workflow, for the proposed supervised secondary control, applied in this section is shown in Figure-6.3. The workflow is divided into two factions 'Offline' and 'Online'. The training stage directly contributes in the online stage. Once the algorithm is trained it works as a function in the

online stage. The crisscrossing lines have been used to show information exchange between these two stages.

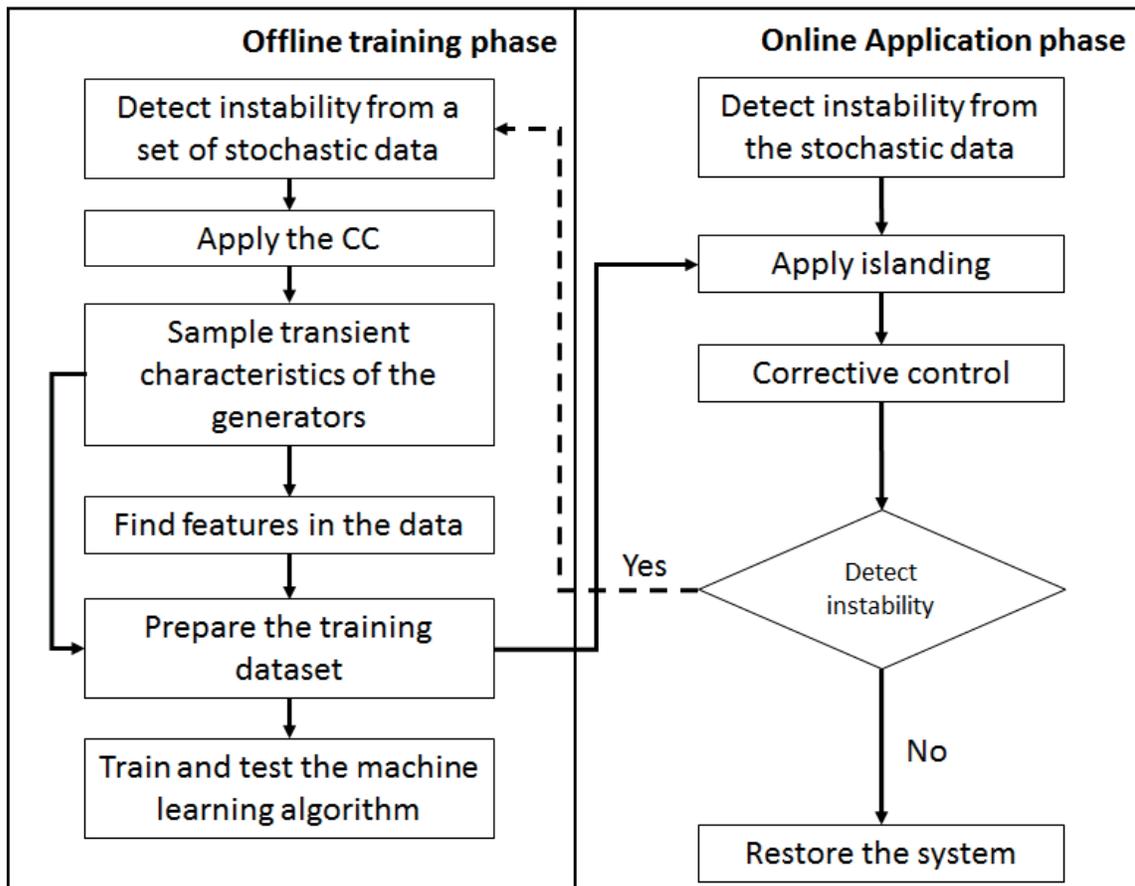


Figure 6.3: Workflow of The Proposed Method

The data preparation stage is based on sampled stochastic scenarios of random wind power and demand. The wind power is considered proportional to the wind speed. A K-means clustering technique is used to categorize each stochastic scenario into a cluster. For simplicity this study deploys 9 clusters considering 3 levels in wind speed and 3 levels in demand. The cluster is prepared using a matrix of normalized wind speed and demand data. Figure-6.4 shows the clusters prepared for simulation optimization. Each colour represents one cluster.

Based on the different stochastic scenarios and fault locations a sensitivity analysis is also car-

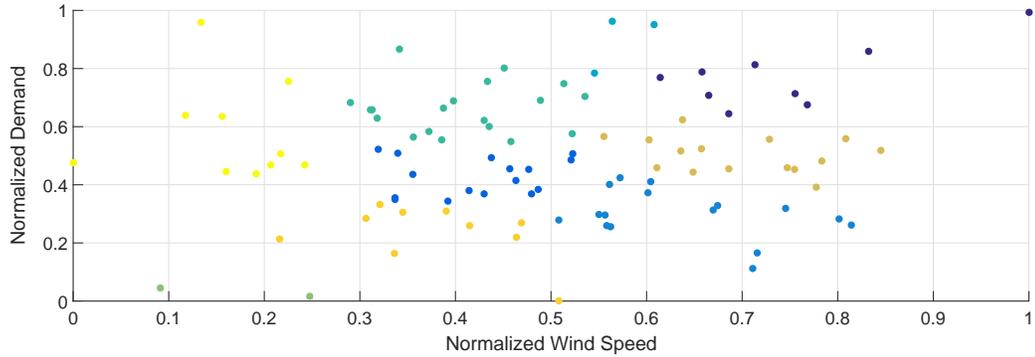


Figure 6.4: Preparation of the **Nine** Stochastic Clusters

ried out. The sole purpose of introducing a sensitivity analysis is to calibrate the optimization routine. The P-V and Q-V characteristics have been considered under three different contexts pre-fault steady state, post-fault transient state and post-fault steady state condition to estimate the control parameters [144]. One instance of the sensitivity curve is shown in Figure-6.5;

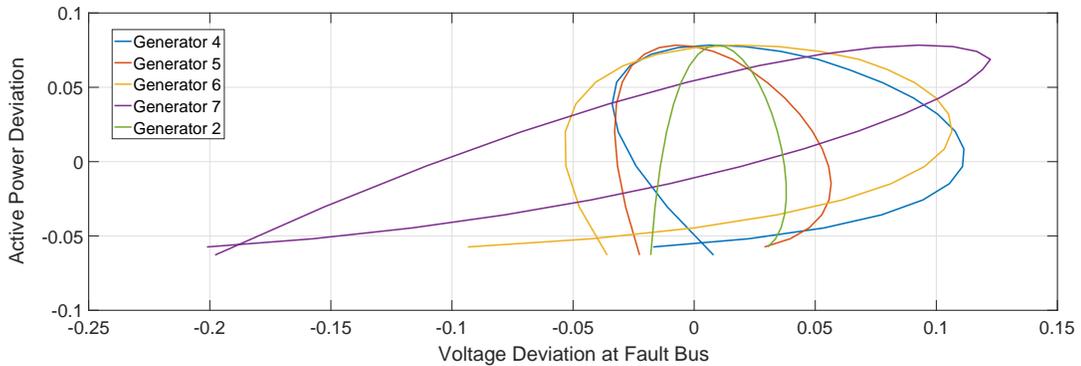


Figure 6.5: Sensitivity  $\partial P_m / \partial V_{fbus}$  Curve of Fault-11

In Figure-6.5 four sensitivity curves, of the generators  $\{ G7, G6, G4, G5 \}$  closer to the short circuit fault, are compared to that of a generator  $\{ G2 \}$  not affected by the fault. The analysis shows the aperture of the curve increases if the effect of the fault is greater on any machine. The aperture is calculated using the numerical integration by trapezoid method. The limit of each response curve is considered between two extreme points over the  $\mathbf{X}$ -axis or the  $\Delta V_{fbus}$  axis. Each curve is then

equally distributed into two halves around the extreme point as convex(Upper) and concave (Lower) perimeters. Finally, the overall area is calculated by subtracting the area under the concave curve from the convex;  $Area = A_{convex} - A_{concave}$ .

$$\int_{V1}^{V2} f(x_{convex})dx \equiv \frac{V2 - V1}{2N} \sum_{n=1}^N (f(x_n) + f(x_{n+1})) \quad (6.1)$$

$$\int_{V1}^{V2} f(x_{concave})dx \equiv \frac{V2 - V1}{2N} \sum_{n=1}^N (f(x_n) + f(x_{n+1})) \quad (6.2)$$

The calculated area is considered as a feature in the generator data. Each contingency develops a set of feature data of the generators. Under one stochastic scenario the area feature is normalized between the largest and the lowest absolute values. Depending on the feature data the upper bounds of the optimization variables are set while stabilizing the system with a cost function of;

$$Objective : \min(\epsilon_{V\ fault} = \sum_{i=1}^N [V_{prefault} - V_{postfault}(i)]^2) \quad (6.3)$$

For optimization in the proposed CVC, Genetic Algorithm (GA) has been chosen [145]. GA is optimizing the supervised active and reactive power generation from each generator during the self-healing mode for a window of ‘4-seconds’. The forced power generation is maintained to damp the rotor angle instability and the voltage fluctuation. Once the oscillation is damped, the supervised secondary control is removed [8]. The objective function is subject to the following constraints;

$$P_{dpfcl} + P_{dpfncl} + P_{Tloss} \leq \sum_{i=1}^{n_g} pg_i + \sum_{i=1}^{n_{ren}} pren_i \quad (6.4)$$

where  $P_{dpfcl}$  and  $P_{dpfncl}$  are the critical and non-critical loads. Due to a sudden change in demand, in some cases, deviation in power generation can be observed. In this study the post fault demands are kept constant and any large deviation is neglected. Thus,  $pg_{i,t} - pg_i, t0$  is considered zero for the half an hour span pre-defined for the stochastic data points. Therefore, the upper ramp

rate  $UR_i$  and lower ramp rate  $LR_i$  have been neglected. To maintain post fault frequency, the stability sufficient spinning reserve should be available.  $\sum_{i \in G} SR_{i,t} \geq SSR_t; \forall t \in T$ .

Where,  $SR_{i,t}$  is the available spinning reserve of individual diesel generator at  $t$ -th half hour, and  $SSR_t$  is the system wide required spinning reserve. Each generator also follows the generator output constraints, which means that generation does not exceed its upper limit  $pg_{i,t} \leq PG_{i,max}; \forall t \in T$  and also for the renewable energy generators  $pgren_{i,t} \leq PGren_{i,max}; \forall t \in T$ . Furthermore, wind and synchronous generators are subject to their active and reactive power limit  $0 \leq P_{scenario_n}^{wind} \leq P_{max,scenario_n}^{wind}$  and  $P_{i,min}^G \leq P_i^G \leq P_{i,max}^G; 0 \geq Q_{scenario_n}^{wind} \geq Q_{min,scenario_n}^{wind}$  and  $Q_{i,min}^G \leq Q_i^G \leq Q_{i,max}^G$ .

After each optimization routine is completed a stability database of stochastic wind and demand data with different fault location is prepared. The data table contains the information of forced active and reactive power generation from each generator during each contingent scenario. Once the data table is prepared, an *Adaptive Neuro Fuzzy Inference System* is trained 'offline', as shown in Figure-6.6. The fuzzy inference system is considered as having dual inputs ' $x$ ' and ' $y$ ' that result in an output of ' $z$ '. For a first order Sugeno fuzzy model typical if-then rules can be established as the following;

Rule 1: If  $x_1 = a_1$  and  $x_2 = b_1$  then output  $z_1 = p_1 \times x_1 + q_1 \times x_2 + r_1$

Rule 2: If  $x_1 = a_2$  and  $x_2 = b_2$  then output  $z_2 = p_2 \times x_1 + q_2 \times x_2 + r_2$  The terms  $p_i, q_i$  and  $r_i$  are the linear parameters of the consequent 'THEN' part of the first order fuzzy inferencing model [53].

The outcome of the proposed algorithm finally can be stated as an optimized, supervised, active and reactive power output from the affected generators in a fault stricken segment of a microgrid.

## 6.4 Result Analysis

The data preparation step develops a table based on the stochastic scenarios of wind speed and demand. The Table-6.1 shows some of the distributed control incidents for the Generator-7 and

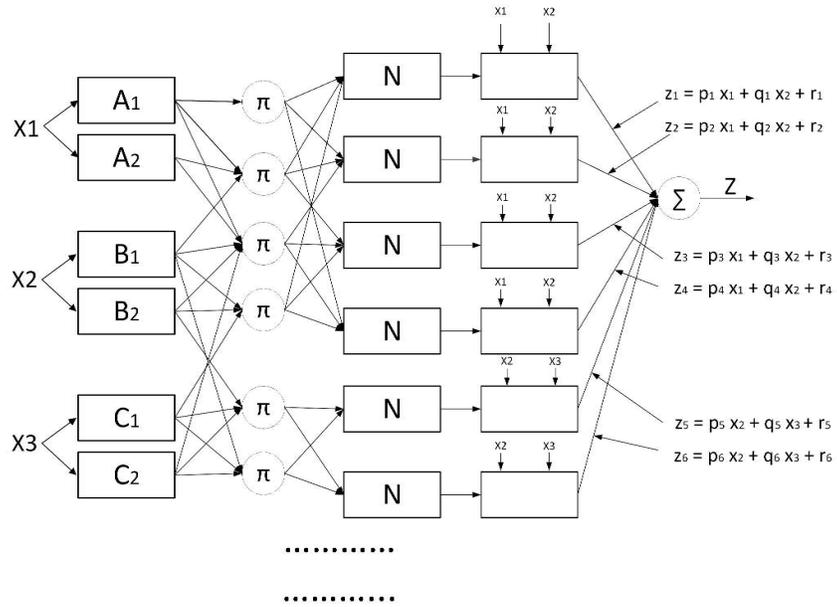


Figure 6.6: Five-layer Three-input Four-Rule fuzzy model (Sugeno) [1]

Generator-9.

Table 6.1: Stochastic Data Preparation Stage for Generator-7&9 (Selected Scenarios)

Wind Power	Demand	Cluster	Fault Location	Sensitivity Feature	Optimized Active Power (P.U.)	Optimized Re-Active Power (P.U.)
High	Me	2	B-16	0.4712	G7: 0.70	G7: 0.62
Low	Me	8	B-16	0.4034	G7: 0.75	G7: 0.56
Med	Low	4	B-26	0.1729	G7: 0.45	G7: 0.57
---	---	-	---	---	---	---
High	Med	2	B-16	0.2052	G9: 0.41	G9: 0.62
Low	Med	8	B-16	0.2284	G9: 0.39	G9: 0.62
Med	Low	4	B-26	0.5913	G9: 0.73	G9: 0.55

The scenario-based data consists of the stochastic cluster, fault location and sensitivity feature, which is fed to the ANFIS model, to predict the optimized active and reactive power generation from each generator in a distributed order. One incident of ANFIS based prediction for the

Generator-7 is shown in Figure-6.7.

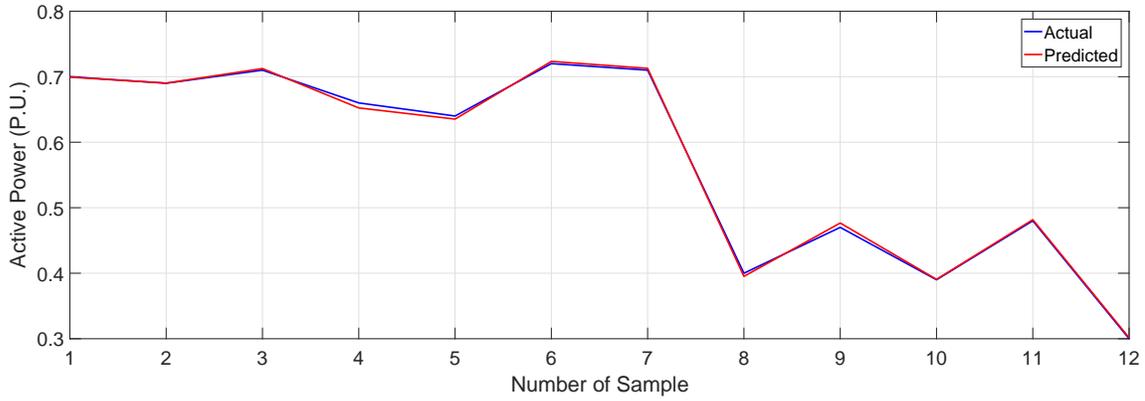


Figure 6.7: ANFIS Based Prediction for Generator-7

The result for predicting the optimized power for Generator-7 is quite accurate. However, a minor classification error is observed in the data sample 4. From an individual machine perspective, such a minor classification error can be neglected.

The overall performance of the proposed method is shown in Figure-6.8, during a critical fault introduced in the Bus-16. The critical fault is cleared by disconnecting the transmission line between the Bus-16 and the Bus-17, as well as the Bus-16 and the Bus-15. During the process of clearing the fault two segments are created. The *Segment-1* consists of the synchronous machines G4, G5, G6, G7; while the *Segment-2* has the rest of the machines connected. The contingency introduces rotor angle instability in the *Segment-1* thus supervised control is imposed on the 4 mostly affected synchronous machines.

Furthermore, in Figure-6.9 the rotor angle deviation compared to the system slack bus and rotor speed of the Generator-7 are elaborately presented. The solid *red* line marks the beginning of the short circuit fault; the dashed *red* line marks the clearance of the fault. The solid *blue* line marks the identification of rotor angle instability and starting of the supervised control scheme while the dashed *blue* line marks the end of the supervised control scheme and handing over to the primary control of the system. The solid *green* line marks the re-closing of the transmission line between

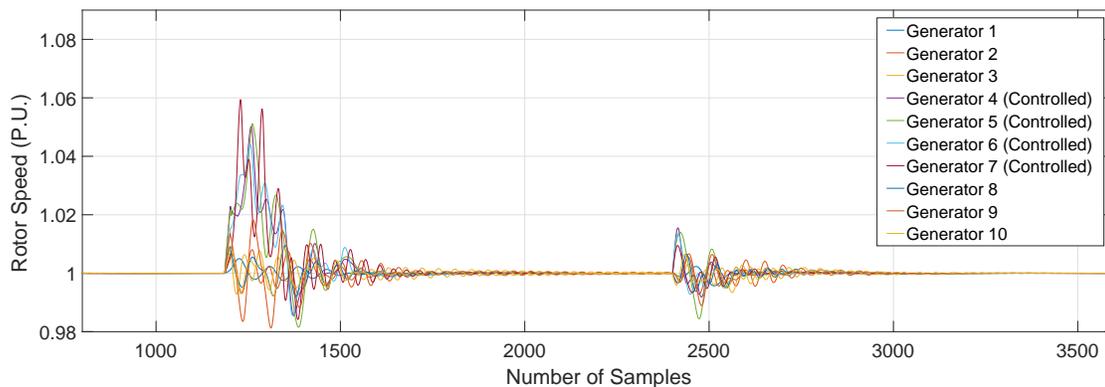


Figure 6.8: Stabilizing the Segmented Microgrid with the Proposed Algorithm

Bus-15, Bus-16 and Bus-17.

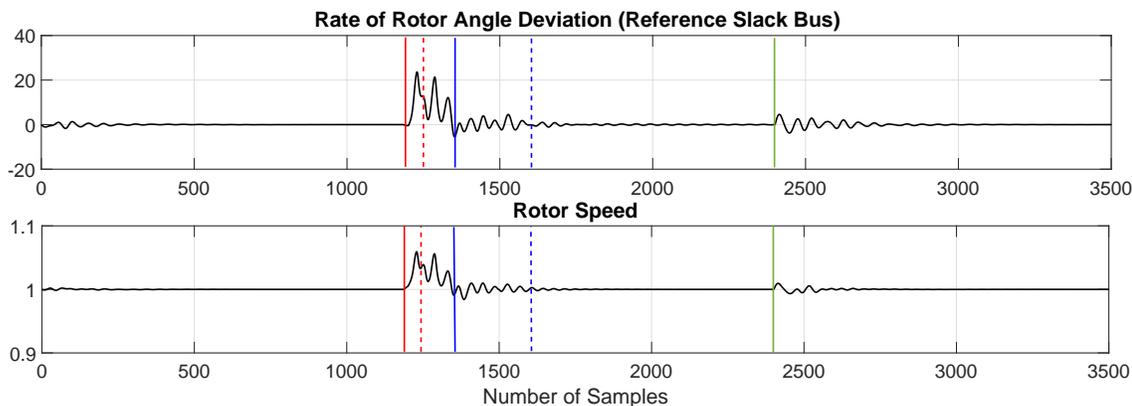


Figure 6.9: Generator-7 Data

Finally, in Figure-6.10 a comparison between the fault bus voltages is shown. The comparison is carried out among the proposed algorithm with optimized CVC, without the CVC optimization (pre-fault fixed value of the control parameters) and without the proposed algorithm. It is observed that the proposed algorithm has better damping coefficient thus better resiliency than the rests.

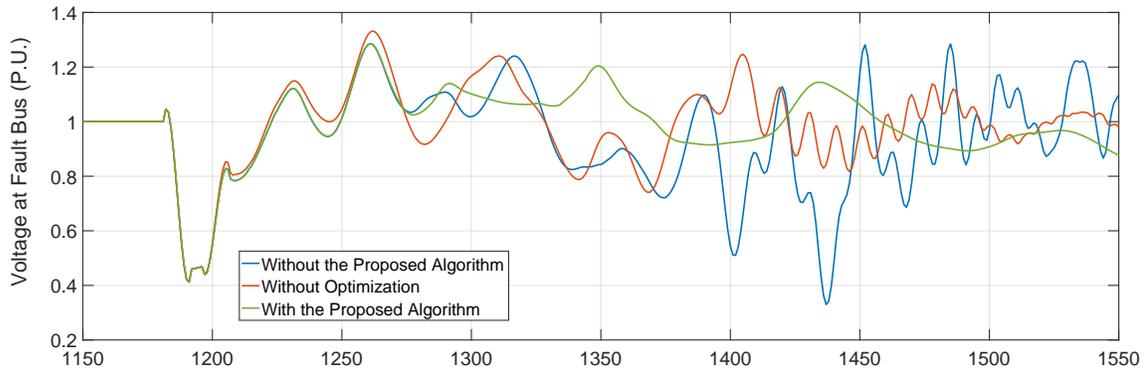


Figure 6.10: Comparison of Voltages at the Fault Bus With Different Methods

## 6.5 Analyzing Limitations

The proposed algorithm successfully eliminates rotor angle instability and restores system voltage in a segmented microgrid. The method works on a distributed platform without any intervention from the central station during a post-fault operation. However, the algorithm is tested on a sampled post-fault scenario with only 9 clusters. A comprehensive analysis with more variations would provide further insight. The study also does not consider classification errors. Thus, the full ramification of a few clusters against accuracy is not addressed in this study. Besides these minor issues, the overall performance of the algorithm on IEEE-39 bus test system for self-healing is promising. Therefore, an exploration of different possible power system features is required to test the effectiveness of the proposed method.

# Chapter 7

## Selection of Appropriate Feature Space

This chapter focuses, on the selection of different types of features and their impact on a self-healing microgrid. The complexity of the self-healing microgrid has been further increased in order to carry out this investigation.

### 7.1 Different Types of Features

Self-healing is a lucrative feature in the service restoration process of a power system. In recent studies, under the category of post-fault islanding scenarios, this self-healing mechanism is getting significant attention [97]. The prominent trend in addressing self-healing is established through the local or distributed control. It is because distributed control is quite effective in faster decision making. However, applying local or distributed control is quite dependant on the dynamic characteristics of the network under analysis. On the other hand, dynamic characteristics are heavily influenced by the pre-contingent stochastic parameters, fault analysis, demand response (DR) and the performance of the algorithms assessing power system security [109].

In previous studies, different methods and systems have been investigated for fault detection, isolation, and service restoration (FDIR). Some of these methods are Central Controller (CC), Dis-

tribution Automation System (DAS), Automatic Controlled Switching (ACSs), Fault Passage Indication (FPI), Corrective Voltage Control (CVC) and the Emergency Demand Response Program (EDRP) [146]. Most of these approaches are data intensive programs, thus prime candidates for machine learning algorithm-based systems, including Artificial Neural Network (ANN), Support Vector Machine (SVM), Random forest. If the underlying events during any critical contingency are properly identified [147], the identification can lead towards developing intelligent self-healing strategies. The process can be developed through expert systems, with a large scenario-based dataset. The performance of the machine learning based systems with a large data set is commendable. However, machine learning systems are affected by processing time and memory. In a fast restoration system, it is not desired, especially for those grids that require adaptive online decision making. Besides, for a large network, performing the dynamic security assessment (DSA) is an increasingly complex problem affected by numerous criteria [148, 149]. Under these contexts, the traditional machine learning applications are progressively being less effective.

Although the traditional machine learning algorithms are trained on an offline basis, generation of large-scale data often makes this approach unattractive for online DSA programs. Therefore, many studies conducted often recommend alternate solutions such as feature selection [21]. The feature selection process is implemented to modify a data set to increase the accuracy of prediction [107–109]. However, it adds redundancy in the algorithmic steps [110]. In reference [109], an energy function-based feature selection is proposed to outperform the raw inputs, to train the machine learning algorithms with a smaller dataset. This strategy proves to be highly effective, but a limited dataset is insufficient while addressing multiple stochastic scenarios because in a hybrid microgrid, stochastic parameters make it challenging to assess dynamic security using analytical approaches. One solution to address this challenge is a simulation-based Monte Carlo method. This approach can establish the stability boundary of a grid with multiple stochastic parameters, [150, 151]. Once the stability boundary is established for a coherent group of stochastic

parameters a limited data set can be used for online DSA programs. This study considers this approach.

Multiple power system features have been introduced in this study to develop a distributed machine learning platform. The underlying objective is to find patterns to classify different dynamic events. The events chosen for this analysis are focused on stabilizing a sectionalized power system and restoring the grid within a shorter duration, without any major service interruptions. The events are categorized based on two aspects: stabilization of a section and restoration of the section by reconnecting it to the rest of the system. The feature-based machine learning platform is used to classify events and make decisions. Each decision refers to a set of actions that will ensure system restoration without further instability. Each of these decisions is based on different stochastic scenarios. For generating those scenarios, a Monte Carlo method-based simulation strategy has been deployed. An IEEE-39 bus 10-machine test system as shown in Figure-7.1 is used in this study [143]. During a critical contingency, the system can be divided into multiple sections, which is necessary to eliminate instability. While developing the algorithm, these segments in the IEEE-39 bus system have been individually considered as microgrids.

## 7.2 Machine Data and Additional Features

The IEEE-39 bus system is suitable for stability analysis. For non-dispatchable energy generation, a wind power plant based on an induction generator is considered and connected in bus-36 at the proximity of Generator-7. The demand and wind speed are randomly varied to do a Monte Carlo based simulation, to prepare a stability database with the synchronous generator data. The parameters chosen are in per unit quantity and these are rotor speed  $\omega$ , rotor angle deviation  $d\delta$ , active power generated  $P_E$ , reactive power generated  $Q_E$  and terminal voltage  $E_G$ . For simplicity, only the cases where the system can be stabilized and restored by dividing it into two segments are considered. Based on these cases and the generator data four features are prepared:

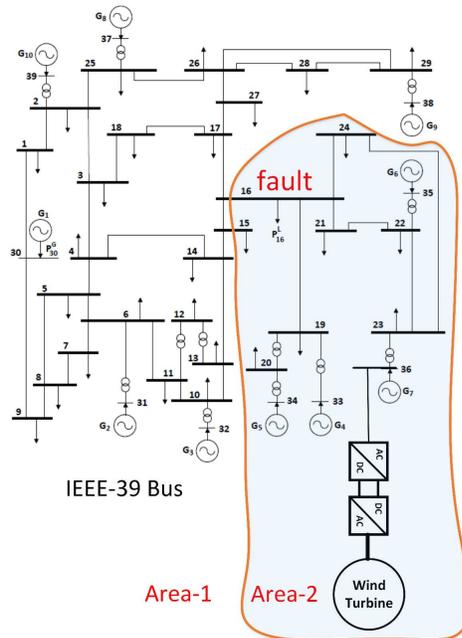


Figure 7.1: The Microgrid Model

1. Prominence of local maxima in terminal voltage [126–129]
2. Available frequencies in the time series voltage data [132, 134, 135]
3. Principle components in the generator data of the islanded sections [147]
4. Sensitivity  $\partial P_E / \partial V_f$  and  $\partial Q_E / \partial V_f$ . where,  $V_f$  is the voltage at the fault bus [22, 150].

Figure-7.2 shows an example of the terminal voltage at Generator-7 under different scenarios, during the beginning of the rotor angle instability. The similarity in the time series data is so overwhelming that it makes application of a machine learning algorithm quite difficult. However, each scenario has its own features that can be used to distinguish them.

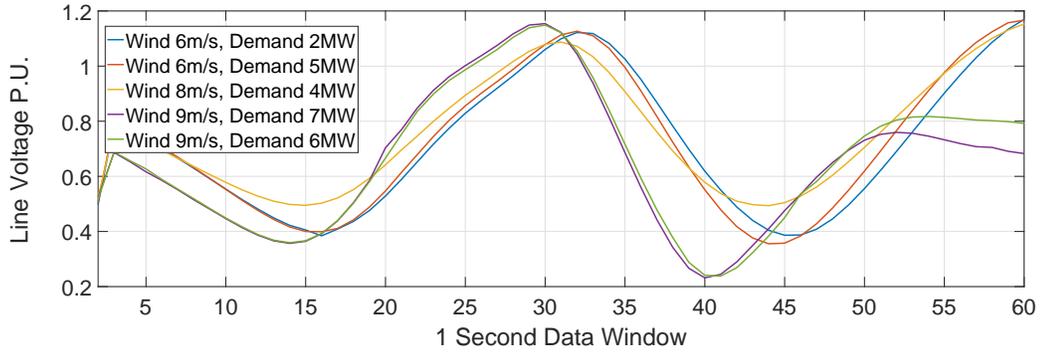


Figure 7.2: Time Series Data of Generator-7 Under Different Scenarios

### 7.3 Corrective Voltage Control Data

The primary control in an isolated grid may not be sufficient as it may cause frequency deviation after a non-critical fault, even if the system is in steady state. Therefore, the secondary control implements a slower dynamic to fine tune the control parameters and prevent frequency fluctuations. However, under a critical fault when the system wide rotor angle instability is observed, a secondary control often fails to stabilize the system [23, 118]. This study implements a machine learning driven supervised secondary control during the post fault scenarios, where secondary control cannot maintain stability after a critical fault. The supervised secondary control is considered as the topmost part of the hierarchical control scheme in this microgrid [7, 8, 152]. The hierarchical supervised control and the overall workflow has already been discussed in the earlier sections, Figure-4.2 and Figure-6.3.

### 7.4 The Machine Learning Platform

To understand the full ramifications of using a feature data set, a machine learning platform has been introduced. The machine learning classifies decisions based on the stochastic scenarios that can restore an unstable and sectionalized microgrid. An ensemble of bagged decision trees has

been chosen as the classification algorithm.

### 7.4.1 Features for Pre-processing

To prepare the features in each data window, sixty samples per second have been considered. Topographical prominence has been chosen as the first feature. The method followed here was inspired by [129]. A normalized window of *1-second* duration in the time series data was chosen to carry out the feature extraction. For simplicity transitional windows have been ignored.

First, the lowest contour line in the *1-second* data window circulating a local maximum point has been detected. Then the height of that point is measured regarding that contour line. Figure-7.3 shows the method of finding prominent local maxima from a 1-second window of the terminal voltage of one of the generators with an example data set. Once several prominent peaks are detected, their average heights are considered as a feature for that data window.

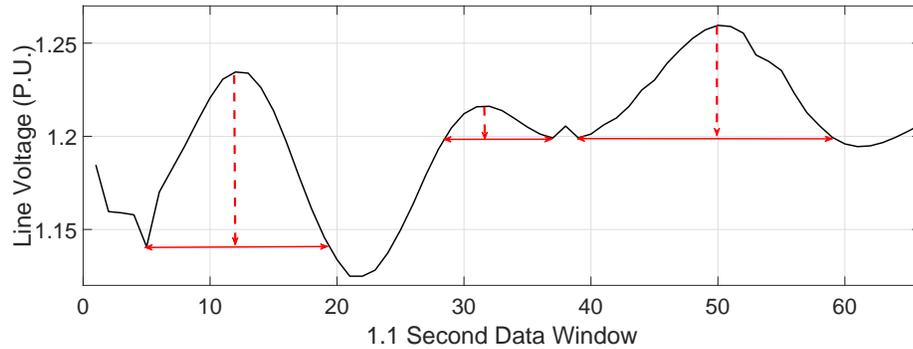


Figure 7.3: Peak prominence and width inside a predefined data window of the Terminal voltage of Generator-1

The second feature is based on the available frequencies in any oscillation, observed in the 1-second data window. This feature is extracted through a Discrete Fourier Transformation (DFT) technique;  $x[n] = \sum_{k=0}^{N-1} \frac{1}{N} \tilde{X}[k] e^{-jk(2\pi/N)/n}$ ,  $K = 0, 1, \dots, N - 1$ . Here,  $\tilde{X}[k]$  is the amplitudes and  $x[n]$  is the linear combination of the complex exponentials with that amplitude. Decision making based on frequency data has often been observed in the field of harmonics and power

quality analysis [132, 134, 135]. In this study, a frequency spectrum of the first 10 Hz has been chosen as the featured attribute as shown in Figure-7.4 (with a small displacement in the X-axis for the purpose of visualization). The chosen attribute is the speed of the generators. Once the magnitudes of the prominent peaks are calculated, all the magnitudes are summed up to be used as a feature for that data window,  $\sum_{i=1}^{10} M_{fi}$ ; where,  $M_{fi}$  is the  $i$ -th magnitude of the frequency feature.

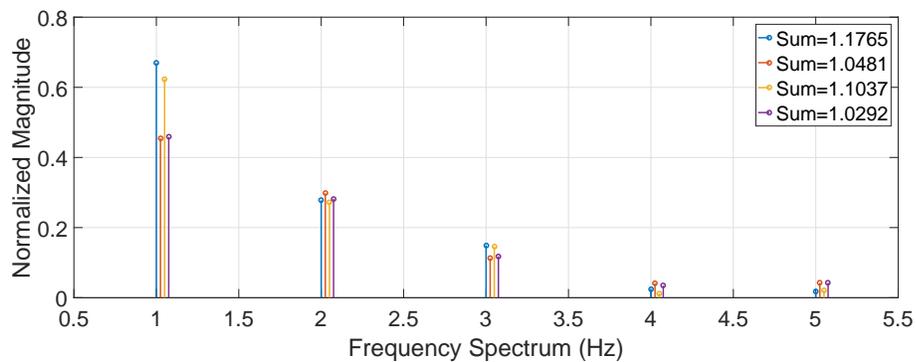


Figure 7.4: Available Frequencies

The third feature is the principle component of the machine data attributes, observed in one segment of the network. The Principle Component Analysis (PCA) is used to find the first component account for most of the variations in data. The data matrix  $X$  has  $m$  number of observations with  $n$  number of attributes. The  $m$ -number of orthonormal basis functions are represented by  $w$  which is of  $n \times n$  dimensions. The data matrix is then represented as  $X = TW'$ , where  $T$  is  $(t_{1,i}, \dots, t_{m,i})$  is used to form the coherent clusters. To create a feature out of this information, a mean of the absolute value of the most significant principle component has been considered for the data window. Figure-7.5 shows the threshold observed while considering this feature to understand stability. 'PC-1' is the first principle component.

A sensitivity data based on the voltage at fault bus and the active power generation from the subject generator is considered as the fourth feature [144]. One instance of the sensitivity curve is shown in Figure-7.6;

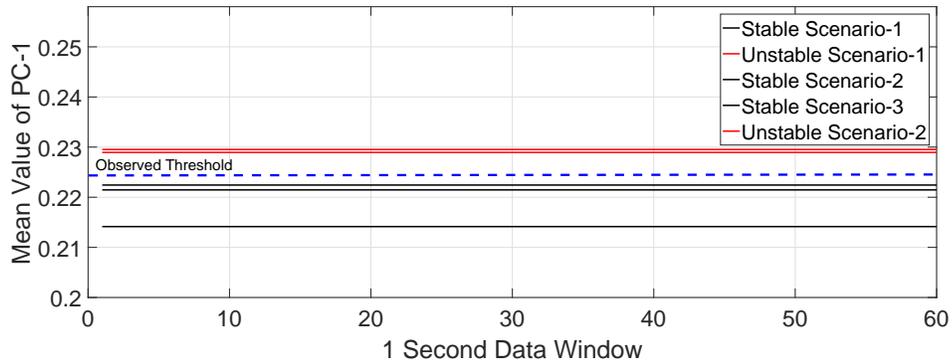


Figure 7.5: Available Frequencies

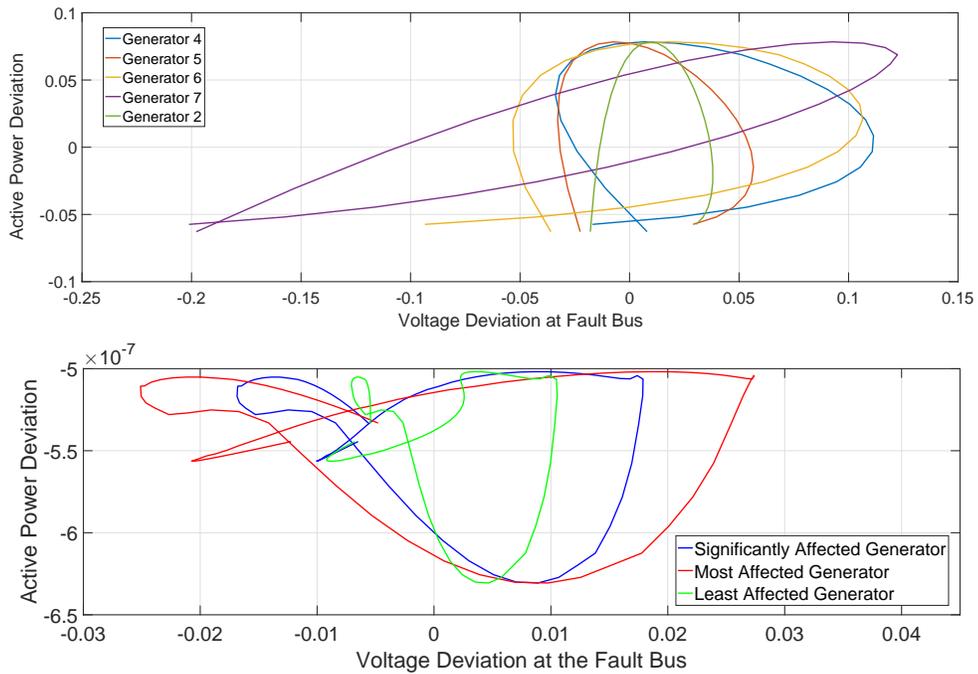


Figure 7.6: Sensitivity  $\partial P_m / \partial V_{fbus}$  Curve

In Figure-7.6, three sensitivity curves of the generators  $\{ G7, G1, G8 \}$  are shown. It clearly shows the aperture of the curve varies depending on the impact of fault. To calculate the aperture, the trapezoid method of numerical integration is used. The limit of each response curve is considered between two extreme points over the X-axis or the  $\Delta V_{fbus}$  axis. Two extreme points develop convex (Upper) and concave (Lower) perimeters. Then the area under the curves are calculated.

The overall area of the aperture is then assumed by subtracting the area under the concave curve from the convex;  $Area = A_{convex} - A_{concave}$ .

$$\int_{V1}^{V2} f(x_{convex})dx \equiv \frac{V2 - V1}{2N} \sum_{n=1}^N (f(x_n) + f(x_{n+1})) \quad (7.1)$$

$$\int_{V1}^{V2} f(x_{concave})dx \equiv \frac{V2 - V1}{2N} \sum_{n=1}^N (f(x_n) + f(x_{n+1})) \quad (7.2)$$

## 7.4.2 Multiclass Classifier

The Multiclass classifier is an ensemble of bagged decision trees trained using the stochastic parameters and the features. The bagged decision trees weigh and reweigh the predictor variables and their estimates. For example; a function estimate  $\hat{g}_{ens} = \sum_{k=1}^M c_k \hat{g}_k(\cdot)$  is obtained based on the  $k - th$  reweighed data with the combined linear estimation co-efficient  $c_k$ . This method is deployed to eliminate classification error due to estimation variance and statistical bias, especially in a stochastic scenario [136]. The ensemble implemented in this study is prepared using one hundred fully grown trees by splitting the attribute data into one hundred training sets,  $D_1, D_2, \dots, D_{100}$ . This approach helps to obtain an improved composite model.  $M_i (1 \geq i \geq 100)$  classifiers vote by predicting a class and the ensemble selects the final class from those votes. The overall technique along with *Random Forest* and *Random Subspace* also includes *Bagging* and *Boosting* [107].

## 7.5 Results

The proposed analysis is carried out on three timelines: timeline for detecting post fault rotor angle instability, timeline for CVC to stabilize the instability, and timeline representing the post restoration period. The two latter periods can either be stable or unstable. Each of these timelines has been assessed using Monte Carlo based simulation strategy, by randomly generating wind power

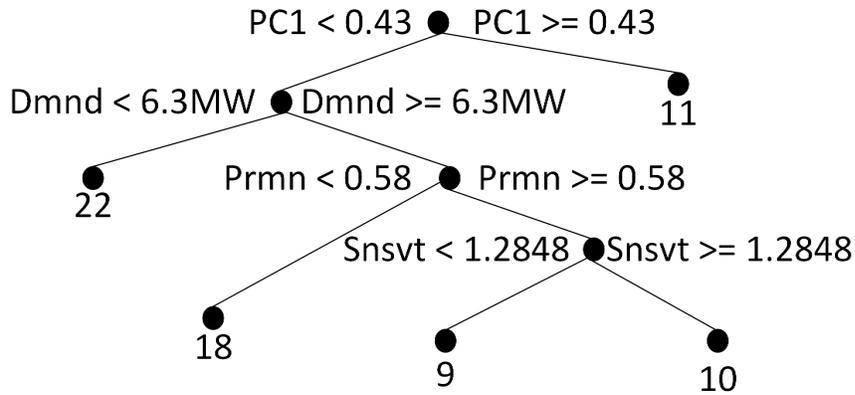


Figure 7.7: A Part of the Classification Tree

and demand. In Figure-7.8, all the different timelines have been shown. Timeline-2 data is the candidate for feature extraction and timeline-3 is the candidate data for evaluating the performance of the algorithm.

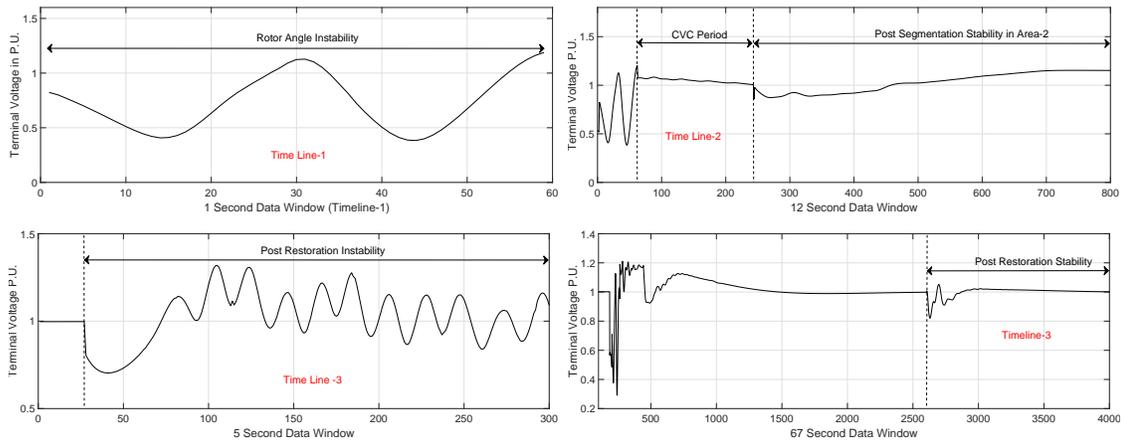


Figure 7.8: Events

Based on the randomly selected stochastic scenarios 22 decisions have been identified in this study. Those decisions can maintain a post restoration stability. Table-7.1 shows an example data set with features that have been used to predict the decision class to stabilize the restored network. *PCAI* is the mean of absolute value of the first principle component, *Snsvt* is the area of the aperture

Table 7.1: Decisions

Decision	Wind Speed	Non Critical Demand	PCA1	Snsvt	Prmn	FF
1	6	3.5e8	0.24	1.4384	67.8908	1.25974
2	6	1.45e8	0.22	1.8333	0.659	1.24694
3	6	2.05e8	0.225	1.6361	0.566	1.00874
4	8	5e8	0.2617	1.4013	0.609	0.98874
5	11	6.0e8	0.2389	1.3250	0.606	1.18874
..	..	..	..	..	..	..
..	..	..	..	..	..	..
21	10	5.1e8	0.2182	1.3495	0.642	1.48874
22	11	7.9e8	0.2340	1.4345	0.593	1.25566

calculated from the sensitivity data, *Prmn* is the peak prominence and *FF* is the frequency feature measured from the available  $0\text{ Hz} - 10\text{ Hz}$  frequencies in the oscillation.

To address the stochastic nature of the proposed algorithm, a Monte Carlo based simulation is carried out for a limited 1560 scenarios. 1000 scenarios have been used for training the algorithm and rest of the 560 scenarios have been used for testing. Figure-7.9 shows the average prediction accuracy of the algorithm with the 560 test cases. With a limited data set the accuracy of the algorithm is promising. In the Figure-7.9 a comparison with the proposed method (extracted features) and without the proposed method has also been shown. The process without the proposed method only uses 1st principle component as raw data.

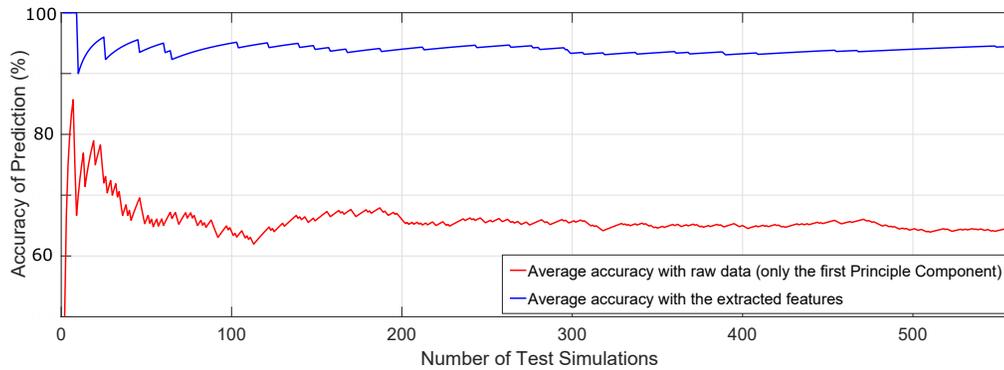


Figure 7.9: Prediction Accuracy with the Test Data

However, the rest of the 5% cases where the algorithm misclassified the decisions, the system

recovery could not be achieved. Figure-7.10 shows a comparison between two cases where in one the algorithm successfully identified the decision and in the next the algorithm could not.

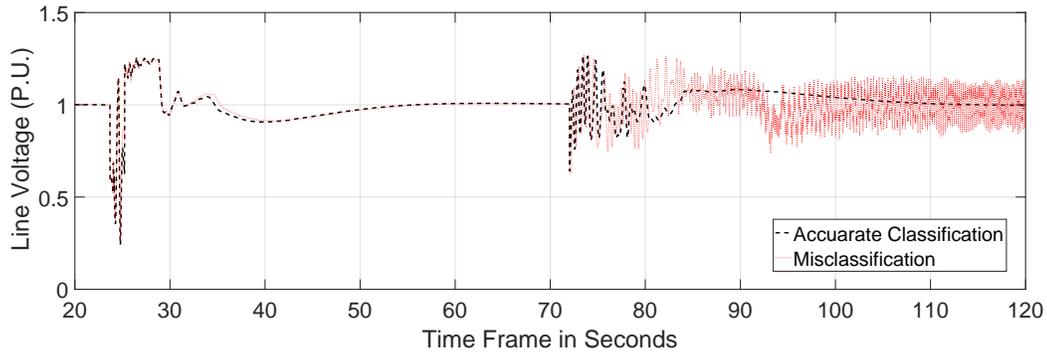


Figure 7.10: Ramification of the Classification Error

## 7.6 Overall Performance Evaluation

Overall the proposed method for analysing decision making, for post-fault restoration has shown promising results. The method works with high accuracy, despite a limited number of training data being provided. One limitation of the study is dimensionality explosion. With the increased number of faults, the system needs to be divided into more than two areas. The proposed training data set does not deal with those scenarios. In the author's future studies, the method is extended towards multiple segment restoration strategies, with a centralized controller having a communication network.

IEEE-39 bus test system has been used to demonstrate the capabilities of the proposed method. Furthermore, the proposed algorithm is tested with a high degree of stochastic influence in the network. This study is highly significant for a stand-alone microgrid where feature-based data analysis is implemented for service restoration.

To understand the performance of an ensemble of machine learning algorithms, in case of multiple decision-making scenarios, a case study with multi-objective function can be helpful. The

following chapter deals with an economic dispatch problem in a self-healing microgrid. The underlying concept is to assess the performance of a machine learning algorithm, under multi-objective functions.

# Chapter 8

## Ensemble Method for Multi-objective Application

This chapter deals with a multi-objective function, and addresses the challenges associated to that, by using an ensemble of machine learning algorithms. A classical economic dispatch problem, in a post-contingent network with stochastic parameters, has been chosen as the use-case scenario. This chapter has been accepted in Elsevier electric power system research.

( Miftah Al Karim, Jonathan Currie, Tek-Tjing Lie, A machine learning based optimized energy dispatching scheme for restoring a hybrid microgrid, Electric Power Systems Research, Volume 155, 2018, Pages 206-215, ISSN 0378-7796, <https://doi.org/10.1016/j.epsr.2017.10.015>. )

### 8.1 Decision Making With Multiple Constraints

A microgrid, having the capability of integrating distributed and renewable energy sources, can use bidirectional energy and information flow. But maintaining bidirectional power flow often introduces security issues in the modern microgrids [153]. The advent of large-scale solar power generation in residential areas, as well as integration of large-scale wind energy sources, contribute

even more to this security. A three-phase fault in such a stochastic scenario can easily cause protection system failure leading towards cascading outages. Thus, a renewably powered microgrid demands novel ways of energy dispatching methods [154]. Traditionally, under these conditions, a central station addresses the tertiary regulation that typically has an interval of 24 hours. However, in a modern microgrid, integration of distributed generators (DG) is compelling to adopt a more local, decentralized approach [106]. It is because long-term planning often is prone to large-scale errors leading towards service interruptions. The alternate approach thus would be to build short-term forecasting systems for a nation's dispatchable energy sources [155]. One of the most sought out methods in this field is the application of machine learning algorithms [76,81,156,157]. The neural network, as well as many ensemble techniques, are often discussed and implemented for such a purpose [76,81,156]. However, most study limits itself in analysing the accuracy of the algorithm by comparing actual and forecast data. Thus, very few studies have been conducted to forecast power system security which can take service restorative (SR) measures [158]. On the other hand, considerable efforts have already been made to improve SR plans by implementing multi agent-based systems (MAS), advanced metering infrastructure (AMI), knowledge-based systems, linear programming, progressive hedging (PH), etc [155,159,160]. These methods are computationally complex and often are not suitable for uncertain fault durations, involving integer decision variables such as a binary indication of the presence of rotor angle instability. This issue can be resolved by considering two types of uncertainties, i.e., a critical fault during vulnerable and non-vulnerable periods. These periods than can be further decomposed into multiple scenario-based stochastic programs [158]. However, such methods would demand accurate forecasting of the system security. The previous studies often ignored this integration and assumed that the local demands, generated from Distributed Resources (DR) are either updated or kept unchanged during a critical fault [161]. However, to minimize the service interruptions, the process of updating the data generation online, right after the fault is cleared, is quite important [162].

This study intends to bridge this gap between security forecasting and service restoration plans by implementing a novel architecture. The key argument of this study is to take advantage of heuristic search over the logic reasoning or empirical judgment [163], by implementing a predictive analytical method that forecasts the vulnerability of a system, if a critical three-phase short circuit fault occurs at that instance. The prediction is made based on the available wind power, solar power and the noncritical (controllable) loads. Machine learning driven optimization is used to implement an autonomous restoration scheme after a critical three-phase fault followed by loss of a generation. The proposed method considers the distributed generations and demands the prediction of system security. The security assessment is carried out within short intervals considering the possibilities of a major three-phase fault taking place shortly [164, 165]. Monte Carlo simulation is used to generate the POS database and train the proposed algorithm.

This analysis explores the idea that under different energy demand and distributed energy generation, the impact of a three-phase fault can either be critical or non-critical [164]. The proposed method figures out that criticality and prepares a set of optimized contingent scenarios for restoring the system. The algorithm is based on machine learning that manipulates an optimization platform to achieve lowest possible operating cost after ensuring voltage quality throughout the network in a post-fault contingency [165]. The database of POS is prepared by a Monte Carlo simulation method. The goal is achieved by implementing an ensemble of bagged decision tree-based systems to perform forecasting followed by a genetic algorithm (GA) for the service restoration. The security above is assessed by a binary security index Probability of Stability (POS) after an event of a short circuit fault resulting in isolating a generator bus at the affected area [27]. The database of POS is prepared by a Monte Carlo simulation method.

A vulnerability analysis should have a hierarchical structure with a primary goal of protection and a secondary goal of economic restoration.

The application of machine learning algorithms in power system is not a novel concept, and

there have been several success stories. In reference [166], applications of an artificial neural network (ANN) is reviewed that solves various energy-related problems. Problems, such as modelling and designing a solar steam generation plant, predicting the local concentration ratio, finding intercept factors of parabolic-trough collectors, predicting performances of solar water-heating system, etc. The application of neural networks also includes diagnosis of faults, forecasting loads, tuning the stabilizers, modelling machines, planning expansions, etc. The strengths and abilities of a neural network based system vary regarding pattern classification, response speed, generalization and prediction of trends, etc. [76, 167–169].

Data-driven control is the main theme in this work. One of the usages of the forecasted data is to predict severity and take necessary measures to maintain stability, which falls under a classification problem [164], and for the classification task here a genetic algorithm driven random forest is used. Different key operational strategies have been combined for running this method. These strategies are addressed in a cascaded scheme to achieve the final goal of predicting the mode of operation of multiple generators. The modes are namely ‘ON’, ‘OFF’ or ‘Continuous’ under different forecasted data clusters. The mode of operation is a direct result of optimized decision-making process where the final objective is to restore the subject microgrid, which is operating in a stand-alone mode after a major three-phase fault has occurred.

## 8.2 The Microgrid Model

The proposed microgrid as shown in Figure-8.1, has one hydro turbine based synchronous generator, two backup diesel generators (synchronous) **G1=4MW** and **G2=3MW**, one asynchronous generator representing the wind farm and a voltage source converter based solar power plant. The loads are lumped on a common transmission grid. This distribution approach is inspired by the microgrids used in reference [99, 159]. However, this model differs in dividing the loads into two parts; critical and noncritical loads [170]. The critical loads are comparable to the base load of

a system that remains unchanged throughout a certain period. While the controllable loads are unpredictable in nature and sudden changes can be observed at any given time. In this study the residential loads have been considered critical and noncritical thus load shedding actions can be applied to it. The total demand in this microgrid is higher than the total capacity of the synchronous generator models used as power plants. It signifies that the power quality and stability of the system time to time depends on the wind power plant and the solar power plant. The wind power plant is modelled as an induction generator-based variable speed wind turbine. The solar power plant is modelled as a current source and placed closer to the residential load. The three-phase model has diodes, internal resistance, and leakage current followed by a voltage source converter (VSC) as presented in reference [101]. In this thesis, the VSC based solar plant is only implemented as an intermittent energy source. These two distributed and intermittent sources do not have any power system stabilizers installed in them thus maintaining stability in the system is carried out through the synchronous generator-based models. To capture the full dynamics of bi-directional energy flow both the solar and wind turbine plants are designed not to have any energy storage device.

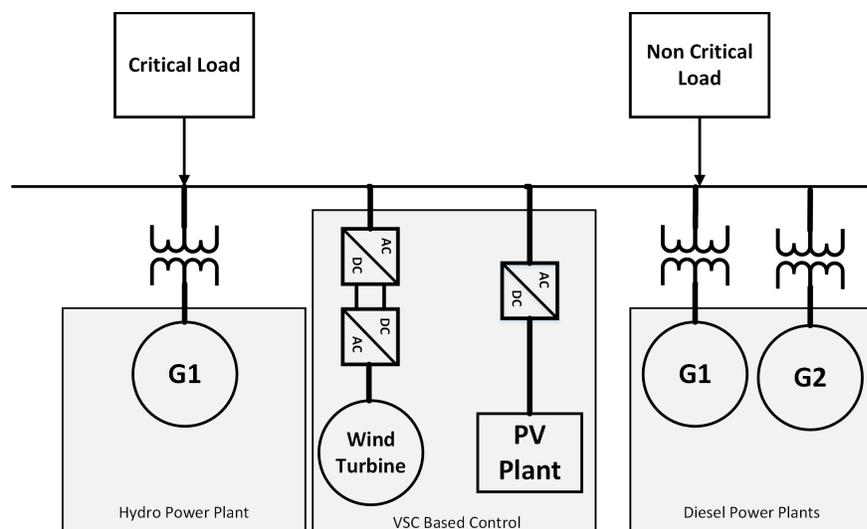


Figure 8.1: The Microgrid Model

The generator buses are represented using the typical second order swing-equation;

$$M_i \ddot{\delta}_i = P_{mi} - P_{gi}; i \in generator_{1:3} \quad (8.1)$$

Here,  $\delta_i$  is the generator rotor angle,  $P_{mi}$  is the mechanical power input,  $P_{gi}$  is the electrical power output,  $M_i$  generator's inertia coefficient and  $D_i$  is the generator's damping coefficient. The overall operation is subject to;

$$P_{gi} - P_{li} - \sum_{j=1}^3 U_i U_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) = 0, \quad (8.2)$$

$$Q_{gi} - Q_{li} - \sum_{j=1}^3 U_i U_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) = 0, \quad (8.3)$$

To test the security of the system and develop the POS database, a critical three phase fault is placed near the hydro station.

Two scenarios are established based on whether the fault is cleared within the critical clearing time or not. Depending on that the post fault response is either stable or unstable. The following part of this article is based on the scenario where the fault is not cleared within its critical clearing time. If the power system stabilizer installed in the hydro plant fails to remove the rotor angle, instability arises right after the three-phase fault generator bus is disconnected. In Figure-8.2 a comparison between the two scenarios with a time delay is shown. The model accommodates the concept of *failure to operate* as protection system failure. The *failure to operate* mode is observed in the power system stabilizer installed in the hydro plant. This mode also ensures loss of a critical load connected to that bus.

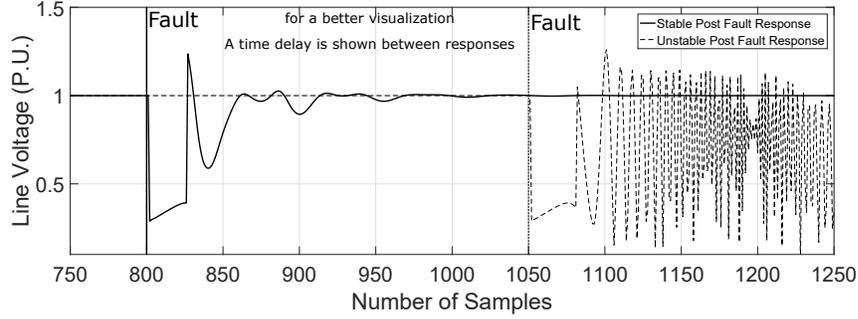


Figure 8.2: Stable and Unstable Post Fault Responses

## 8.3 Data Model

To build up the forecasting system three data models have been prepared. Once the data model has been prepared, a random data set of  $N$  number of data points is generated and fed to the aforementioned microgrid model in Figure-8.1 in order to carry out the Monte Carlo simulation for developing the Probability of Stability (POS) table by introducing the three-phase fault; here  $N = 10000$ . Figure-8.3 shows the application of the data model in this study. The following sections discuss the development of different data models.

### 8.3.1 Wind Energy Model

While preparing the wind turbine generating model, certain conditions have been considered such as: the model only takes wind speed as input variable, the velocity of wind is uniformly distributed on the surface of each blade and the air density is constant during each calculation period. The fundamental equation to calculate the active power from wind velocity is;

$$P_W = C_{total} \frac{1}{2} \rho A v^3 \quad (8.4)$$

where,  $P_W$  is active power output,  $C_{total}$  is overall efficiency of the wind turbine,  $\rho$  is air density,

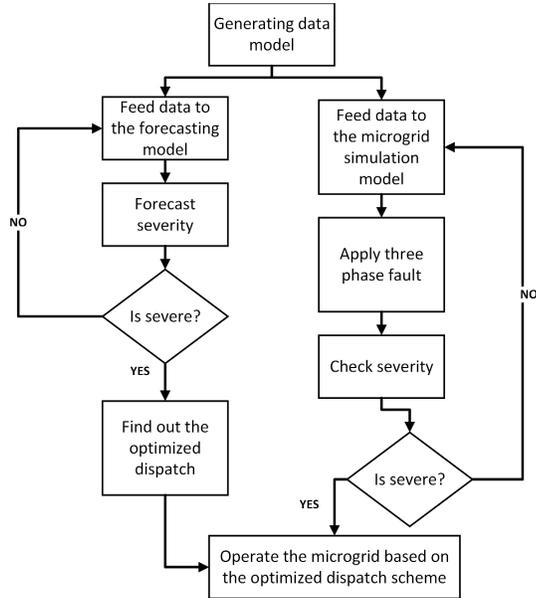


Figure 8.3: Implementing the Data Model into the Proposed microgrid

$A$  is swept area and  $v$  is the wind velocity. The wind velocity at any certain height is calculated by  $v_h = v_r \left(\frac{h}{h_r}\right)^\alpha$  and  $\alpha$  is the power law exponent. Where,  $v_h$  is the speed at hub height,  $v_r$  is the speed at reference height.

In some of the earlier research the wind velocity has also been considered as a function of temperature [76]. Further considerations are made;

$$P_W = 0; v(t) < v_{ci} \quad (8.5)$$

$$P_W = 0; v(t) > v_{co} \quad (8.6)$$

where,  $v_{ci}$  is cut-in speed and  $v_{co}$  is cut-out speed. The energy model used in this study is a 3.8 MW wind farm driven by wind velocity as input information. The data model is influenced by the Hau Nui Wind Farm in Christchurch, New Zealand; the model has a cut in speed of  $3m/s$  [171,172].

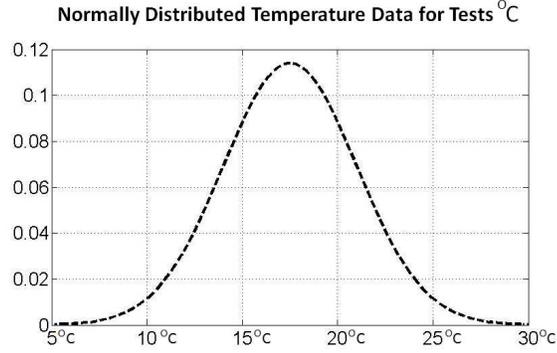


Figure 8.4: Normal probability distribution of temperature (Degree Celsius)

### 8.3.2 Solar Energy Model (Photovoltaic)

To simulate the characteristics of solar power generation the following irradiance based photovoltaic cell model of a solar cell is used;

$$I = I_{ph} - I_s \left[ e^{\frac{V_{OC} + IR_s}{N_1 V_t}} - 1 \right] - I_{s2} \left[ e^{\frac{V_{OC} + IR_s}{N_2 V_t}} - 1 \right] - \frac{V_{OC} + IR_s}{R_p} \quad (8.7)$$

$$V_{OC}(t, \beta) = V_{OC-STC} - K_V T_C(t) \quad (8.8)$$

where  $V_{OC}$  is the open circuit voltage of the PV-module,  $I_{ph}$  is the solar-induced current that can be further explained by  $I_{ph} = I_{ph0} \frac{I_r}{I_{r0}}$   $I_r$  is the irradiance in  $W/m^2$ ,  $I_{ph0}$  is the solar current obtained for irradiance  $I_{r0}$ ;  $I_s$  and  $I_{s2}$  are the saturation currents of the Diode-1 and Diode-2 inside;  $N_1$  and  $N_2$  are the quality factors diodes;  $V_t = \frac{kT}{q}$  is the thermal voltage, ( $k$  is Boltzmann constant,  $T_C$  is device temperature in Kelvin) and  $K_V$  is the open circuit voltage temperature coefficient;  $T_C = T_A + (NOCT - 20 \text{ deg}) \frac{I_r(t, \beta)}{800}$ .  $R_s$  and  $R_p$  are the series and parallel resistances [77].  $\beta$  is the tilt angle and  $T_A$  is the ambient temperature. The overall output power from the plant is given as;  $P_{array}(t, \beta) = \eta_{PV} N_S N_P P_{PV}(t, \beta)$ . Here,  $N_S$  and  $N_P$  are the total number of modules

connected in series and parallel,  $\eta$  is the conversion efficiency.  $P_{PV} = V_{OC}I$ ; is the instantaneous power output from each PV-module. The solar power plant is designed to have a maximum of 6 MW power capacity.

### 8.3.3 Electrical Load Model

The electrical load model used in this study is a function of the base load as well as the ambient temperature. The ambient temperature influences the residential controllable loads while it is considered to have no impact on the critical uncontrollable loads. The impact of temperature on the residential load is modelled as a 3<sup>rd</sup> order polynomial system. [84] shown in Figure-8.5.

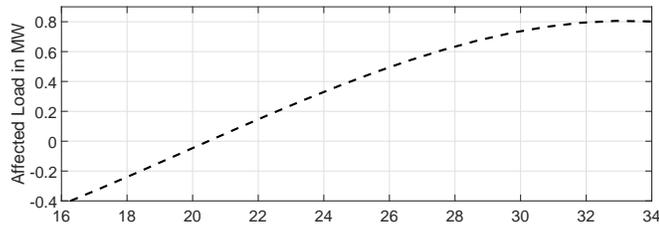


Figure 8.5: Influence of the Ambient Temperature on the Controllable Load

The controllable load model is assumed in MW as;

$$P_{Load} = f(T) = B_{CL} + a_0 + a_1T + a_2T^2 + a_3T^3 \quad (8.9)$$

where  $B_{CL}$  is the base controllable load that is not influenced by temperature,  $a$  = multiplying constants,  $T$  = temperature in degree Celsius. The coefficients are  $a_0, a_1, a_2, a_3 = -0.5632, -1.185e^{-1}, 1.09e^{-2},$  The overall controllable residential load model is shown in Figure-8.6.

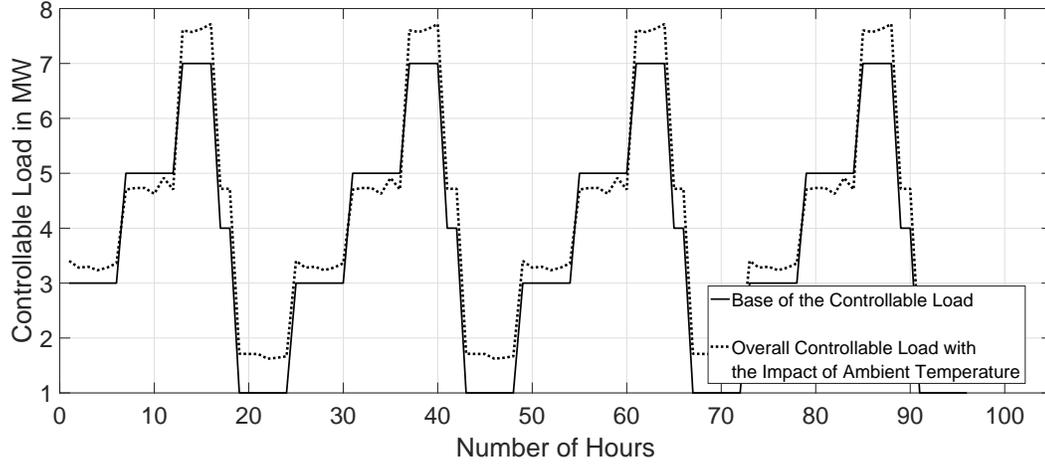


Figure 8.6: Controllable Load Model

### 8.3.4 Security Index Probability of Stability

Due to the possibilities of having numerous post-disturbance initial conditions, generalizing *service restoration* plans is quite a challenge [163]. This study thus assumes that the post-fault initial conditions only differ from the pre-fault operational conditions in the topological layer and in the form of bus isolation at the affected segment. However, all the other available parameters are kept similar. The start-up time for the backup diesel generators is considered negligible. Even so, the possibilities are endless. To address this issue a Monte Carlo simulation-based approach is taken to prepare the POS database [173]. Wind, solar and consumption data have been randomly selected within a predefined range and fed to the simulation model to create the events. In every simulated event, a three-phase fault is introduced and cleared after the **critical clearing time**, which is observed **300** milliseconds for the proposed microgrid. Depending on the energy flow in the system, a rotor angle instability is developed. As the stochastic data of wind, solar and load can produce infinite combinations, a k-means clustering approach is taken to prepare finite data segments. The POS is then calculated based on each of the data segments or clusters as;  $POS_k = \sum i N_k \frac{S_{ik}}{N_k}$ . Here,  $k$  is the current data cluster,  $N_k$  is the total number of simulations in each cluster.  $S_{ik}$  is **1** if the

system becomes unstable and **0** if the system remains stable. Overall **10000** simulation is carried out to develop the POS for 0, 1, 2... $k$  number of clusters of the system [165, 174]. The POS table is shown in Table-8.1.

Table 8.1: Data Table With Clusters Centroids

Wind Power MW	Solar Power MW	Controllable Load MW	Cluster	$POS_k$
1.73	3.47	8.31	1	16.00%
2.77	5.55	9.7	2	64%
...	...	...	...	...
...	...	...	...	...
2.17	4.3	9	45	80%
0.6	1.2	6.8	81	0.0%
2.93	5.87	9.91	96	48%

The overall POS under a critical three-phase fault is represented in Figure-8.7 using a kernel density function due to the finite data sample used for simulation. Due to the smoothing operation, negative and more than 100% probability is observed within the maximum confidence interval. However, the proposed **100** data clusters are of different data size. Thus the method is implemented recursively, in every iteration eliminating a data cluster having more than **100** data points.

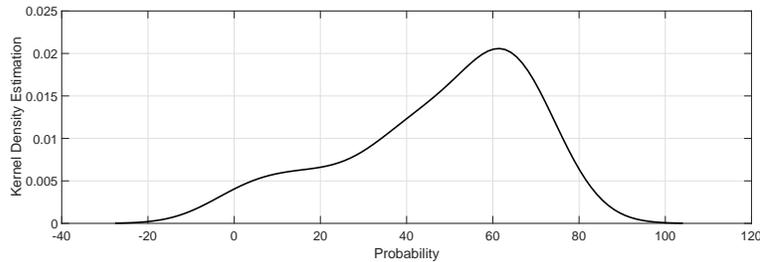


Figure 8.7: Kernel Density Estimation

### 8.3.5 Data for Fault Analysis

The proposed microgrid considers four instances of the line voltage; pre-fault line voltage  $V_{pre}$ , during fault line voltage  $V_{flt}$ , line voltage during the post-fault restoration  $V_{rst}$  and line voltage af-

ter the post-fault restoration  $V_{post}$ . This study aims to accurately predict system operation strategies that will lead towards achieving a stable and optimal post-fault line voltage. The transition period from the fault state towards the restoration state has been considered similar for all the operational strategies. Incidents such as starting up of a backup generator are considered under the conditions when energy balance cannot be made.  $\sum_{i=1}^{n_g} pg_i + \sum_{i=1}^{n_{ren}} pre_n_i \leq P_{dpfcl} + P_{dpfncl} + P_{Tloss}$ ; where,  $P_{dpfcl}$  is post-fault critical load,  $P_{dpfncl}$  is the post-fault non critical load,  $P_{Tloss}$  is the transmission line loss,  $pre_n_i$  power generated from the renewable energy sources,  $n_g$  is number of hydro generators,  $n_{ren}$  is the number of renewable energy sources. Under such conditions, post-fault initial states are re-evaluated, and the operation scheme initiates a temporal restoration process comprising of **30 to 60** minutes. The fault location is set near the generator mostly catering to the base load. The generator along with the nearby load is considered disconnected in the process of clearing the fault. Depending on different times of the day, the loss of load in the generator bus is considered within a range of **6 MW to 10 MW**.

## 8.4 Classification Process for Optimal Voltage

After preparing the forecasting models, an unsupervised K-means cluster has been implemented. The purpose of using an unsupervised cluster is to narrow down the target data before preparing other machine learning algorithms. Unsupervised clustering would make sure that the data falls under one category at a time. Here twelve predefined categories have been chosen for the clusters. These twelve clusters have been selected from the combinations of: high and low wind power, high and low solar power and high, medium and low residential loads, shown in

The POS data are shown in Table-8.1 is used to initiate the optimal decision-making process. While testing the proposed algorithm, random instances for wind power, solar power, and noncritical load have been chosen and then compared with the POS above table. If the instance matches to any of the clusters having a probability of higher than 5% of being unstable have been considered

as a potential candidate for the optimized restoration analysis. The unsupervised ‘K-means’ cluster discussed earlier also minimizes the leaf nodes of the classification tree used. Each cluster or data segment is associated to one of the sixteen predefined restoration actions for the case of loss of generation implemented in this study. The sixteen actions are defined in Table-8.2. The actions are described through a set of binary numbers, where for the backup generators **1** means *start* and **0** means *do not start*. While for the non critical loads **1** means *shedding loads* and **0** means *not shedding loads*. The diesel generators are presented as to the swing equation model where the capacities are subject to;  $P_{gimin} \leq P_{gi} \leq P_{gimax}$ ,  $Q_{gimin} \leq Q_{gimax}$ ,  $0 \leq P_{li} \leq P_{di}$ ,  $0 \leq Q_{li} \leq Q_{di}$ ,  $V_{tmin} \leq V_{tmax}$ . Here,  $P_{li}$  and  $Q_{li}$  are the load after system restoration and  $P_{di}$  and  $Q_{di}$  are the actual active and reactive demand,  $V_t$  is the generator terminal voltage. So the overall load shedding under any of the sixteen actions is  $P_{di} - P_{li}$ .

Table 8.2: Sixteen Probable Restoration Schedules

Decision	Start Diesel-1	Start Diesel-2	Shed 1.5 MW	Shed 2.5 MW
1	0	0	0	0
2	0	0	0	1
3	0	0	1	0
4	0	0	1	1
..	..	..	..	..
..	..	..	..	..
5	0	1	0	0
6	0	1	0	1
7	0	1	1	0
8	0	1	1	1
9	1	0	0	0
10	1	0	0	1
11	1	0	1	0
12	1	0	1	1
13	1	1	0	0
14	1	1	0	1
15	1	1	1	0
16	1	1	1	1

The post-fault restoration executes one of these sixteen actions. The difference between the

diesel generator-1 and the diesel generator-2 is that the diesel generator-1 has higher rated power than generator-2; **4 MW** capacity compares to that of **3 MW** capacity. While restoring the system, it has been considered. The primary objective is to maintain an equal pre and post-fault voltage state. The actions above have been thus evaluated against this objective of maintaining system-wide desired voltage level. The objective was achieved by implementing an exhaustive search-based optimization technique. Here a total of **9600** cases have been simulated to observe the impact of the all sixteen actions.

$$Objective : \min(\epsilon_V = \sum_{i=1}^N [V_{prefault} - V_{postfault}(i)]^2) \quad (8.10)$$

Here  $i$  is the data point in consideration.  $N$  is the total number of data points considered once the fault is cleared.  $V_{prefault}$  is the stable line voltage before the fault and  $V_{postfault}$  is the post-fault stable line voltage. The action with the lowest voltage deviation is the optimized target decision for the classification algorithm.

The third step is to prepare a genetic algorithm-based system to make the optimized decision for the above mentioned twelve clusters. The purpose of the decision-making process is to select one of any six post-fault conditions or actions for system restoration before the base generator can be integrated again.

Table 8.3: Data Table With Clusters Centroids

Number	Action
1	Start two diesel generators without load shed
2	Start two diesel generators with load shed
3	Start the diesel generator-1 without load shed
4	Start the diesel generator-1 with load shed
5	Start the diesel generator-2 without load shed
6	Start the diesel generator-2 with load shed

### 8.4.1 Preparing the Classification Tree

Once all the optimized restoration actions have been identified for the **9600** data points, an ensemble of bagged decision trees has been trained. In Figure-8.8 one limited instance of the ensemble tree is shown;

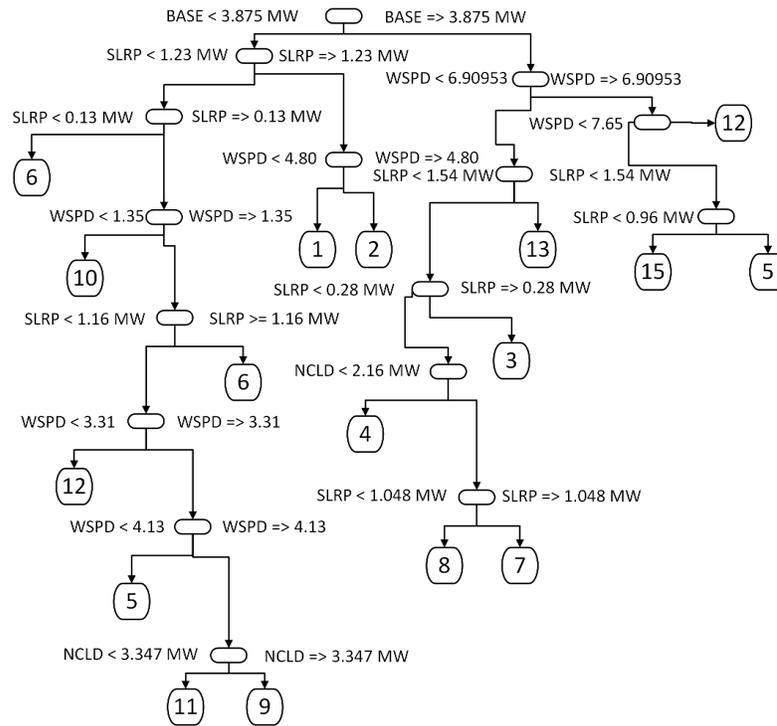


Figure 8.8: The Decision Tree Classifier (Limited Scenarios)

However, the classification algorithm is not immune to classification errors [26]. As a microgrid operated at its design limit is highly vulnerable any misclassification may lead towards unstable post fault operation. Therefore, a decision ranking system has been proposed in this study to prevent the probability of system instability due to misclassification. The ranking system is a cluster of decision hierarchy that has been prepared from Table-8.1 and Table-8.2 based on the voltage deviation  $\epsilon_V$  observed previously. In this study thirteen different combinations have been identified in the top

five decision hierarchy. Based on these combinations a hierarchical cluster is prepared to validate any predicted decision by the ensemble tree. Table-8.4 shows the ranking hierarchical cluster.

Table 8.4: The Hierarchical Ranking Table

1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	HCluster
5	4	9	7	11	1
9	7	11	6	10	2
..	..	..	..	..	..
..	..	..	..	..	..
1	3	7	11	10	3
11	9	5	10	8	4
7	6	11	9	10	5
5	10	7	11	6	6
12	8	10	9	13	7
8	11	9	6	12	8
15	14	16	13	12	9
13	15	14	16	12	10
9	5	2	1	6	11
7	3	13	2	15	12
13	15	14	12	16	13

The dendrogram prepared from the hierarchical cluster is shown in Figure-8.9

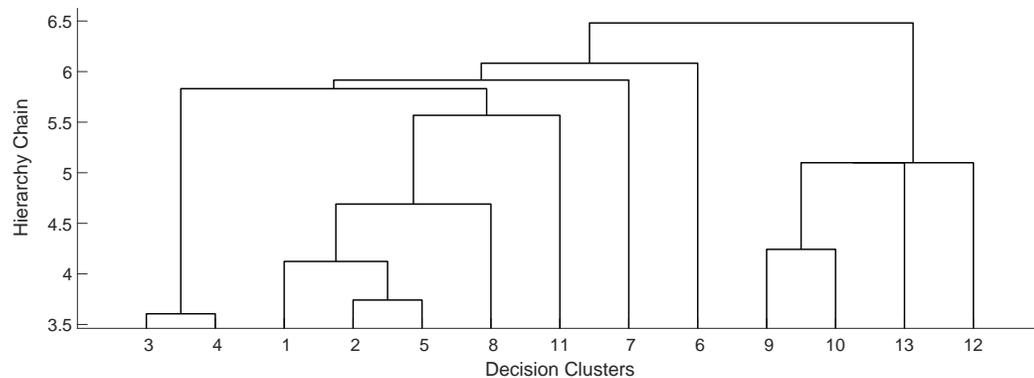


Figure 8.9: Hierarchical Cluster

The hierarchical cluster indicates the coherency among different decision clusters. Three additional ensemble of bagged decision trees have been trained with the identical data set with two different target decisions. Decisions having the Ranks 2<sup>nd</sup>, 3<sup>rd</sup>. Then a temporary decision set is

prepared  $D1^{st}, D2^{nd}, D3^{rd}$ . The decision set is then compared with the hierarchical ranking Table (Table-8.4). Where  $N_P^{th}$  stands for the probable predicted decision,  $N_{Ri}^{th}$  is the combination of decisions observed in the ranking table from historical data and  $HC$  stands for the hierarchical clusters.

$$D_{SETi} = \{1_P^{st} \cap 1_{Ri}^{st}, 2_P^{nd} \cap 2_{Ri}^{nd}, 3_P^{rd} \cap 3_{Ri}^{rd}\}; i \in HC_{1:13} \quad (8.11)$$

$D_{SETi}$  is then used for validation. Three conditions have been observed in the validation process;

1. If all three decisions match to one of the combination presented in the ranking table the overall prediction is considered accurate and the hierarchical cluster is preserved as the predicted cluster.
2. If top two decisions match then the the hierarchical cluster is preserved as the predicted cluster.
3. If the top decision and the bottom decision match the hierarchical cluster is preserved as the predicted cluster.

Apart from these three conditions any other condition is considered erroneous and the process is repeated.

## 8.4.2 Process for Economic Dispatch

The secondary objective of the optimized grid restoration strategy is economic dispatch (ED). The objective function for the ED is to minimize total generation cost of the microgrid.

$$\min : \sum_t \sum_{i \in G} C_i(pg_{i,t}) + \sum_t \sum_{k \in D} C_k(pd_{k,t}); \forall t \in T \quad (8.12)$$

where,  $C_i(pg_{i,t})$  is the cost function of the  $i$ -th generator at  $t$ -th half hour.  $C_i$  is formulated as a quadratic function  $C_i(pg_{i,t}) = a_{0,i} + a_{1,i}pg_{i,t} + a_{2,i}pg_{i,t}^2$ , where,  $a_0$ ,  $a_1$  and  $a_2$  are the coefficients. A similar approach has also be taken for the cost of shedding load  $C_i(pd_{i,t})$ . The cost function is subject to the inequality constraints at any given time after the fault has been cleared;

$$P_{dpfcl} + P_{dpfncl} + P_{Tloss} \leq \sum_{i=1}^{n_g} pg_i + \sum_{i=1}^{n_{ren}} preni \quad (8.13)$$

where  $P_{dpfcl}$  and  $P_{dpfncl}$  are the critical and non-critical loads. Due to sudden change in demands, in some cases deviation in power generation can be observed. In this study the post fault demands are kept constant and any large deviation is neglected. Thus  $pg_{i,t} - pg_i, t0$  is considered zero for the half an hour span pre-defined for the stochastic data points. Therefore, the upper ramp rate  $UR_i$  and lower ramp rate  $LR_i$  have been neglected. To maintain post fault frequency stability, sufficient spinning reserve should be available.

$$\sum_{i \in G} SR_{i,t} \geq SSR_t; \forall t \in T \quad (8.14)$$

where,  $SR_{i,t}$  is the available spinning reserve of individual diesel generator at  $t$ -th half hour, and  $SSR_t$  is the system wide required spinning reserve. Each generators also follows the generator output constraints, which means the generation does not exceed its upper limit  $pg_{i,t} \leq PG_{i,max}; \forall t \in T$  and also for the renewable energy generators  $pgren_{i,t} \leq PGren_{i,max}; \forall t \in T$  [175, 176].

For solving the economic dispatch problem, genetic algorithm (GA) is used. Here the GA applies **five** steps: population initialization, evaluation, selection, crossover, and mutation. In the proposed method the GA has generations of **51**, the problem type is defined as **linearconstraints**, and the number of evaluation of the fitness function is carried out **10201**. The dataset has been pre-processed by the ensemble of bagged decision trees mentioned earlier. The classifier, based on the stochastic data of POS introduces the upper boundary for the design variables namely, dispatch

from two diesel generators and shedding two loads. The upper boundary is prepared based on the decision table (Table-8.2). For example if the classifier takes *Decision-13*, then the two diesel generators will be switched on under the condition of  $0 \leq pg_{i,t} \leq PG_{i,max}; \forall t \in T$  and no load shedding will take place thus,  $pd_{i,t} = 0; \forall t \in T$  and  $i \in \{1, 2\}$  [145].

The overall architecture of the proposed algorithm is shown in Figure-8.10. The algorithm has two parts. Offline detection and mitigation of unstable use cases is presented on the left hand faction of the figure. On the right hand side, the online application phase is shown. One of the key reasons to select this architecture is to gradually train the proposed classification algorithm. Once the instability is spotted several consecutive steps are taken to generate a data set that contains the optimized decisions. The algorithm is trained with these optimal decisions. The offline stage also collects data and stores it. Depending on the detection of instability the system data is stored in different locations and used appropriately in the online stage.

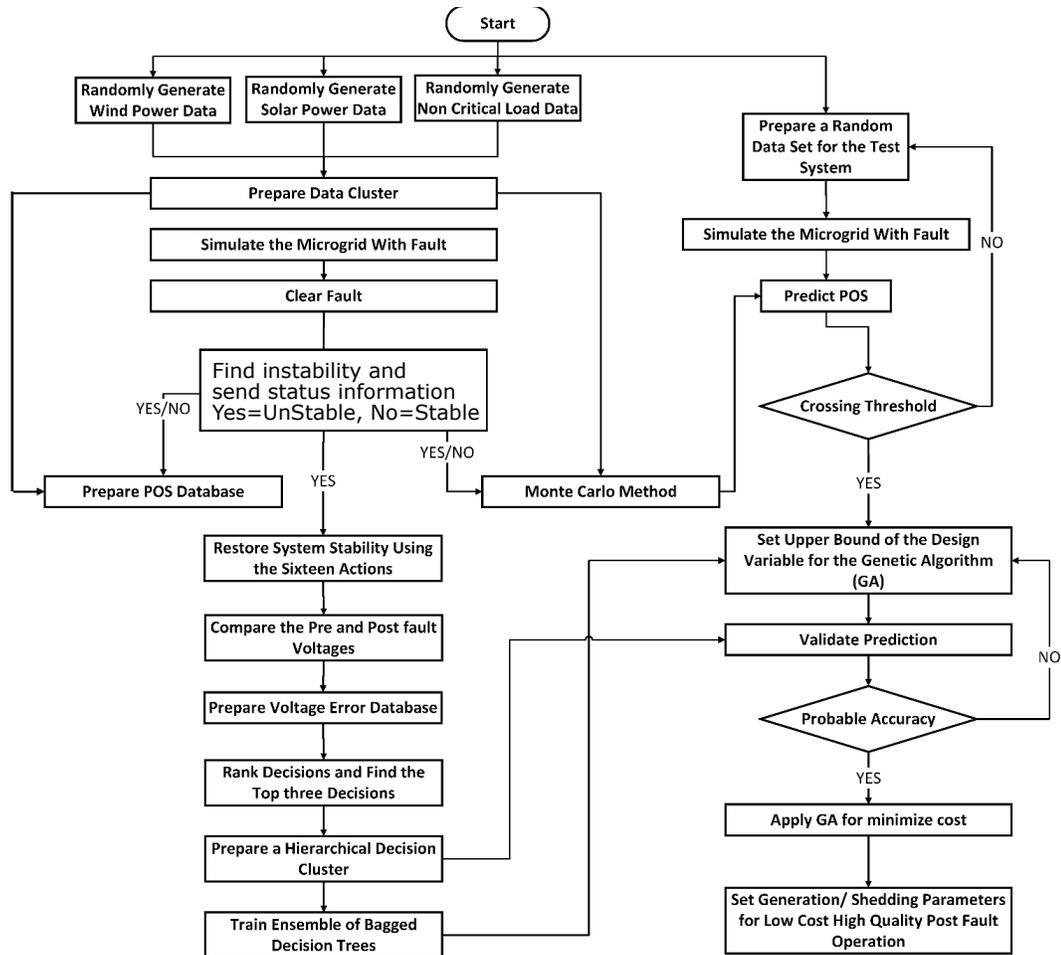


Figure 8.10: Work-Flow of the Proposed Algorithm

## 8.5 Results

### 8.5.1 Estimation of Instability

The application of Monte Carlo method is opening a scope of interpreting the underlying post-fault scenario. The method relative to each data point presents unique characteristics which either

leads the system towards instability or the system remains resilient. The Monte Carlo simulation is carried out with **10000** data points. Out of the **100** data clusters, **6**, randomly selected clusters have been chosen to analyse the overall outcome of this study.

Figure-8.11 shows the process of estimating POS of all the proposed clusters. It is based on multiple simulated scenarios under one cluster. Each simulation result is added up and the average of POS is updated gradually till it converges towards obtaining the final POS data.

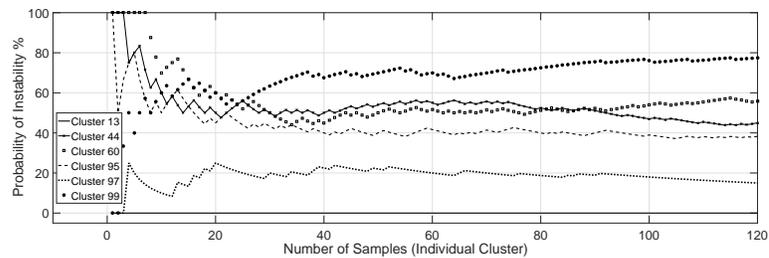


Figure 8.11: POS of the **100** Clusters and the Monte Carlo method of convergence

Figure-8.12 shows the POS of all the proposed clusters. Once the convergence is observed, the POS data is stored in a table against each of the clusters. The clusters having a very lower probability of being unstable has not been considered for further analysis.

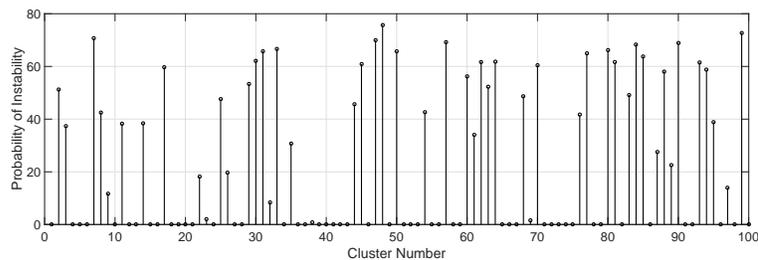


Figure 8.12: POS of the Clusters

Any data cluster having a low POS is ignored in the process of training the ensemble of bagged decision tree. For example *POS cluster-13* has a zero probability of being unstable. Therefore, no restorative measure has been taken for all the instances in *POS cluster-13*.

## 8.5.2 Selection of Restorative Action

The proposed ensemble of bagged decision tree method is invoked if the threshold probability of 5% is crossed. Figure-8.13 demonstrates the result of the classification algorithms in finding out the 1<sup>st</sup>, the 2<sup>nd</sup> and the 3<sup>rd</sup> top decisions with the lowest possible errors. The top decision database is used to control the upper bound of the GA.

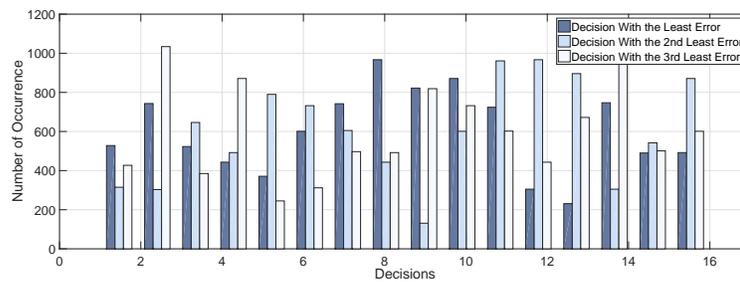


Figure 8.13: Top Three Decisions after 10k simulations

The GA is applied to observe the lowest possible operational cost. The results are explained with the previously mentioned six POS clusters. The key observation made in this study is that the most economical dispatch under duress is not always producing the best possible voltage profile. Thus, the machine learning algorithm is implemented to find out the suitable initial conditions for the GA that will maintain post-fault system stability. Table-8.5 shows the three different possibilities while deciding and Table-8.6 shows the comparisons between the operational costs. The comparison is made with actions only considering economic dispatch, the voltage quality and the proposed method.

Table 8.5: Three Different Actions

POS Cluster	13	44	60	95	97	99
Decision With the Lowest Cost but poor voltage profile	NA	11	1	8	10	10
Decision With the Best Voltage Profile	NA	5	3	12	5	7
Acceptable Voltage Profile at Lower Cost	NA	9	3	9	11	9

Table 8.6: Normalized Operational Costs

POS Cluster	13	44	60	95	97	99
Decision With the Lowest Cost	NA	0	0	0	0	0
Decision With the Best Voltage Profile	NA	1	1	1	1	1
Acceptable Voltage Profile at Lower Cost	NA	0.96	1	0.62	0.89	0.13

Depending on the available wind, solar power and controllable loads the actions mentioned above can either be different or can be overlapping. Figure-8.14 and Figure-8.15 elaborately shows all three decisions and their impacts on the **SIX** clusters. In *cluster-13* the probability of system instability is **zero**. Thus, the post-fault restorative actions are not invoked. For the other clusters like *cluster-44*, the algorithm is applied. In all the cases the decisions with the lowest cost have the highest voltage deviation. It is a clear indication that GA for cost optimization alone is not sufficient to maintain voltage quality. On the other hand, the exhaustive optimization method for maintaining post-fault voltage quality based on the decision hierarchy recommended decisions have the highest operational costs. The proposed algorithm is mitigating these two extremes by taking appropriate actions that produce stable output at a lower cost.

## 8.6 Analyzing the Limitations

This study demonstrates a probabilistic optimization model for restoring a power system after a major three-phase fault. The underlying objective of managing resource under a contingent scenario, with a machine learning based system is analysed and presented. The standard GA based optimization system is manipulated using the machine learning-based classifier, and comprehensive results are achieved.

The control strategies adopted in this study is grid specific. Therefore, data pre-processing plays a vital role in this study. The analyses advocates considering different approaches for different systems. However, the Monte Carlo simulation-based approach is implemented in this study to

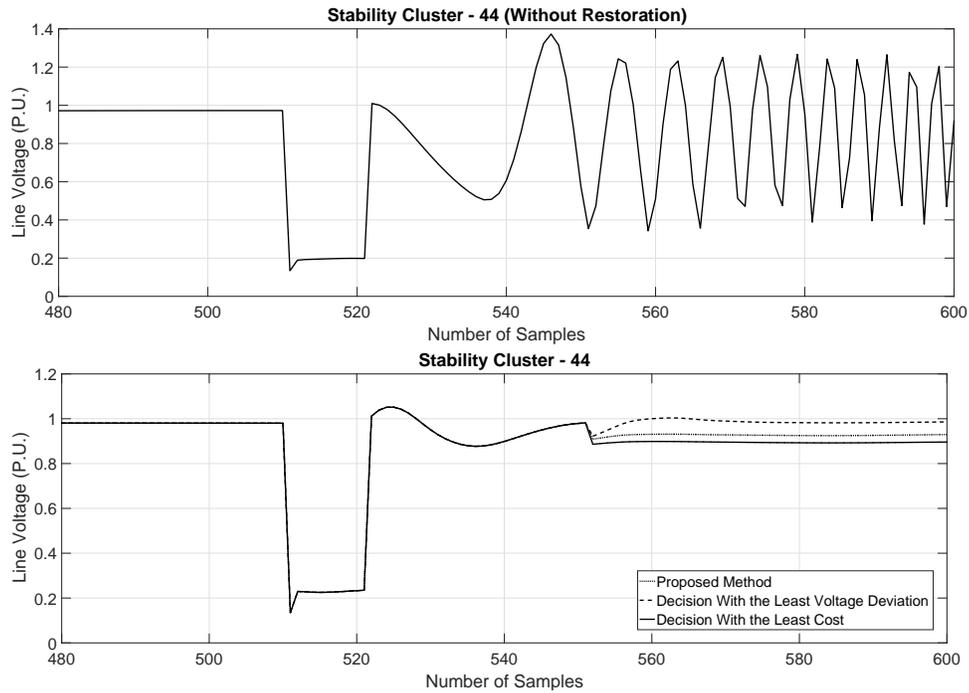


Figure 8.14: Results With and Without the Proposed Method

understand the feasibility of developing a data-driven generic restoration strategy. The challenge is to create a bridge between the finite numerical simulation and the infinite post-fault scenarios.

The proposed classifier also considers classification errors and simplified rectification techniques. The rectification process is an important factor while considering a machine learning algorithm for maintaining a system-wide stable operation under all the stochastic scenarios. The method for rectification implemented in this study was comprehensive. However, whether such a simplified technique is adequate for a larger context needs to be evaluated. Also, further investigations need to be carried out to understand the full ramifications of misclassifying data.

The performance of the GA is quite satisfactory for the proposed microgrid, and the contingencies analysed. More scenarios would help to compare and to understand the performance of the GA optimization method under other contingencies, not considered in this study. Overall it can be stated that the proposed Monte Carlo simulation and machine learning algorithm driven optimization model, for restoring a stand-alone microgrid after a major short circuit fault, is demonstrating

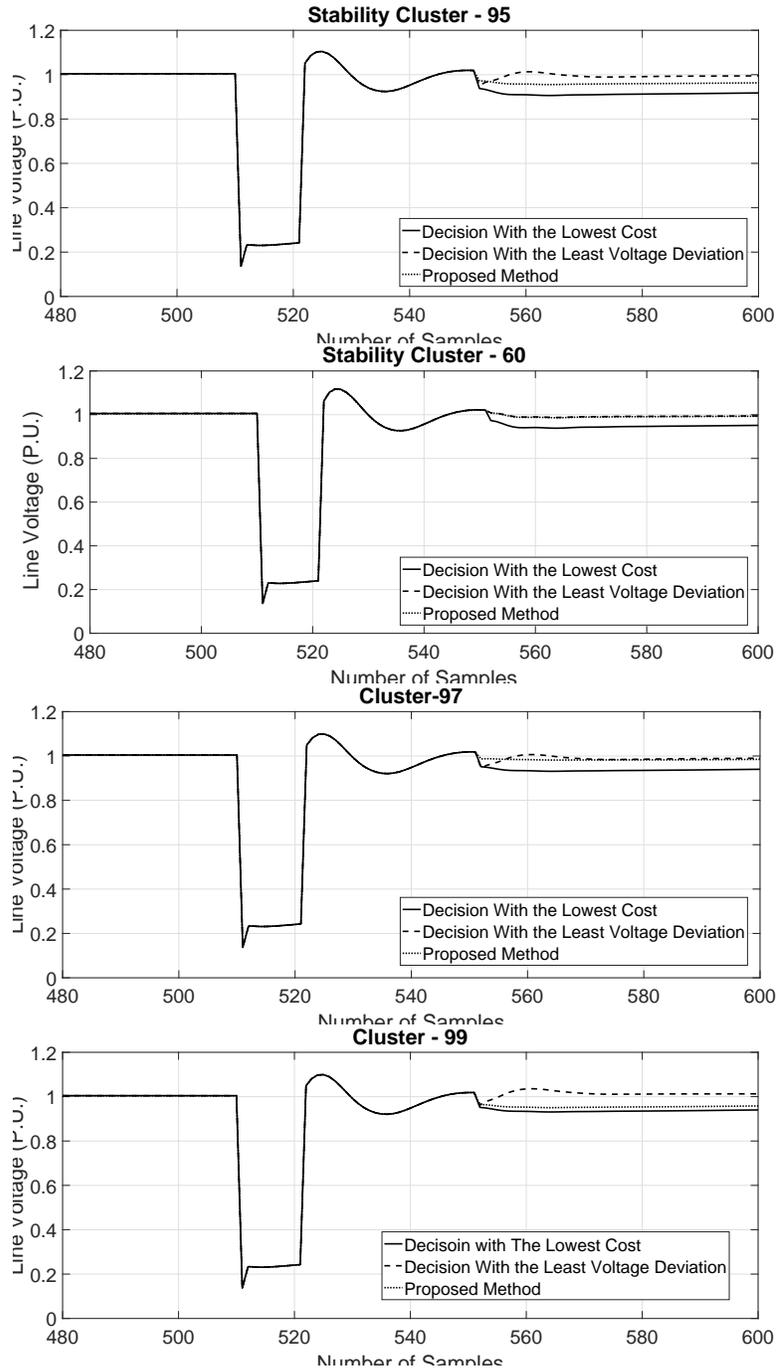


Figure 8.15: Different Decision Making Processes and Impacts (The Remaining Clusters)

a very promising outcome.

# Chapter 9

## Conclusion and Future Work

### 9.1 Concluding Remarks

Machine learning is a well-established domain in much applied scientific research. It has the capability of tracing out meaningful patterns in a given dataset, which in turn can help make critical decisions. However, traditional cleaning, training, and testing-based applications of machine learning cannot succeed in many scenarios as machine learning also has its limitations. This work investigated such use-cases, where a traditional application of machine learning algorithms is not meaningful. In a sequence of five different case studies, this research established a novel ensemble of machine learning algorithms to solve complex power system problems. It has also recommended different machine learning platforms for different types of use-cases.

The investigations led to an understanding that, power system data, especially the dynamic power system data, has many different characteristics. Exploring those characteristics helps a machine learning platform to improve accuracy and efficiency. The purpose of this study is to adopt a point of view that is application-centric. The applications had to be meaningful and complex enough for intuitive decision-making. Such, applications have been sought out and successfully resolved by applying machine learning algorithms. Several areas where the implementation of

machine learning based systems is apparent have been described in this thesis.

While addressing different data preparation techniques, it became evident that cross-disciplinary approaches are mandatory in a competitive environment. However, the selection of appropriate interdisciplinary approaches also depends on case study or events. A critical observation is, therefore, required for understanding the suitability of an algorithm in a particular scenario. In several case studies, this thesis has established such an approach. It is also advocated that a certain degree of scepticism is in-fact quite rewarding. In Chapter-3 **Stochastic Data model**, it is discussed that machine learning algorithms are not a good candidate to solve any and every problem in power systems. Especially for low-level machine control in real-time scenarios, the performance of machine learning algorithms is not at all overwhelming. However, as shown in the later chapters of this thesis, in a complex and sequential decision-making scenario, machine learning based platforms excel. In a nutshell, this is the major contribution of this work.

While the performance of the proposed ensemble methods is commendable, it is also important to keep in mind that power system operation is also influenced by other acts. Factors from other engineering faculties such as database management, communication channels and networking can affect the flow of data. Those factors can also affect the characteristics or the so-called 'features' used in this study. Therefore, it will be interesting to further incorporate other factors while analysing any system phenomenon and making a valid decision. Overall, before summarizing this thesis, it is imperative to state that, influences on power system data from other disciplines such as database management, communication, etc. have been ignored. Only the influences from system dynamics, data preparation techniques, and different algorithms have been considered.

In summary, a data-driven decision-making process has its fair share of challenges. This thesis, in multiple stages, addresses those issues. In the process, several key areas have been identified. In Chapter-4 **Machine Learning for Supervised Hierarchical Platform**, the ramifications of the speed of operation in decision making have been observed. In a hierarchical structure, if faster

decision-making tools are not deployed, system stability cannot be maintained. In a small-scale system, a centralized structure may bring value, however with added areas it will lose speed. The simplified patterns observed in a small-scale network also evolves with the addition of different types of distributed generators and loads. Therefore, the concept developed in this chapter is generic regarding hierarchy but specific in terms of decision making. Such specificity demands a distributed architecture.

In the Chapter-5, **Simplified Feature Selection for Dynamic Data**, the proposition of an independent and distributed system operation is presented. The distribution of resources depends on local patterns. Understanding those and forming a decision coherency requires specific knowledge extraction. This chapter successfully delivered that by proposing a simplified feature extraction technique that not only understands local patterns but also the influences those from neighbouring systems.

Metrics defining system restoration accuracy is a concept needed to be addressed in case of a larger network. For example, if the decision-making accuracy is not attained to a level of near perfection, then decision-making itself can lead towards a system-wide instability. In Chapter-6 **Data Augmentation for a Segmented Grid** a concept called corrective voltage control based optimization technique has been considered, as a native structure to define decision-making accuracy, in a multi-machine system. The evolution of features with stochastic data such as fault location, wind power, etc. is observed and dealt with in this chapter. The stochastic parameters can increase dimensionality. And with increased dimensionality training an algorithm becomes challenging. However, with an increased number of case studies and appropriate data preparation, this challenge can be overcome.

In Chapter-7 **Selection of Appropriate Feature Space** an analysis and comparison among different features have been presented. An association is brought by solving complex multi-area restoration problems. The problems are again stochastic, while being heavily influenced by the

inter-machine dynamics. Regardless, the fault location for the proposed feature-based machine learning algorithms are capable of restoring grid-segments and then the grid itself. However, it is also observed that misclassification, even if it occurs in rare conjunction of stochastic parameters has severe ramifications. Distributed control fails under those scenarios.

A backup plan thus becomes mandatory and in Chapter-8 **Ensemble Method for Multi-objective Application** an additional focus is set on it. Mitigation of any misclassification error is important. Thus, the machine learning platform must be self-correcting. This has been achieved at a concept level in this chapter by using hierarchical machine learning algorithm. While making decisions, the challenges of multi-objectivity has also been addressed.

The work carried out in this thesis cultivates the idea of how machine learning platforms can be prepared or manipulated for dynamic data analysis, to assist the self-healing process in a system. This thesis can be used as a reference for further expansion in the field of machine-learning based self-healing functionalities in a stochastic network. The potential is immense. Therefore, it is hoped that subsequent researchers can take these principles to a pathway for developing real-life self-healing applications.

## 9.2 Future Scopes

This work could only scratch the surface of self-healing process in a standalone microgrid using machine learning algorithms. In the future, this work can be taken even further through various means. One of the approaches would be to apply deep learning and artificial intelligence. In this method, the sequential decision-making process can be presented as a Markov, chain model. The process then will train a deep learning algorithm as a reinforcement learner. The stochastic parameters can help develop a policy gradient-based system for training. The process can provide a classification probability of each restoration process in a segmented grid, by building a  $n$ -the order Markov model, while the deep learning platform will associate the policy gradient and find

the optimum decision in each stochastic scenario. This method can be assumed as an extension of Chapter-8.

The Chapter-7 **Selection of Appropriate Feature Space** can further be extended by reviewing feature extension method applicable for different power system operational issues. This work can bring comparison among different techniques. Modern machine learning methods and their use in feature extraction can also be brought out. For example, deep learning algorithms have a work-flow that allows them to find suitable features. Ramifications of such approaches can also be introduced.

Further analysis in the online decision-making process can also be carried out. This approach will be an extension of Chapter-6 **Data Augmentation for a Segmented Grid**. Rather than moving towards deeper learning methods, how a simplified algorithm can be efficiently distributed can also be addressed.

The Chapter-5 **Simplified Feature Selection for Dynamic Data** can be further extended by implementing an optimization method that can address both feature selection and distributed control. Feature selection is used in this work to complement distributed control. However, feature selection can become resource heavy for an online system. A balance needs to be sought out. Therefore, an optimization routine can be developed for selecting appropriate features for online operations. It is hoped that this work has developed enough ground, for the subsequent researchers to carry it on and fathom the aspects which have not yet been discussed.

# Bibliography

- [1] F.-J. Chang and Y.-T. Chang, “Adaptive neuro-fuzzy inference system for prediction of water level in reservoir,” *Advances in Water Resources*, vol. 29, no. 1, pp. 1–10, 2006.
- [2] T. Guo and J. V. Milanovic, “Identification of power system dynamic signature using hierarchical clustering,” in *PES General Meeting— Conference & Exposition, 2014 IEEE*, pp. 1–5, IEEE, 2014.
- [3] R. A. Aliev, R. R. Aliev, B. Guirimov, and K. Uyar, “Dynamic data mining technique for rules extraction in a process of battery charging,” *Applied Soft Computing*, vol. 8, no. 3, pp. 1252–1258, 2008.
- [4] R. Aliev, B. Fazlollahi, B. Guirimov, and R. Aliev, *Recurrent Fuzzy Neural Networks and Their Performance Analysis*. INTECH Open Access Publisher, 2008.
- [5] J. Awan and D. hyo Bae, “Application of adaptive neuro-fuzzy inference system for dam inflow prediction using long-range weather forecast,” in *Digital Information Management (ICDIM), 2013 Eighth International Conference on*, pp. 247–251, Sept 2013.
- [6] “RetScreen software version-4 <http://www.retscreen.net /ang/version4.php>.” Accessed: 2015-03-12.
- [7] A. Bidram and A. Davoudi, “Hierarchical structure of microgrids control system,” *Smart Grid, IEEE Transactions on*, vol. 3, no. 4, pp. 1963–1976, 2012.

- [8] M. A. Karim, J. Currie, and T. T. Lie, “A distributed machine learning approach for the secondary voltage control of an islanded micro-grid,” in *2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, pp. 611–616, Nov 2016.
- [9] J. He, Y. W. Li, and F. Blaabjerg, “An enhanced islanding microgrid reactive power, imbalance power, and harmonic power sharing scheme,” *IEEE Transactions on Power Electronics*, vol. 30, pp. 3389–3401, June 2015.
- [10] T. Guo and J. V. Milanović, “Online identification of power system dynamic signature using pmu measurements and data mining,” *IEEE Transactions on Power Systems*, vol. 31, pp. 1760–1768, May 2016.
- [11] P. C. Wang, “Packet classification with multiple decision trees,” in *2015 21st Asia-Pacific Conference on Communications (APCC)*, pp. 626–631, Oct 2015.
- [12] H. Liu, X. Chen, K. Yu, and Y. Hou, “The control and analysis of self-healing urban power grid,” *IEEE Transactions on Smart Grid*, vol. 3, pp. 1119–1129, Sept 2012.
- [13] R. Li, X. Zhang, Q. Zhao, and G. Liu, “Economic optimization of self-healing control of power grid based on multi-agent system,” in *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pp. 1334–1337, Oct 2016.
- [14] M. Amin, “A smart self-healing grid: in pursuit of a more reliable and resilient system [in my view],” *IEEE Power and Energy Magazine*, vol. 12, no. 1, pp. 112–110, 2014.
- [15] H. Wang and Y. Liu, “Hierarchical case-based decision support system for power system restoration,” in *IEEE Power Engineering Society General Meeting, 2004.*, pp. 1115–1119 Vol.1, June 2004.

- [16] A. Golshani, W. Sun, Q. Zhou, Q. P. Zheng, and J. Tong, "Two-stage adaptive restoration decision support system for a self-healing power grid," *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1–1, 2017.
- [17] E. Kim, J. Lee, L. He, Y. Lee, and K. G. Shin, "Offline guarantee and online management of power demand and supply in cyber-physical systems," in *2016 IEEE Real-Time Systems Symposium (RTSS)*, pp. 89–98, Nov 2016.
- [18] J. Ma, X. Wang, and X. Lan, "Small-signal stability analysis of microgrid based on perturbation theory," in *2012 Asia-Pacific Power and Energy Engineering Conference*, pp. 1–4, March 2012.
- [19] J. Heaton, "An empirical analysis of feature engineering for predictive modeling," in *South-eastCon 2016*, pp. 1–6, March 2016.
- [20] M. Boudour and A. Hellal, "Large scale power system dynamic security assessment using the growing hierarchical self-organizing feature maps," in *Industrial Technology, 2004. IEEE ICIT'04. 2004 IEEE International Conference on*, vol. 1, pp. 370–375, IEEE, 2004.
- [21] M. Mohammadi and G. B. Gharehpetian, "Application of core vector machines for on-line voltage security assessment using a decisiontree-based feature selection algorithm," *IET Generation, Transmission Distribution*, vol. 3, pp. 701–712, August 2009.
- [22] J. Zhang, X. Zheng, Z. Wang, L. Guan, and C. Y. Chung, "Power system sensitivity identification x2014;inherent system properties and data quality," *IEEE Transactions on Power Systems*, vol. 32, pp. 2756–2766, July 2017.
- [23] A. Rabiee, A. Soroudi, B. Mohammadi-ivatloo, and M. Parniani, "Corrective voltage control scheme considering demand response and stochastic wind power," *IEEE Transactions on Power Systems*, vol. 29, pp. 2965–2973, Nov 2014.

- [24] T. J. Overbye and J. D. Weber, "Visualization of power system data," in *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, pp. 7–pp, IEEE, 2000.
- [25] D. Song, B. Wen, X. Yang, Y. Gu, S. Wei, and S. Ma, "A new method for processing and application of wide area measurement big data in a power system," in *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2015 International Conference on*, pp. 137–140, IEEE, 2015.
- [26] M. Al Karim, M. Chenine, K. Zhu, and L. Nordstrom, "Synchrophasor-based data mining for power system fault analysis," in *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pp. 1–8, IEEE, 2012.
- [27] N. Yu, S. Shah, R. Johnson, R. Sherick, and M. Hong, "Big data analytics in power distribution systems,"
- [28] J. Zhan, J. Huang, L. Niu, X. Peng, D. Deng, and S. Cheng, "Study of the key technologies of electric power big data and its application prospects in smart grid," in *Power and Energy Engineering Conference (APPEEC), 2014 IEEE PES Asia-Pacific*, pp. 1–4, IEEE, 2014.
- [29] M. Ghosal and V. Rao, "Fusion of pmu and scada data for dynamic state estimation of power system," in *North American Power Symposium (NAPS), 2015*, pp. 1–6, Oct 2015.
- [30] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, *et al.*, "Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions," *Power Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 1387–1401, 2004.
- [31] S. E. McHann, "Grid analytics: How much data do you really need?," in *Rural Electric Power Conference (REPC), 2013 IEEE*, pp. C3–1, IEEE, 2013.

- [32] M. Kezunovic, L. Xie, and S. Grijalva, "The role of big data in improving power system operation and protection," in *Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium*, pp. 1–9, IEEE, 2013.
- [33] K. Tomsovic and D. Faleo, "Advanced power system controls using intelligent systems," in *Power Engineering Society Summer Meeting, 2000. IEEE*, vol. 1, pp. 336–339, IEEE, 2000.
- [34] M. Kazerooni, H. Zhu, and T. J. Overbye, "Literature review on the applications of data mining in power systems," in *Power and Energy Conference at Illinois (PECI), 2014*, pp. 1–8, IEEE, 2014.
- [35] M. Ghodsi, "A brief review of recent data mining applications in the energy industry," *International Journal of Energy and Statistics*, vol. 2, no. 01, pp. 49–57, 2014.
- [36] H. Mori, "State-of-the-art overview on data mining in power systems," in *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, pp. 33–34, Oct 2006.
- [37] T. J. Overbye, E. M. Rantanen, and S. Judd, "Electric power control center visualization using geographic data views," in *Bulk Power System Dynamics and Control-VII. Revitalizing Operational Reliability, 2007 iREP Symposium*, pp. 1–8, IEEE, 2007.
- [38] T. J. Overbye, R. P. Klump, and J. D. Weber, "A virtual environment for interactive visualization of power system economic and security information," in *Power Engineering Society Summer Meeting, 1999. IEEE*, vol. 2, pp. 846–851, IEEE, 1999.
- [39] B. Zhang, W. Wu, H. Wang, and C. Zhao, "An efficient security assessment system based on pc cluster for power system daily operation planning validation," in *Power and Energy Society General Meeting, 2010 IEEE*, pp. 1–8, IEEE, 2010.

- [40] G. Lin, W. Tongwen, and Z. Yao, "New method for power system dynamic stability analysis based on a novel unsupervised clustering algorithm," in *Power and Energy Engineering Conference, 2009. APPEEC 2009. Asia-Pacific*, pp. 1–4, IEEE, 2009.
- [41] A. Chakraborty, "Wide-area damping control of power systems using dynamic clustering and tcsc-based redesigns," *Smart Grid, IEEE Transactions on*, vol. 3, no. 3, pp. 1503–1514, 2012.
- [42] S. Dutta and T. J. Overbye, "Optimal wind farm collector system topology design considering total trenching length," *Sustainable Energy, IEEE Transactions on*, vol. 3, no. 3, pp. 339–348, 2012.
- [43] K. M. Rogers and T. J. Overbye, "Clustering of power system data and its use in load pocket identification," in *System Sciences (HICSS), 2011 44th Hawaii International Conference on*, pp. 1–10, IEEE, 2011.
- [44] J.-H. Shin, B.-J. Yi, Y.-I. Kim, H.-G. Lee, and K. H. Ryu, "Spatiotemporal load-analysis model for electric power distribution facilities using consumer meter-reading data," *Power Delivery, IEEE Transactions on*, vol. 26, no. 2, pp. 736–743, 2011.
- [45] E. Karapidakis and N. Hatziargyriou, "Online preventive dynamic security of isolated power systems using decision trees," *Power Systems, IEEE Transactions on*, vol. 17, no. 2, pp. 297–304, 2002.
- [46] E. Leonidaki, D. Georgiadis, and N. Hatziargyriou, "Decision trees for determination of optimal location and rate of series compensation to increase power system loading margin," *Power Systems, IEEE Transactions on*, vol. 21, no. 3, pp. 1303–1310, 2006.

- [47] P. G. Axelberg, I. Y.-H. Gu, and M. H. Bollen, "Support vector machine for classification of voltage disturbances," *Power Delivery, IEEE Transactions on*, vol. 22, no. 3, pp. 1297–1303, 2007.
- [48] L. Fugon, J. Juban, and G. Kariniotakis, "Data mining for wind power forecasting," in *European Wind Energy Conference & Exhibition EWEC 2008*, pp. 6–pages, EWEC, 2008.
- [49] P. D. Diamantoulakis, V. M. Kapinas, and G. K. Karagiannidis, "Big data analytics for dynamic energy management in smart grids," *Big Data Research*, 2015.
- [50] R. Mallik, N. Sarda, H. Kargupta, and S. Bandyopadhyay, "Distributed data mining for sustainable smart grids," *Proc. of ACM SustKDD*, vol. 11, pp. 1–6, 2011.
- [51] S. Chan, K. Tsui, H. Wu, Y. Hou, Y.-C. Wu, and F. Wu, "Load/price forecasting and managing demand response for smart grids: Methodologies and challenges," *Signal Processing Magazine, IEEE*, vol. 29, pp. 68–85, Sept 2012.
- [52] F. Luo, Z. Dong, G. Chen, Y. Xu, K. Meng, Y. Chen, and K. Wong, "Advanced pattern discovery-based fuzzy classification method for power system dynamic security assessment," *Industrial Informatics, IEEE Transactions on*, vol. 11, no. 2, pp. 416–426, 2015.
- [53] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *Smart Grid, IEEE Transactions on*, 2015.
- [54] M. Aghamohammadi, F. Mahdavi-zadeh, and R. Bagheri, "Power system dynamic security classification using kohonen neural networks," in *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pp. 1–7, IEEE, 2009.
- [55] I. Genc, R. Diao, V. Vittal, S. Kolluri, and S. Mandal, "Decision tree-based preventive and corrective control applications for dynamic security enhancement in power systems," *Power Systems, IEEE Transactions on*, vol. 25, no. 3, pp. 1611–1619, 2010.

- [56] T. Assis, A. Nohara, and T. Valentini, "Power system dynamic security assessment through a neuro-fuzzy scheme," in *Intelligent System Applications to Power Systems, 2009. ISAP'09. 15th International Conference on*, pp. 1–6, IEEE, 2009.
- [57] G. Dudek, "Pattern-based local linear regression models for short-term load forecasting," *Electric Power Systems Research*, vol. 130, pp. 139–147, 2016.
- [58] C.-L. Niu, X.-N. Yu, J.-Q. Li, and W. Sun, "The application of operation optimization decision support system based on data mining in power plant," in *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, vol. 3, pp. 1830–1834, IEEE, 2005.
- [59] A. Ukil and R. Zivanovic, "Automated analysis of power systems disturbance records: Smart grid big data perspective," in *Innovative Smart Grid Technologies-Asia (ISGT Asia), 2014 IEEE*, pp. 126–131, IEEE, 2014.
- [60] S. Liu, *Dynamic-data driven real-time identification for electric power systems*. PhD thesis, Citeseer, 2009.
- [61] J. Meng, D. Sun, and Z. Li, "Applications of data mining time series to power systems disturbance analysis," in *Advanced Data Mining and Applications*, pp. 749–760, Springer, 2006.
- [62] R. Billinton, H. Chen, and R. Ghajar, "Time-series models for reliability evaluation of power systems including wind energy," *Microelectronics Reliability*, vol. 36, no. 9, pp. 1253–1261, 1996.
- [63] F. Martinez Alvarez, A. Troncoso, J. C. Riquelme, and J. S. Aguilar Ruiz, "Energy time series forecasting based on pattern sequence similarity," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 23, no. 8, pp. 1230–1243, 2011.

- [64] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espínola, “Forecasting next-day electricity prices by time series models,” *Power Systems, IEEE Transactions on*, vol. 17, no. 2, pp. 342–348, 2002.
- [65] C. Haomin, L. Peng, G. Xiaobin, X. Aidong, C. Bo, X. Wei, and Z. Liqiang, “Fault prediction for power system based on multidimensional time series correlation analysis,” in *Electricity Distribution (CICED), 2014 China International Conference on*, pp. 1294–1299, IEEE, 2014.
- [66] I. Mado, A. Soeprijanto, *et al.*, “Design of robust-fuzzy controller for smib based on power-load cluster model with time series analysis,” in *Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS), 2014*, pp. 8–15, IEEE, 2014.
- [67] Y. Wang, G. Cao, S. Mao, and R. Nelms, “Analysis of solar generation and weather data in smart grid with simultaneous inference of nonlinear time series,” in *Computer Communications Workshops (INFOCOM WKSHPS), 2015 IEEE Conference on*, pp. 600–605, IEEE, 2015.
- [68] H. A. Alsafih and R. Dunn, “Determination of coherent clusters in a multi-machine power system based on wide-area signal measurements,” in *Power and Energy Society General Meeting, 2010 IEEE*, pp. 1–8, IEEE, 2010.
- [69] S. Brin and L. Page, “Dynamic data mining: Exploring large rule spaces by sampling,” *Stanford InfoLab*, 1999.
- [70] J. Zhang, T. Li, D. Ruan, and D. Liu, “Neighborhood rough sets for dynamic data mining,” *International Journal of Intelligent Systems*, vol. 27, no. 4, pp. 317–342, 2012.
- [71] Y. Shi, *Dynamic data mining on multi-dimensional data*. PhD thesis, State University of New York at Buffalo, 2005.

- [72] A. Oonsivilai and M. El-Hawary, "Power system dynamic load modeling using adaptive-network-based fuzzy inference system," in *Electrical and Computer Engineering, 1999 IEEE Canadian Conference on*, vol. 3, pp. 1217–1222, IEEE, 1999.
- [73] P. Esling and C. Agon, "Time-series data mining," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, p. 12, 2012.
- [74] T.-c. Fu, "A review on time series data mining," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 164–181, 2011.
- [75] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.
- [76] Z. Liu, W. Gao, Y.-H. Wan, and E. Muljadi, "Wind power plant prediction by using neural networks," in *Energy Conversion Congress and Exposition (ECCE), 2012 IEEE*, pp. 3154–3160, Sept 2012.
- [77] M. G. Villalva, J. R. Gazoli, *et al.*, "Comprehensive approach to modeling and simulation of photovoltaic arrays," *IEEE Transactions on Power Electronics*, vol. 24, no. 5, pp. 1198–1208, 2009.
- [78] "Mathworks documentation." <http://www.mathworks.com/help/>. Accessed: 2015-03-09.
- [79] A. Mellit, A. HadjArab, N. Khorissi, and H. Salhi, "An anfis-based forecasting for solar radiation data from sunshine duration and ambient temperature," in *Power Engineering Society General Meeting, 2007. IEEE*, pp. 1–6, IEEE, 2007.
- [80] C. Chen, S. Duan, T. Cai, and B. Liu, "Online 24-h solar power forecasting based on weather type classification using artificial neural network," *Solar Energy*, vol. 85, no. 11, pp. 2856–2870, 2011.

- [81] H. Hippert, C. Pedreira, and R. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, pp. 44–55, Feb 2001.
- [82] A. Bakirtzis, V. Petridis, S. Kiartzis, M. Alexiadis, and A. Maissis, "A neural network short term load forecasting model for the Greek power system," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 858–863, 1996.
- [83] R. Karki, P. Hu, and R. Billinton, "A simplified wind power generation model for reliability evaluation," *IEEE Transactions on Energy Conversion*, vol. 21, pp. 533–540, June 2006.
- [84] X. Kong and A. Khambadkone, "Modeling of a Pem Fuel-Cell Stack for Dynamic and Steady-State Operation Using ANN-Based Submodels," *IEEE Transactions on Industrial Electronics*, vol. 56, pp. 4903–4914, Dec 2009.
- [85] G. A. Pagani and M. Aiello, "Towards decentralization: A topological investigation of the medium and low voltage grids," *Smart Grid, IEEE Transactions on*, vol. 2, no. 3, pp. 538–547, 2011.
- [86] J. Yuan, J. Ma, L. Zhao, and H. Liu, "A real-time balancing market clearing model for grid-connected micro-grid including energy storage system," in *Smart Electric Distribution Systems and Technologies (EDST), 2015 International Symposium on*, pp. 32–36, Sept 2015.
- [87] J. Lopes, C. Moreira, and A. Madureira, "Defining control strategies for microgrids islanded operation," *Power Systems, IEEE Transactions on*, vol. 21, no. 2, pp. 916–924, 2006.
- [88] T. R. Mtshali, G. Coppez, S. Chowdhury, and S. P. Chowdhury, "Simulation and modelling of PV-wind-battery hybrid power system," in *2011 IEEE Power and Energy Society General Meeting*, pp. 1–7, July 2011.
- [89] N. Argaw, R. Foster, and A. Ellis, "Renewable energy for water pumping applications in rural villages," *New Mexico State, United States*, 2003.

- [90] A. Elmitwally and M. Rashed, "Flexible operation strategy for an isolated pv-diesel micro-grid without energy storage," *IEEE Transactions on Energy Conversion*, vol. 26, pp. 235–244, March 2011.
- [91] S. Singh, A. K. Singh, and S. Chanana, "Operation and control of a hybrid photovoltaic-diesel-fuel cell system connected to micro-grid," in *Power India Conference, 2012 IEEE Fifth*, pp. 1–6, Dec 2012.
- [92] Y. Y. Hong, M. C. Hsiao, Y. R. Chang, Y. D. Lee, and H. C. Huang, "Multiscenario under-frequency load shedding in a microgrid consisting of intermittent renewables," *IEEE Transactions on Power Delivery*, vol. 28, pp. 1610–1617, July 2013.
- [93] M. Kohansal, G. B. Gharehpetian, M. Rahmatian, M. Abedi, and M. J. Sanjari, "The effect of load condition on stability in isolated microgrids," in *2012 Second Iranian Conference on Renewable Energy and Distributed Generation*, pp. 28–32, March 2012.
- [94] S. Xing, "Microgrid emergency control based on the stratified controllable load shedding optimization," in *Sustainable Power Generation and Supply (SUPERGEN 2012), International Conference on*, pp. 1–5, Sept 2012.
- [95] A. Y. Saber and G. K. Venayagamoorthy, "Smart micro-grid optimization with controllable loads using particle swarm optimization," in *2013 IEEE Power Energy Society General Meeting*, pp. 1–5, July 2013.
- [96] J. Xu, S. Tan, and S. K. Panda, "Optimization of economic load dispatch for a microgrid using evolutionary computation," in *IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society*, pp. 3192–3197, Nov 2011.
- [97] F. Shahnia, R. P. S. Chandrasena, S. Rajakaruna, and A. Ghosh, "Primary control level of parallel distributed energy resources converters in system of multiple interconnected au-

- tonomous microgrids within self-healing networks,” *IET Generation, Transmission Distribution*, vol. 8, pp. 203–222, February 2014.
- [98] M. Babazadeh and H. Karimi, “Robust decentralized control for islanded operation of a microgrid,” in *2011 IEEE Power and Energy Society General Meeting*, pp. 1–8, July 2011.
- [99] B. S. Hartono, Y. Budiyanto, and R. Setiabudy, “Review of microgrid technology,” in *QiR (Quality in Research), 2013 International Conference on*, pp. 127–132, June 2013.
- [100] M. Amin, “Toward self-healing energy infrastructure systems,” *IEEE Computer Applications in Power*, vol. 14, pp. 20–28, Jan 2001.
- [101] P. C. Sekhar, S. Mishra, and R. Sharma, “Data analytics based neuro-fuzzy controller for diesel-photovoltaic hybrid ac microgrid,” *IET Generation, Transmission Distribution*, vol. 9, no. 2, pp. 193–207, 2015.
- [102] M. Savaghebi, A. Jalilian, J. C. Vasquez, and J. M. Guerrero, “Secondary control scheme for voltage unbalance compensation in an islanded droop-controlled microgrid,” *Smart Grid, IEEE Transactions on*, vol. 3, no. 2, pp. 797–807, 2012.
- [103] S. Brahma, R. Kavasseri, H. Cao, N. R. Chaudhuri, T. Alexopoulos, and Y. Cui, “Real-time identification of dynamic events in power systems using pmu data, and potential applications ;models, promises, and challenges,” *IEEE Transactions on Power Delivery*, vol. 32, pp. 294–301, Feb 2017.
- [104] I. Kouveliotis-Lysikatos and N. Hatziargyriou, “Distributed economic dispatch considering transmission losses,” in *2017 IEEE Manchester PowerTech*, pp. 1–6, June 2017.
- [105] T. Nagata and H. Sasaki, “A multi-agent approach to power system restoration,” *IEEE Transactions on Power Systems*, vol. 17, pp. 457–462, May 2002.

- [106] Z. Wang and J. Wang, "Self-healing resilient distribution systems based on sectionalization into microgrids," *IEEE Transactions on Power Systems*, vol. 30, pp. 3139–3149, Nov 2015.
- [107] C. Zhang, Y. Li, Z. Yu, and F. Tian, "Feature selection of power system transient stability assessment based on random forest and recursive feature elimination," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 1264–1268, Oct 2016.
- [108] K. Niazi, C. Arora, and S. Surana, "Power system security evaluation using ann: feature selection using divergence," *Electric Power Systems Research*, vol. 69, no. 2, pp. 161–167, 2004.
- [109] J. Geeganage, U. D. Annakkage, T. Weekes, and B. A. Archer, "Application of energy-based power system features for dynamic security assessment," *IEEE Transactions on Power Systems*, vol. 30, pp. 1957–1965, July 2015.
- [110] O. Abedinia, N. Amjady, and H. Zareipour, "A new feature selection technique for load and price forecast of electrical power systems," *IEEE Transactions on Power Systems*, vol. 32, pp. 62–74, Jan 2017.
- [111] F. M. Tehrani, G. Shahgholian, and H. Pourghassem, "Dynamic study and stability analyze of damping cohefision and reactance in tcsc controller connected on optimization smib system," in *2011 IEEE 3rd International Conference on Communication Software and Networks*, pp. 270–274, May 2011.
- [112] M. D. Hausmann and H. Ziekow, "The potential of household specific feature selection for analysing smart home time-series data," in *Smart SysTech 2016; European Conference on Smart Objects, Systems and Technologies*, pp. 1–6, July 2016.

- [113] P. Ferraro, E. Crisostomi, M. Raugi, and F. Milano, "Analysis of the impact of microgrid penetration on power system dynamics," *IEEE Transactions on Power Systems*, vol. PP, no. 99, pp. 1–1, 2016.
- [114] D. J. Hill and I. M. Y. Mareels, "Stability theory for differential/algebraic systems with application to power systems," *IEEE Transactions on Circuits and Systems*, vol. 37, pp. 1416–1423, Nov 1990.
- [115] P. Arunagirinathan, H. A. Abdelsalam, and G. K. Venayagamoorthy, "Remote power system stabilizer tuning using synchrophasor data," in *2014 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*, pp. 1–7, Dec 2014.
- [116] R. Grondin, I. Kamwa, G. Trudel, L. Gerin-Lajoie, and J. Taborda, "Modeling and closed-loop validation of a new pss concept, the multi-band pss," in *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No.03CH37491)*, vol. 3, p. 1809 Vol. 3, July 2003.
- [117] E. Vittal, M. O'Malley, and A. Keane, "Rotor angle stability with high penetrations of wind generation," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 353–362, 2012.
- [118] Z. Wang, E. B. Makram, G. K. Venayagamoorthy, and C. Ji, "Dynamic estimation of rotor angle deviation of a generator in multi-machine power systems," *Electric Power Systems Research*, vol. 97, pp. 1–9, 2013.
- [119] T. Nagata, Y. Tao, H. Sasaki, and H. Fujita, "A multiagent approach to distribution system restoration," in *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No.03CH37491)*, vol. 2, p. 660 Vol. 2, July 2003.
- [120] L. Sun, C. Zhang, Z. Lin, F. Wen, Y. Xue, M. A. Salam, and S. P. Ang, "Network partitioning strategy for parallel power system restoration," *IET Generation, Transmission Distribution*, vol. 10, no. 8, pp. 1883–1892, 2016.

- [121] D. R. Medina, E. Rappold, O. Sanchez, X. Luo, S. R. R. Rodriguez, D. Wu, and J. N. Jiang, “Fast assessment of frequency response of cold load pickup in power system restoration,” *IEEE Transactions on Power Systems*, vol. 31, pp. 3249–3256, July 2016.
- [122] N. Mithulananthan, C. A. Canizares, J. Reeve, and G. J. Rogers, “Comparison of pss, svc, and statcom controllers for damping power system oscillations,” *IEEE transactions on power systems*, vol. 18, no. 2, pp. 786–792, 2003.
- [123] M. H. Bollen and I. Gu, *Signal processing of power quality disturbances*, vol. 30. John Wiley & Sons, 2006.
- [124] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.
- [125] P. H. R, N. K. K, C. T. L, N. K, and B. R. M, “Feature engineering on forest cover type data with ensemble of decision trees,” in *2015 IEEE International Advance Computing Conference (IACC)*, pp. 1093–1098, June 2015.
- [126] R. Ueda, T. Arai, K. Sakamoto, Y. Jitsukawa, K. Umeda, H. Osumi, T. Kikuchi, and M. Komura, “Real-time decision making with state-value function under uncertainty of state estimation-evaluation with local maxima and discontinuity,” in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pp. 3464–3469, April 2005.
- [127] I. S. Honkala and T. K. Laihonon, “The probability of undetected error can have several local maxima,” *IEEE Transactions on Information Theory*, vol. 45, pp. 2537–2539, Nov 1999.
- [128] E. W. Abel, N. S. Arkidis, and A. Forster, “Inter-scale local maxima-a new technique for wavelet analysis of emg signals,” in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286)*, vol. 3, pp. 1471–1473 vol.3, Oct 1998.

- [129] J. Griffié, L. Boelen, G. Burn, A. P. Cope, and D. M. Owen, “Topographic prominence as a method for cluster identification in single-molecule localisation data,” *Journal of biophotonics*, vol. 8, no. 11-12, pp. 925–934, 2015.
- [130] E. Barocio, B. C. Pal, N. F. Thornhill, and A. R. Messina, “A dynamic mode decomposition framework for global power system oscillation analysis,” *IEEE Transactions on Power Systems*, vol. 30, pp. 2902–2912, Nov 2015.
- [131] Y. Y. Hong and L. H. Lee, “Analysis of equivalent 10 hz voltage flicker in power systems,” *IEE Proceedings - Generation, Transmission and Distribution*, vol. 146, pp. 447–452, Sep 1999.
- [132] M. V. E. Xavier, W. d. C. Boaventura, and M. D. Noronha, “High performance power quality monitoring system,” in *2016 17th International Conference on Harmonics and Quality of Power (ICHQP)*, pp. 974–979, Oct 2016.
- [133] I. Kamwa, S. Samantaray, and G. Joos, “Catastrophe predictors from ensemble decision-tree learning of wide-area severity indices,” *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 144–158, 2010.
- [134] Z. Hu, M. Xiong, H. Shang, and A. Deng, “Anti-interference measurement methods of the coupled transmission-line capacitance parameters based on the harmonic components,” *IEEE Transactions on Power Delivery*, vol. 31, pp. 2464–2472, Dec 2016.
- [135] G. L. G. de Pinho, G. S. Wojichowski, and C. D. P. Crovato, “Tool for power quality assessment of motor fed by variable speed electric drivers,” in *2016 17th International Conference on Harmonics and Quality of Power (ICHQP)*, pp. 109–114, Oct 2016.
- [136] P. Bühlmann, “Bagging, boosting and ensemble methods,” in *Handbook of Computational Statistics*, pp. 985–1022, Springer, 2012.

- [137] I. Guyon and A. Elisseeff, "An introduction to feature extraction," *Feature extraction*, pp. 1–25, 2006.
- [138] P.-N. Tan, M. Steinbach, and V. Kumar, "Classification: basic concepts, decision trees, and model evaluation," *Introduction to data mining*, vol. 1, pp. 145–205, 2006.
- [139] S. Zubić, A. Wahlroos, J. Altonen, P. Balcerak, and P. Dawidowski, "Managing post-fault oscillation phenomenon in compensated mv-networks," in *13th International Conference on Development in Power System Protection 2016 (DPSP)*, pp. 1–6, March 2016.
- [140] L. Goel, Q. Wu, and P. Wang, "Reliability enhancement of a deregulated power system considering demand response," in *2006 IEEE Power Engineering Society General Meeting*, pp. 6 pp.–, 2006.
- [141] V. Ajarapu and A. S. Meliopoulos, "Preventing voltage collapse with protection systems that incorporate optimal reactive power control," *PSERC Publication*, pp. 08–20, 2008.
- [142] A. Rabiee, A. Soroudi, and A. Keane, "Risk-averse preventive voltage control of ac/dc power systems including wind power generation," *IEEE Transactions on Sustainable Energy*, vol. 6, pp. 1494–1505, Oct 2015.
- [143] T. Athay, R. Podmore, and S. Virmani, "A practical method for the direct analysis of transient stability," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, pp. 573–584, March 1979.
- [144] S. Lian, S. Morii, T. Ishii, and S. Kawamoto, "Voltage stability and sensitivity analysis considering dynamic load for smart grid," in *2010 Innovative Smart Grid Technologies (ISGT)*, pp. 1–6, Jan 2010.
- [145] D. Whitley, "A genetic algorithm tutorial," *Statistics and computing*, vol. 4, no. 2, pp. 65–85, 1994.

- [146] A. Zidan, M. Khairalla, A. M. Abdrabou, T. Khalifa, K. Shaban, A. Abdrabou, R. E. Shatshat, and A. M. Gaouda, "Fault detection, isolation, and service restoration in distribution systems: State-of-the-art and future trends," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–16, 2017.
- [147] Y. Zhou, R. Arghandeh, I. Konstantakopoulos, S. Abdullah, A. von Meier, and C. J. Spanos, "Abnormal event detection with high resolution micro-pmu data," in *2016 Power Systems Computation Conference (PSCC)*, pp. 1–7, June 2016.
- [148] S. Jin, Z. Huang, R. Diao, D. Wu, and Y. Chen, "Comparative implementation of high performance computing for power system dynamic simulations," *IEEE Transactions on Smart Grid*, vol. 8, pp. 1387–1395, May 2017.
- [149] S. K. Khaitan, "A survey of high-performance computing approaches in power systems," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pp. 1–5, July 2016.
- [150] A. dos Santos and M. T. C. de Barros, "Sensitivity of voltage sags to network failure rate improvement," in *2016 Power Systems Computation Conference (PSCC)*, pp. 1–7, June 2016.
- [151] G. E. Constante-Flores and M. Illindala, "Data-driven probabilistic power flow analysis for a distribution system with renewable energy sources using monte carlo simulation," in *2017 IEEE/IAS 53rd Industrial and Commercial Power Systems Technical Conference (I CPS)*, pp. 1–8, May 2017.
- [152] T. T. Nguyen and V. L. Nguyen, "Power system security restoration by secondary control," in *2007 IEEE Power Engineering Society General Meeting*, pp. 1–8, June 2007.
- [153] X. Zhan, T. Xiang, H. Chen, B. Zhou, and Z. Yang, "Vulnerability assessment and reconfiguration of microgrid through search vector artificial physics optimization algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 62, pp. 679–688, 2014.

- [154] Q. Ye, T. Ma, Y. Gu, TaoWang, D. Wang, and Y. Bai, "Research on dispatch scheduling model of micro-grid with distributed energy," in *Electricity Distribution (CICED), 2012 China International Conference on*, pp. 1–5, Sept 2012.
- [155] J. Saez-Gallego and J. M. Morales, "Short-term forecasting of price-responsive loads using inverse optimization," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2017.
- [156] L. Zhuang, H. Liu, J. Zhu, S. Wang, and Y. Song, "Comparison of forecasting methods for power system short-term load forecasting based on neural networks," in *2016 IEEE International Conference on Information and Automation (ICIA)*, pp. 114–119, Aug 2016.
- [157] L. de Andrade, M. Oleskovicz, A. Quaresma Santos, D. Vinicius Coury, and R. Souza Fernandes, "Very short-term load forecasting based on NARX recurrent neural networks," in *PES General Meeting — Conference Exposition, 2014 IEEE*, pp. 1–5, July 2014.
- [158] B. Zhao, X. Dong, and J. Bornemann, "Service restoration for a renewable-powered micro-grid in unscheduled island mode," *IEEE Transactions on Smart Grid*, vol. 6, pp. 1128–1136, May 2015.
- [159] H. F. Habib, T. Yossef, M. Cintuglu, and O. Mohammed, "A multi-agent based technique for fault location, isolation, and service restoration," *IEEE Transactions on Industry Applications*, vol. PP, no. 99, pp. 1–1, 2017.
- [160] Z. Wang and J. Wang, "Service restoration based on ami and networked mgs under extreme weather events," *IET Generation, Transmission Distribution*, vol. 11, no. 2, pp. 401–408, 2017.
- [161] C. Moreira, F. Resende, and J. P. Lopes, "Using low voltage microgrids for service restoration," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 395–403, 2007.

- [162] R. F. Sampaio, L. S. Melo, R. P. S. Leão, G. C. Barroso, and J. R. Bezerra, “Automatic restoration system for power distribution networks based on multi-agent systems,” *IET Generation, Transmission Distribution*, vol. 11, no. 2, pp. 475–484, 2017.
- [163] L. H. Fink, K.-L. Liou, and C.-C. Liu, “From generic restoration actions to specific restoration strategies,” *IEEE Transactions on Power Systems*, vol. 10, no. 2, pp. 745–752, 1995.
- [164] M. Negnevitsky, N. Tomin, V. Kurbatsky, D. Panasetsky, A. Zhukov, and C. Rehtanz, “A random forest-based approach for voltage security monitoring in a power system,” in *PowerTech, 2015 IEEE Eindhoven*, pp. 1–6, June 2015.
- [165] X. Yu and C. Singh, “A practical approach for integrated power system vulnerability analysis with protection failures,” *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1811–1820, 2004.
- [166] S. A. Kalogirou, “Applications of artificial neural-networks for energy systems,” *Applied Energy*, vol. 67, no. 1, pp. 17–35, 2000.
- [167] A. Mishra and L. Ramesh, “Application of neural networks in wind power (generation) prediction,” in *Sustainable Power Generation and Supply, 2009. SUPERGEN '09. International Conference on*, pp. 1–5, IEEE, April 2009.
- [168] A. J. F. T. Li and L. Zhao, “Neural network compensation control for output power optimization of wind energy conversion system based on data-driven control,” *Journal of Control Science and Engineering*, vol. 2012, p. 15, 2012.
- [169] T. Routh, A. Bin Yousuf, M. Hossain, M. Asasduzzaman, M. Hossain, U. Husnaeen, and M. Mubarak, “Artificial neural network based temperature prediction and its impact on solar cell” in *Informatics, Electronics Vision (ICIEV), 2012 International Conference on*, pp. 897–902, May 2012.

- [170] A. Parisio, E. Rikos, and L. Glielmo, “A model predictive control approach to microgrid operation optimization,” *Control Systems Technology, IEEE Transactions on*, vol. 22, no. 5, pp. 1813–1827, 2014.
- [171] [http://www.windenergy.org.nz/hau-nui-wind farm](http://www.windenergy.org.nz/hau-nui-wind-farm), “New zealand wind energy association, hau nui wind farm,” February 2017.
- [172] [http://www.mbie.govt.nz/info-services/sectors-industries/energy/energy-data-modelling/publications/energy-in-new zealand](http://www.mbie.govt.nz/info-services/sectors-industries/energy/energy-data-modelling/publications/energy-in-new-zealand), “Energy in new zealand 2016, the ministry of business, innovation and employment (mbie),” February 2017.
- [173] A. Sankar Krishnan and R. Billinton, “Sequential monte carlo simulation for composite power system reliability analysis with time varying loads,” *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1540–1545, 1995.
- [174] Z. Chu, J. Zhang, O. Kosut, and L. Sankar, “Evaluating power system vulnerability to false data injection attacks via scalable optimization,” in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 260–265, Nov 2016.
- [175] B. M. S. M. Ramadan, T. Logenthiran, R. T. Naayagi, and C. Su, “Hybridization of genetic algorithm and priority list to solve economic dispatch problems,” in *2016 IEEE Region 10 Conference (TENCON)*, pp. 1467–1470, Nov 2016.
- [176] C. Lin, W. Wu, X. Chen, and W. Zheng, “Decentralized dynamic economic dispatch for integrated transmission and active distribution networks using multi-parametric programming,” *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2017.