COMPARATIVE ANALYSIS OF TRADITIONAL MACHINE LEARNING METHODS AND SPIKING NEURAL NETWORKS FOR SPATIO-TEMPORAL DATA MINING

PARAG GANESH NAYAK

A thesis submitted to

Auckland University of Technology

In partial fulfilment of the requirement for the degree of

Master of Computer and Information Sciences (MCIS)

2016

Faculty of Design and Creative Technologies

School of Computer and Mathematical Sciences

Attestation of Authorship

"I hereby declare that this submission is my own work and to the best of my belief and knowledge, it contains no material previously written or published by another person (except acknowledgements or where explicitly defined) nor has the material been submitted for any award or degree or diploma to an institution or university of higher learning."

Parag 28/07/2016

Parag Ganesh Nayak

July, 2016

Contents

List of Figures:-

List of Tables:-

Acknowledgement

I would like to thank my primary supervisor, Professor Nikola Kasabov, and my secondary supervisor, Dr. Enmei Tu, for accepting me as a master’s student and for allowing me to become a part of this revered institute. Professor Kasabov’s comments, teaching and support of my research was very important and critical. He navigated me towards the right research path. Dr Enmei Tu’s comments also helped me gain new insight into my research.
I would also like to thank Joyce D’Mello, the manager of Knowledge Engineering and Discovery Research Institute (KEDRI) for her encouragement and support which helped me during my research study.
I would also like to thank my colleague Reggio Hartono. PhD student for helping me in my research.
I would also like to thank my fellow master’s and PhD students and other staff at KEDRI who supported me during my research study.
Last, but not least, I would like to thank my family for providing me with this opportunity and also for their support and encouragement.

Abstract

A new framework in this study, which uses spiking neural networks for learning spectro-temporal and spatio-temporal data, is the NeuCube. The NeuCube is able to learn and classify and predict data, both in online and offline modes.

NeuCube-based methodology is used, tested and implemented for the classification and regression of spatio-temporal data (seismic data set). The spatio-temporal data consists of both time and space. In this study the spatial data are latitude, longitude and depth. The temporal data is the magnitude of the earthquake. The modelling of the spatio-temporal data is used to predict new patterns from the complex spatio- and spectral temporal data and to make an accurate prediction of events such as the predicting the occurrence of earthquakes. Seismic data is in relation to the occurrence of earthquakes and is acquired by analyzing the surface of the earth through the deployment of various sensors. After deployment, actuation of the sensor's source, which receives the seismic signals, occurs in turn producing raw seismic data.

In this study, the author performs a comparative analysis of the spatio-temporal data with regards to the machine learning algorithms (WEKA), evolving connectionist systems (NeuCom) and spiking neural networks (NeuCube). The comparative analysis between machine learning algorithms and spiking neural networks is based on classification of the data, and the comparative analysis between evolving connectionist systems and spiking neural networks is based on regression/prediction. The seismic dataset used in this study is publicly available on the GeoNet website (www.geonet.org.nz).

Chapter 1:-Introduction

1. Motivation

In science today, the prediction of earthquakes, is still a distant goal, as numerous parameters need to be known in precedence or parallel to be able to make successful earthquake predictions.**(Kumar et al., 2013)** Some theoretical seismologists consider the prediction of earthquakes to be an impossible task. **(Lighthill, 1996)** For the past decade there have been numerous earthquakes around the world claiming huge numbers of casualties and causing damage to economies. As, an earthquake is a natural calamity, predicting it perfectly is impossible. But we can at least try to be near perfect in predicting an earthquake to save lives and to better understand the occurrence of earthquakes. The factors that determine the strength of an earthquake are its magnitude, depth, local geological conditions, secondary effects, depth and architecture. **(Cole, Elliott, Okubo, & Strobl, 2013)**

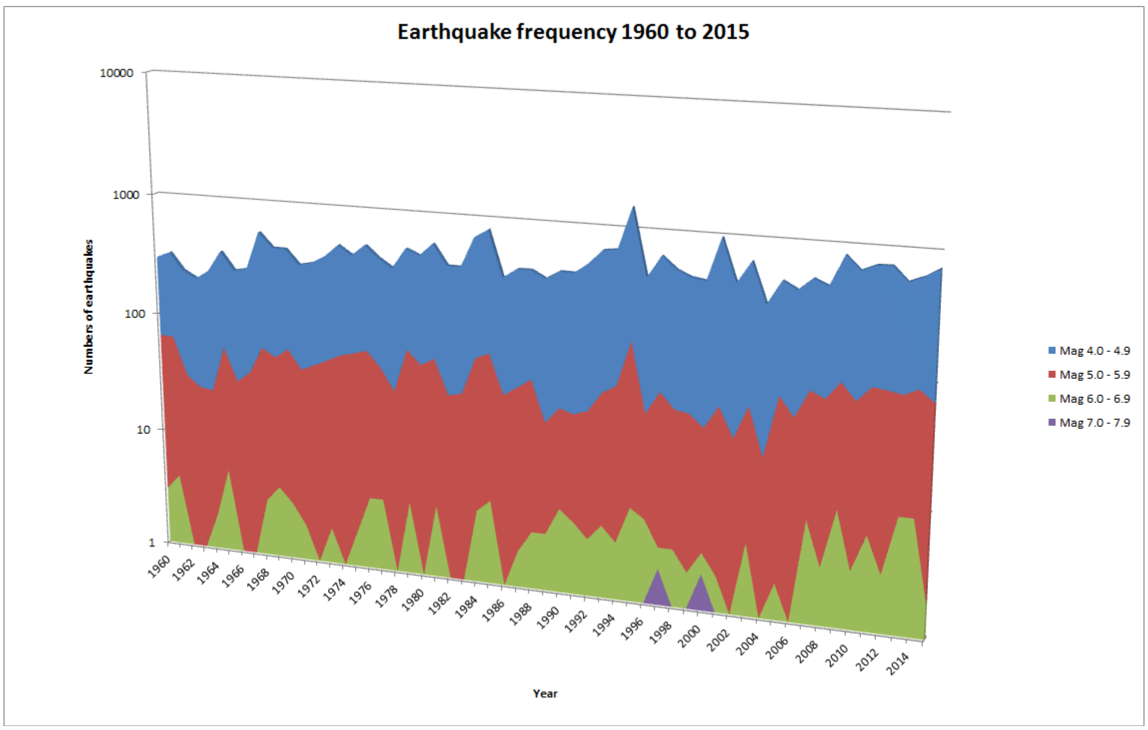

Figure 1:- ***Number of earthquakes in the New Zealand region, by magnitude range, from 1960-2015. (GeoNet, n.d.)***

As shown in the above figure (Figure 1), a lot of earthquakes occur in New Zealand which are not necessarily huge but which can cause a lot of damage to property as well as to humans. Some earthquakes that are low in magnitude can be dangerous due to the depth at which they occur. This curios fact led me to do a comparative analysis between traditional machine learning algorithms, evolving connectionist systems and spiking neural networks to understand which model could be helpful in the classification as well as the prediction of earthquakes.

2. Research Scope and Focus

The research scope of this study is to compare and analyze the classification and regression of seismic data between machine learning algorithms (WEKA), evolving connectionist systems (NeuCom) and spiking neural networks (NeuCube). The focus is mainly on which software/algorithm can be better used to classify or predict an earthquake.

3. Research Question

The Research Question for my study is:-
How do traditional machine learning methods compare with spiking neural networks in terms of:-
a) Accuracy of pattern classification

b) Both early and accurate prediction

c) Deep learning capacity

4. Overview of the Study

Literature review based on spatio-temporal data and earthquakes is presented. Literature review based on seismic data is also presented. Literature review based on machine learning algorithms and their implementation in WEKA is presented. Literature review based on Evolving Connectionist Systems and their implementation in NeuCom is presented. Literature review based on Spiking Neural Networks and their implementation in NeuCube is presented.

A methodology framework has been proposed for this comparative analysis. The seismic data has been collected from the GeoNet website (www.geonet.org.nz) The data that collected was extracted from the Canterbury area of New Zealand only. A 3D visualization was also implemented for the visualization of spikes that have been trained in NeuCube based on seismic data. The results and analysis are based on preliminary experiments done during the study.

5. Structure of the Thesis

The structure of the thesis is as follows:-

Chapter 2:- This chapter describes the spatio-temporal data. Detailed information on seismic data and earthquakes is also given.

Chapter 3:-This chapter presents information on traditional machine learning algorithms and evolving connectionist systems (ECOS). It also provides information about WEKA software and NeuCom software

Chapter 4:- An introduction to spiking neural networks is presented, and the information on the model used for spiking neural networks, i.e. Neucube, is given.

Chapter 5:- This chapter provides information on the methodology framework.

Chapter 6:-The experiment results along with the visualization and the analysis are described in this chapter.

Chapter 7:- This chapter is the conclusion of the study and suggestions for future work.

Chapter 2:- Introduction to Spatio-Temporal Data and a Case Study on Earthquakes Occurring in New Zealand

2.1 Spatio-Temporal Data:-

Spatio-Temporal Data is considered to be robust temporal content that is mostly collected in areas such as Bio-Informatics (e.g. gene and protein expression), Engineering (e.g. speech and audio), Economics (e.g. financial time series and macroeconomics), Ecology (e.g. seismic data and species establishment), Environment (e.g. global warming), Medicine (e.g. risk of disease and patient's recovery time), Neuro-Informatics (e.g. fMRI, EEG) etc. Spatio-Temporal modelling can be used to predict new patterns from complex SSTD (Spatio- and Spectro-Temporal Data) to provide an accurate prediction of Spatio-Temporal events in machine learning systems. **(N. Kasabov, Dhoble, Nuntalid, & Indiveri, 2013)** Numerous problems in nature require spatio and/or spectro-temporal data (SSTD) which includes measuring spatial and / or spectral variables over time. SSTD is defined by a triplet (X, Y, and F), where Y is a set of output dependent variables, F is known as the association function between whole segments of data, and X is a set of independent variables measured over successive discrete time moments. The association function is sampled in a window time frame with the output variable belonging to

$$F: X(d_t) \rightarrow Y$$

Where X (t) = $(x_1(t), x_2(t) \dots x_n(t))$, t=1, 2…

It is significant that a computational model learns and captures whole spatial and spectro-temporal patterns from various data streams to predict accurate future events in relation to new input data. **(N. Kasabov, 2014**) Space and Time are viewed as essential aspects of the phenomenon of the real world. Spatio-temporal data (STD) consists of information in relation to both space and time. **(Liang, Krishnamurthi, Kasabov, & Feigin, 2014)** The spatial scale of the data available is mostly determined by a random grid that is usually larger than the true dimensionality of the system. The main task is to identify the individual semi-autonomous components of the system and to deduce their (potentially weighted and lagged) interconnections from the spatio-temporal data. **(Fountalis, Bracco, Dilkina, Dovrolis, & Keilholz, 2016)**

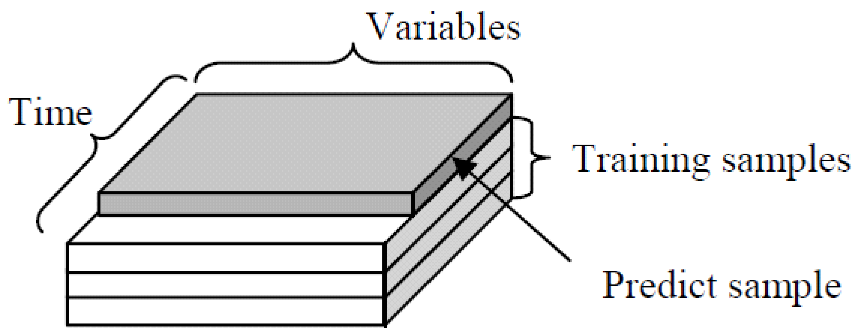

*Figure 2:- **A spatio-temporal data model used for early event prediction. (Tu et al., 2014)***

Early event prediction is important for solving vital ecological tasks defined by spatio-temporal data such as earthquake prediction. The standard task is the early and accurate prediction of an event that will occur in the future based on historical spatio-temporal data. The above figure (Figure 2) illustrates that the time length of the test data (samples used for prediction) and the training data (samples collected in the past) can be different. Predictive modelling of the spatio-temporal data (STD) is a difficult task because it is demanding to model both space and time components of data due to their interrelationship and interconnection. **(Tu et al., 2014)**

2.2 Earthquake:-

2.2.1 What is an earthquake?

An earthquake is a sudden motion or trembling in the crust caused by the abrupt release of accumulated stress along a fault, a break in the earth's crust. **(GNS Science, n.d.)**

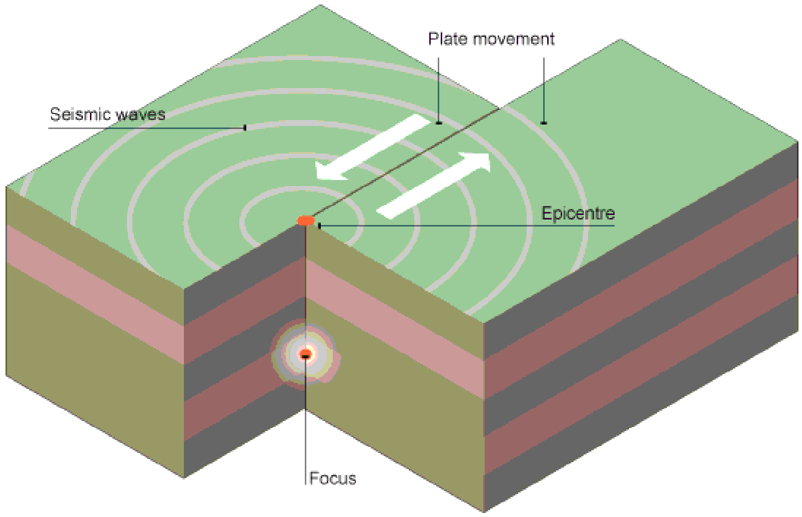

*Figure 3:- **An earthquake (BBC GCSE BiteSize, 2014)***

2.2.2 Types of Earthquakes:-

Deep Earthquakes:-

Underneath the North Island lies the Benioff Zone which delineates the active seismic part of the Pacific Plate that is subdued underneath the Australian Plate. The deep earthquakes that are located in the Fiordland region define the sharp easterly dipping zone that represents the subduction of the Australian Plate underneath the Pacific Plate. **(Anderson & Webb, 1994)**

Hikurangi Margin:-

The deep earthquakes that mark the Hikurangi Benioff Zone extend into the northern South Island and ends at a boundary line south of Westport. Deep earthquakes are most intense in the western region of the map as shown in the figure below (Figure 4) as there are very steep slab dips underneath this region's boundary. The standard location procedures had presented a systematic bias in the area since earthquakes with depths >40 km will be likely to have fixed depths (generally 33km) when they are >30km from the closest seismograph (Figure 5).

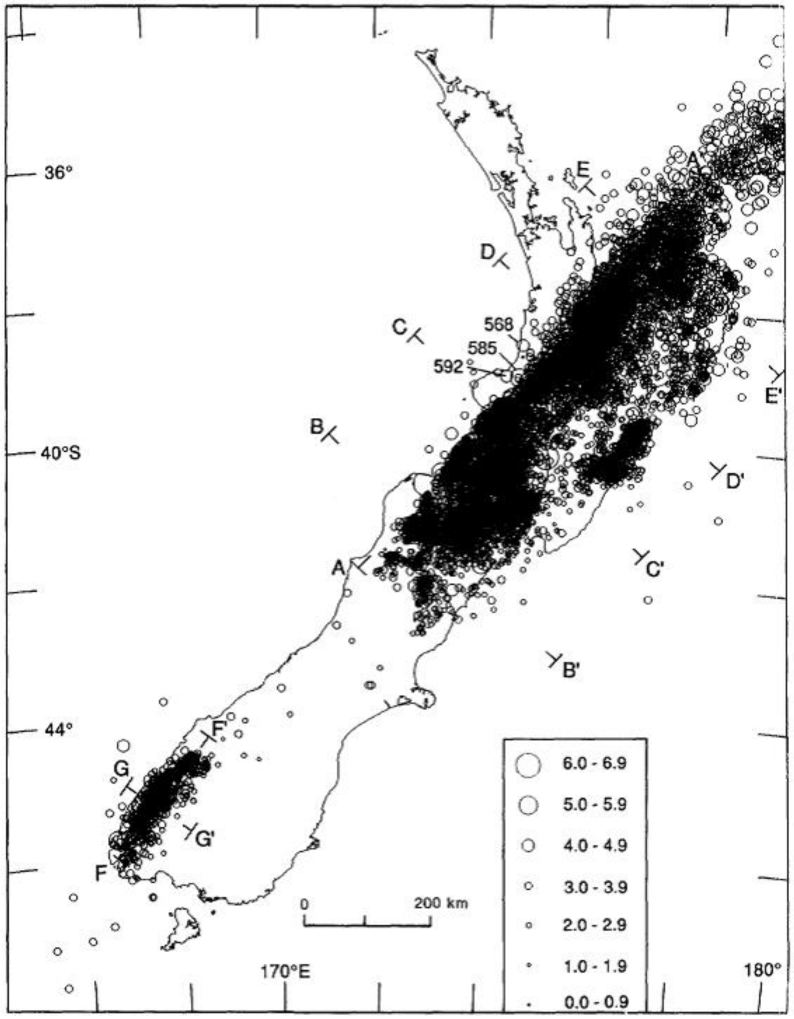

Figure 4:- ***Depth Distribution of Earthquakes in New Zealand (Anderson & Webb, 1994)***

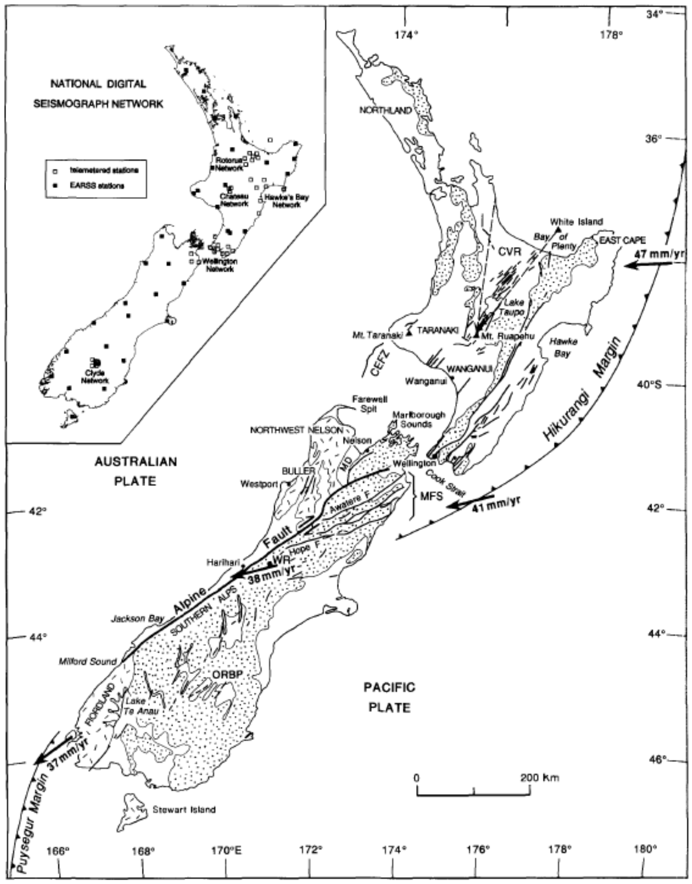

*Figure 5:- **Generalized tectonic map of New Zealand with active faults and major tectonic provinces. (Anderson & Webb, 1994)***

The depth distribution of earthquakes in the Hikurangi margin of New Zealand as shown in the above figure (Figure 4) suggests the persistence of a few features, such as clustering activity which can be related to the subducting slab's physical characteristics. The most visible feature of the deep seismicity area is the shallowing from the north-east to the south-west region of the bottom edge of the Benioff Zone. Supposing, that the length of the attached aseismic slab to the Benioff Zone is analogous everywhere, the shallowing could be interpreted as an indication of the decreasing total slab length, because of low convergence rates in Wellington (41mm/a) in relation to East Cape (47mm/a). The convergence rates were calculated from the Pacific-Australian Euler Pole taken from. **(DeMets, Gordon, Argus, & Stein, 1990)**

In the northern part of South Island, there is a sharp step located at the base of the Benioff zone; earthquakes in the Westport region range to depths of only about 100km, while the seismicity in the north-west of the Nelson region continues to depths of around 230km with isolated events that occur at depths of about 350km. Due to the nature and geometry of the presubduction of the Pacific Plate, there is a dissimilarity in the slab depths, suggesting time difference in the subduction initiation. Underneath the main Benioff Zone lie three groups of deep events. At the north-eastern end underneath White Island lies an isolated event at a depth of about 450km. There are three extremely deep (>550km) earthquakes located within 60 km of each other underneath the north Taranaki region with earthquake magnitudes ranging from 4.4-5.4. At the 670 km discontinuity, deep earthquakes occur in a slice of the detached and sinking slab. **(Anderson & Webb, 1994)**

Fiordland Region:-

Seismicity in the Fiordland region is disproportionately distributed with intense activity describing a Benioff Zone that lies near the west of Lake Te Anau and dropping 80° to the east. The distribution of large events makes the northern part of the Fiordland region appear to be more distinctly seismic than the southern part of the Fiordland region. **(Anderson & Webb, 1994)**

The figure (Figure 6) shown below helps in understanding the numerous deep earthquakes that have occurred in New Zealand in the last 10 years.

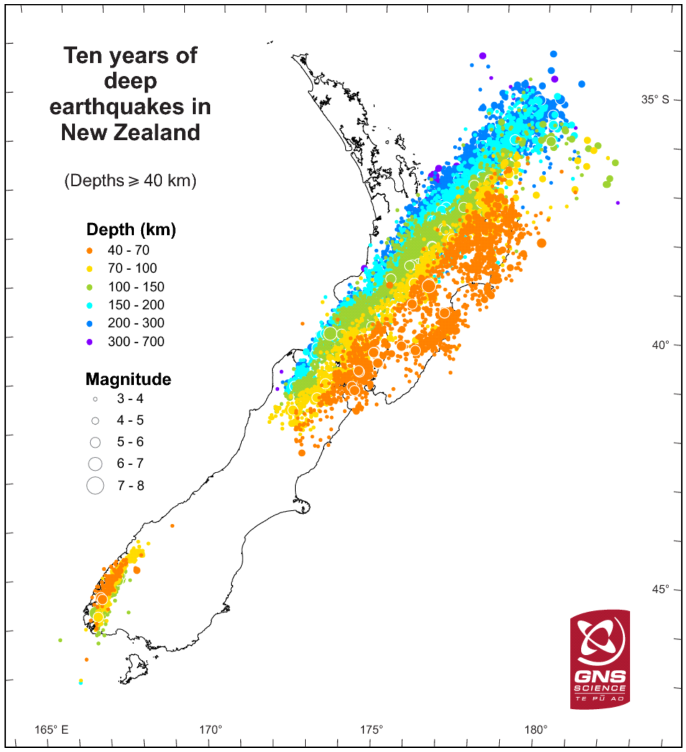

*Figure 6:-**Deep Earthquakes in New Zealand (GNS Science, n.d.-b)***

2.2.2 Shallow Earthquakes:-

Seismicity patterns of shallow events with depths >15km but <40km are dominated by small events, usually at the subducting slab, recorded in the Wellington region. With a lower depth of about 15km at the crustal seismicity, epicentral density in the Wellington region has a low concealing effect, and patterns such as events near Wairarapa Fault and the Cape Campbell cluster become evident.**(Anderson & Webb, 1994)**

Aftershock Sequences:-

Several large shallow earthquakes that have occurred in the past had aftershock sequences due to the shallow (>15km) seismicity pattern. The large, shallow earthquakes that occurred in the North Island during that period were mostly situated on the east coast. The group of four large earthquakes were of different kinds: the first event occurring near the upper part of the subducting plate, whereas the second event, which was the largest, was a mix of strike-slip faulting and thrust in the overlying Australian Plate. Many events are located around the offshore region of the Bay of Plenty. **(Anderson & Webb, 1994)**

North Island:-

The Central Volcanic Region is distinctive for the alignment of shallow earthquakes that drifts north-east from Mt Ruapehu. The alignment that occurs along the eastern stretch of the Central Volcanic Region is susceptible to volcanism between Mt Ruapehu and White Island.

A group of shallow seismicity earthquakes runs east –west from Mt Ruapehu to Mt Taranaki. This group of earthquake events lies between depths of about 15km to 40km, indicating that the region has an active lower crust. This group of seismicity overlaps with a strong gravity gradient that was interpreted as originating from the association of the normal continental crust (35km thick) towards the south against the somewhat thin continental crust (25km thick) located towards the north.

The Cape Egmont Fault Zone includes earthquakes running south-west of Mt Taranaki for a minimum of 50km that tend to coincide with the seafloor scarp, which is53km long and 1-5m high, hitting northeast-southwest. This Fault Zone is mainly normal faulting but the area north-west of Nelson includes earthquakes with reverse faulting mechanisms. **(Anderson & Webb, 1994)**

South Island:-

In the South Island, the dominant part of the structural component of the plate boundary is the Alpine Fault which presents a major seismic hazard for the South Island region. Here, earthquakes are formed due to the partitioned deformation between strike-slip faulting in the Marlborough Fault System and reverse faulting in the Buller region.

Some alignment of the epicenters may be associated with the active service faulting in the central South Island. Numerous earthquakes occur at the intersection between the Alpine and the Awatere Faults where the band of epicenters wanders slightly from the Alpine Fault towards the southern point, merging with the cluster of earthquakes in the Canterbury area.

The central section of the Alpine fault is somewhat aseismic. Fewer events have occurred in the section from Harihari to Jackson Bay, and less action is seen in the northern part of Harihari region. Towards the southern part of Jackson Bay, the shallow seismicity pattern is complicated due to its interaction with the Fiordland Benioff Zone. No tectonic feature is located with the alignments that occur in the Otago Range and Basin Province.

The central Canterbury region appears to be active due to the Marlborough Fault System. The earthquakes in this region, have depths of around <40km that occur within the Pacific Plate in the southern extension of the Hikurangi Subduction Zone. **(Anderson & Webb, 1994)**

- Wellington Region:-

A long lasting telemetered network offers high- quality data on the Wellington region. Numerous small events are recorded, which shows that the Wellington region is more seismically active than any other region in New Zealand.

The pattern of seismicity for shallow earthquakes, i.e. less than 40km shows a common fabric that runs parallel to the structural grain of the north-east running faults in the overlying Australian Plate, but no clear correlation is available concerning any specific fault.

There is a strong activity in relation to depths to 40km in the Wanganui Basin, whereas the Marlborough Sounds region is aseismic. Earthquakes occurring with depths less than 15km show Wanganui Basin is active to shallow depths, however the depth control is very poor for onshore regions. In the Wellington region, earthquakes that occur with depths <15km lie between Wairarapa and the Wellington Faults. Apart from the central band of shallow seismicity, the region between Wellington Fault and Kapiti Island is aseismic.

The cross-section regarding shallow activity in the Wellington region is subjugated by the upper active seismic boundary of the sub ducting Pacific Plate. **(Anderson & Webb, 1994)**

The figure (Figure 7) shown below helps with understanding the numerous shallow earthquakes that have occurred in New Zealand in the last 10 years.

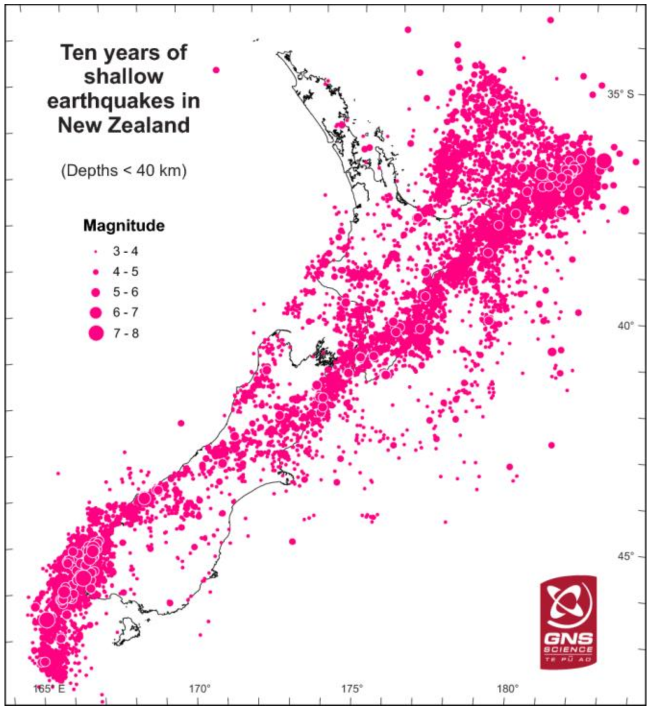

*Figure 7:- **Shallow Earthquakes in New Zealand(GNS Science, n.d.-b)***

•Where do Earthquakes occur?

The Hikurangi Benioff zone is manifested by the intense seismic activity at depths between 150 and 200km underneath the Central Volcanic Region; it has a severe discontinuity between northwest Nelson and it covers as far southwest as Westport. The Fiordland Benioff is noticeably more seismically active in its northern block, and the activity is strenuous in the zone towards the west of Lake Te Anau. The earthquakes also shape the eastern boundary of Central Volcanic Region and create an east-west band from Mt Ruapehu to Mt. Taranaki. The earthquakes extend to deep crustal levels due to the discontinuity in the region. The Cape Egmont Fault zone has been active but the Alpine Fault has been quiet in the section from Harihari to Jackson Bay in the South Island. There has been no seismic activity at the southern end of the Alpine Fault, where two subparallel lineaments seem to form boundaries between the western end of the Otago Range and the Basin Province. Earthquakes shallower than 15km are located between the Wairapapa Faults and Wellington. The contact between two fault systems of different movements near the Carterton area results in the occurrence of groups of earthquakes. The Benioff Zone underlying the Wellington region consists of a gap underneath the Wairarapa region. There is a severe northeast-southwest running boundary that is located between the aseismic Marlborough Sounds and the shallow active seismicity of the Wanganui Basin which is linked to subduction. New Zealand earthquakes are usually located by first using arrivals from the P waves, and then, the S phases are read from the horizontal seismometers. **(Anderson & Webb, 1994)**

This study uses earthquake (seismic data) located by the New Zealand National Seismograph Network operated by GeoNet, which is a collaboration between the Earthquake Commission and GNS Science. The New Zealand National Seismograph Network comprises primary sites that are located at roughly 100km spacing, with additional regional sites at places with geophysical significance.

The figure (Figure 8) below shows combined shallow and deep earthquakes across New Zealand.

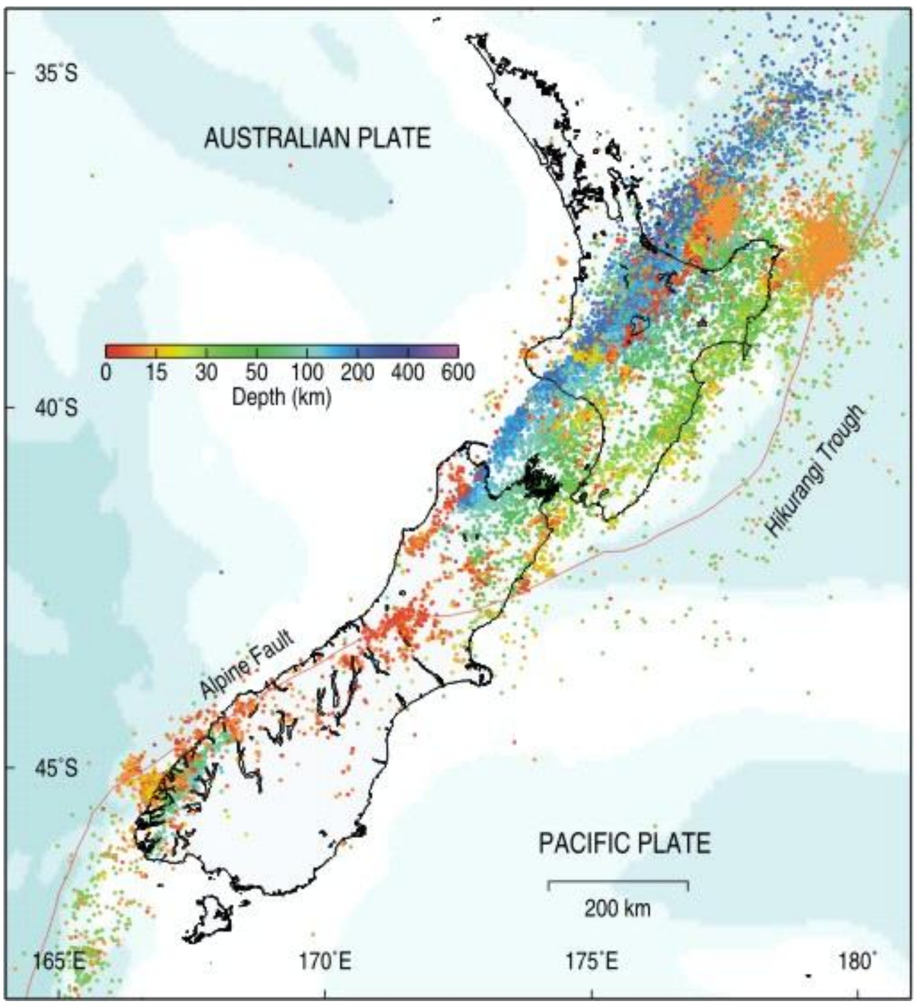

*Figure 8:- **The distribution of New Zealand earthquakes and the boundary of the Pacific and Australian tectonic plates. (Earth Science beta, 2014)***

New Zealanders know that earthquakes might occur at any time, depending on the location of the edge of the Pacific Plate or the edge of the Australian Plate. An extensive zone of deformation exists in most parts of the eastern and central North Island. The active boundary of the Australian-Pacific Plate spreads across New Zealand, producing volcanoes, active deformation, steep terrains and earthquakes. A major event could occur anywhere in the country and could possibly affect the whole economy as well as society because New Zealand is a small country and business, logistics and the infrastructure are all interdependent. **(GeoNet, n.d.)**

- How do earthquakes occur?

Natural hazards resembling earthquakes are mostly the result of propagating seismic waves beneath the surface of the earth. Seismometers located at different geographical locations record the vertical motion of surface waves, to measure earthquakes across the world. The earth's surface, which is also known as the "crust", is distributed into seven large tectonic plates. Several small sub-plates are formed from the larger plates and are in a continuous process of deformation and moving apart. Ground motions are of several types. The convergence and divergence processes help transform the plate boundaries. The divergence of the plate boundaries are formed when the tectonic plates move apart. The mountain ranges arise, from the process of convergence of different densities of tectonic plates. The process of transformation occurs, when the tectonic plates slide away from each other. Major earthquakes are caused by the transformation, convergence and divergence of plate boundaries commonly known as "faults". Stresses are caused by faults in the geological region. As the stresses are released, huge energy patterns produced are referred to as seismic waves or formally "seismic activity". Besides the faults, other causes of earthquakes are nuclear tests, mine blasts, volcanic activity etc. The point of origin from where the earthquake occurs is called the focus point. **(Azam, Sharif, Yasmin, & Mohsin, 2014)** The figure (Figure 9) shown below is another image of explaining how an earthquake occurs.

Earthquakes with less depth and a large magnitude are the most dangerous. Earthquakes that are deep and of a large magnitude are often considered to be less dangerous. These earthquakes can cause inner energy release patterns or can affect other layers, but the effect may not be spontaneous. Earthquakes with a shallow depth and small magnitude can be felt if they are persistent over a certain period of time.**(Azam et al., 2014)**

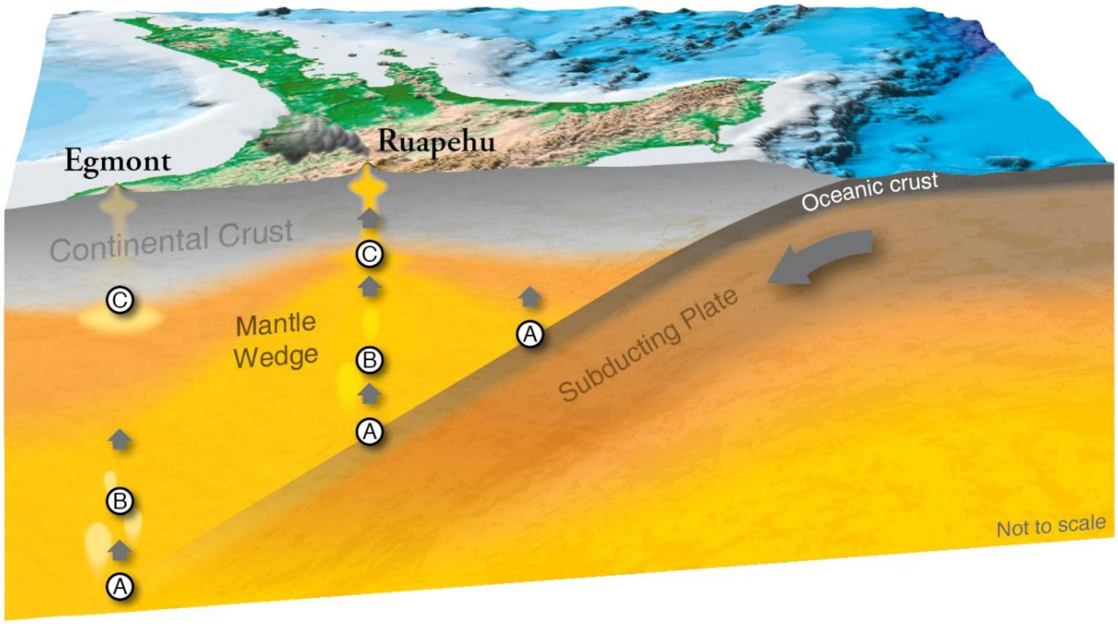

*Figure 9 :- **Plate subduction activity beneath New Zealand. A = zone of intense earthquake activity, B = hot liquid basaltic magmas, C = andesitic, dacitic or rhyolitic magmas.(Resilience, n.d.)***

•Why do Earthquakes occur?

The outer surface of the Earth is like a hard shell that are shattered into pieces, which are known as "tectonic plates". New Zealand is located on the boundary of two of these tectonic plates, the Pacific Plate and the Australian Plate. The plates are constantly crushing against each other, causing stresses to build up in the upper, brittle layers of the plate. Earthquakes occur due to sudden motion occurring in the crust, caused by the rapid release of accumulated stress along the break, or. Fault, in the Earth's crust. Under the region of New Zealand, the Pacific Plate is moving west-south-west at a rate of about 50mm a year. The interaction zone of the entire plate is a source of moderate to large earthquakes in New Zealand. **(GeoNet, n.d.)**

•Predicted types of Earthquakes:-

For earthquake assessment, the magnitude is assigned as the dependent variable whereas the latitude, longitude and focal depth are assigned as independent variables. The most widely cited quantity used to investigate earthquake size is magnitude. There are two predicted types of earthquake i.e. large-magnitude earthquakes and small-magnitude earthquakes.

- Seismic data:-

The method for acquiring and processing seismic data involves a few steps, such as the deployment of seismic sensors, actuation of the seismic source, receiving seismic signals produced by seismic sensors, and then producing seismic data. The acquiring and processing of seismic signals into raw seismic data has been done by GeoNet organization which is collaboration between the Earthquake Commission and GNS Science, New Zealand. Seismic data is acquired for analyzing the subsurface of the Earth and is also acquired in association with hydrocarbon exploration and production activities. Seismic data acquired with regards to analyzing the subsurface of the Earth is either collected from land or water. The seismic data is obtained by using an acoustic source that either consists of explosives or a seismic vibrator on land. The acoustic signals of the seismic data reflected by the numerous geological layers underneath the surface of the earth are called “traces”. These traces are sensed by thousands of geophones or sensors on land and hydrophones at sea. These reflected signals are then recorded and the results are the corresponding raw data. **(Baeten, Ferber, & Lengeling, 2002)**

The difficulty in modelling seismic data listed for small sub-regions is that a large number of important seismic events might not be listed in the historical data list in relation to all time periods used for training the network. **(Panakkat & Adeli, 2009)**

Studying the parameters of seismic data can help to better understand how earthquakes occur and also help with their prediction. Varying the parameters of seismic data can be correlated to the occurrence of earthquakes as well as the prediction of future, strong, main shocks. **(Djarfour et al., 2014)**

One of the main problems of seismic data processing is the removal of noise while preserving the main components of the data. Noise is present in all stages of seismic data from data acquisition to final processing. Noise is undesirable as it not only decreases the performance of the processing of seismic data but also leads to data interpretation errors. The most common noises are wind motion, electrical noise etc. Therefore, filtering methods are used to identify features that distinguish signals from noise. Attenuating and recognizing of noise component in seismic data still remains a challenge and complicates signal processing. Hence, to improve the quality of seismic data we need to develop reliable and useful methods. **(Djarfour et al., 2014)**

The seismic data used in this data set has been acquired from earthquakes that have occurred since 2010. GeoNet has used the GROPE technique to collect the seismic data from 2010 to 2011. After 2011, GeoNet started using the SeisComP3 technique. **(GeoNet, n.d.)**

- *GROPE:-*

The GROPE technique is the usage of P and S phases or the first-arriving crustal P and S phases. Four different depth/velocity structures were used in various parts of the country. **(GeoNet, n.d.)**

- *SeisComP3:-*

SeisComP3 technique is a seismological software used for data acquisition distribution, processing and interactive analysis. As a part of its location functionality, it uses two techniques: **(GeoNet, n.d.)**

- LocSAT:-

This uses one-dimensional model of the crust. The model is based on “earth model iasp91”. The “iasp91 reference model” is a parameterized velocity model that has been constructed to be a summary of the travel time characteristics of the main seismic phases.**(IRIS)**

- NonLinLoc

This is a software package that uses a three-dimensional model of the crust and based on the current model definition of “earth model nz3drx”. **(GeoNet, n.d.)**

Chapter 3:- Introduction to traditional Machine Learning Algorithms and Evolving Connectionist Systems.

- **Machine Learning Algorithms:-**

Machine Learning is a part of information science that is concerned with creating information models with data, representing knowledge, and the interpretation of knowledge and information from objects and processes. Machine learning consists of methods for model creation, knowledge extraction, model validation and feature selection. **(N Kasabov, 2007)**

1.1. WEKA

WEKA is also known as the Waikato Environment for Knowledge Analysis. WEKA has been created by researchers at University of Waikato, New Zealand. WEKA has been implemented in its new form since 1997. GNU General Public License is used in WEKA. The figure (Figure 10) below illustrates the WEKA interface.

*Figure 10:- **WEKA Interface***

The WEKA software has been written in JAVA and consists of a GUI to interact with data files. WEKA is a data-mining tool which provides various machine learning algorithms or techniques to be applied to real-world data-mining applications. The data file usually used in WEKA is ARFF file format. ARFF is also known as Attribute Relation File Format and includes special tags used to indicate differentiations in data files. The various algorithms that are used in WEKA are for classification, data pre-processing, clustering, regression and association rule mining.

Visualization tools are also included in WEKA. New machine learning algorithms can be created as WEKA is an open-source software tool. The main features of WEKA are that there are:-

•76 classification/regression algorithms.

•8 clustering algorithms.

•49 data pre-processing tools.

•10 search algorithms for feature-selection +15 subset/attribute evaluators.

• 3 association rule mining algorithms

•4 graphical user interfaces (GUIs).**(Dash, 2013)**

The four components included in WEKA are:-

•**Explorer: -** The Explorer interface offers a graphical front end to WEKA's components and routines as shown in the figure below (Figure 11). **(Sharma, Alam, & Rani, 2012)**

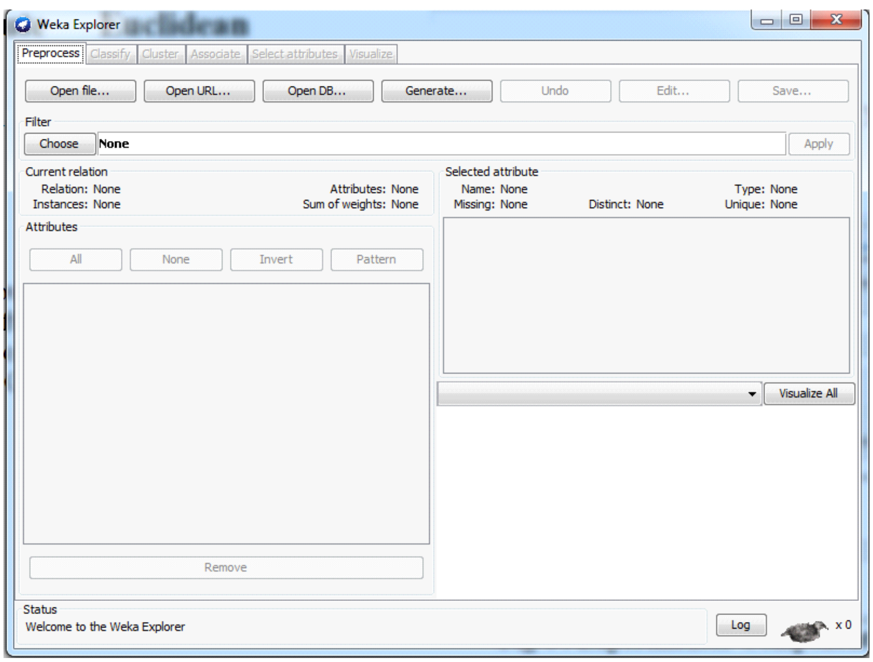

*Figure 11:-**Explorer Interface Component (Sharma, Alam et al. 2012)***

•**Experimenter: -** The Experimenter Component allows users to perform classification experiments as shown in the figure below (Figure 12). **(Sharma et al., 2012)**

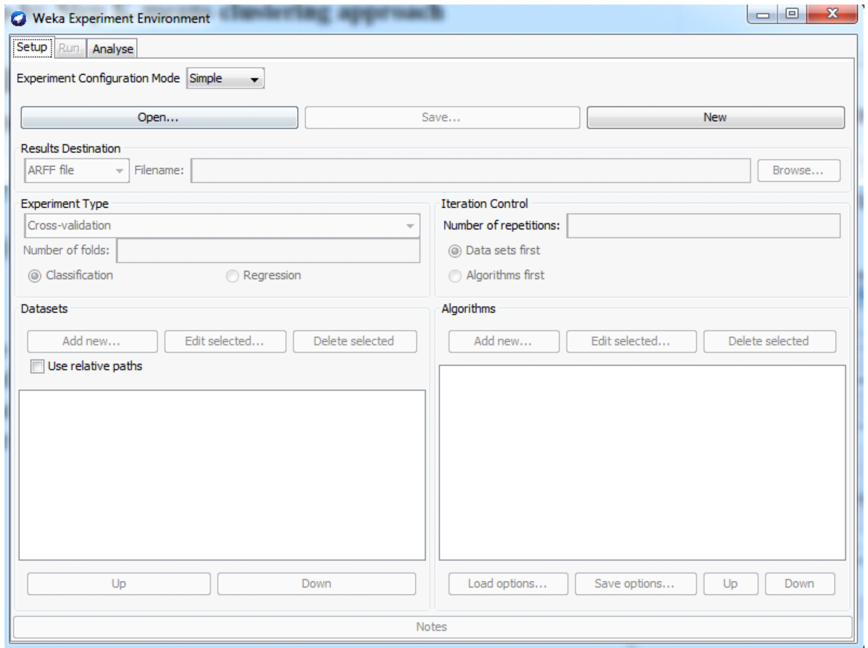

*Figure 12:- **Experimenter Component(Sharma et al., 2012)***

•**Knowledge Flow: -** The Knowledge Flow Component provides an option to the Explorer Component as a graphical front end to various algorithms used in WEKA as shown in the figure below (Figure 13). **(Sharma et al., 2012)**

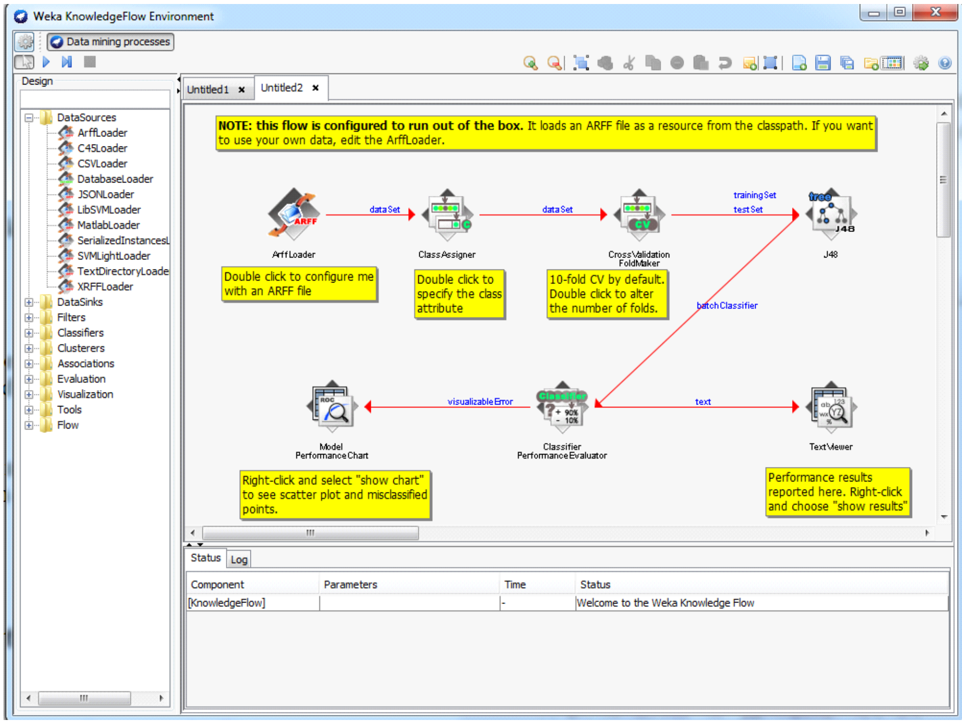

*Figure 13:- **Knowledge Flow Component(Sharma et al., 2012)***

•**Simple CLI: -** Simple CLI is a command line interface related to WEKA's routines as shown in the figure below (Figure 14). **(Sharma et al., 2012)**

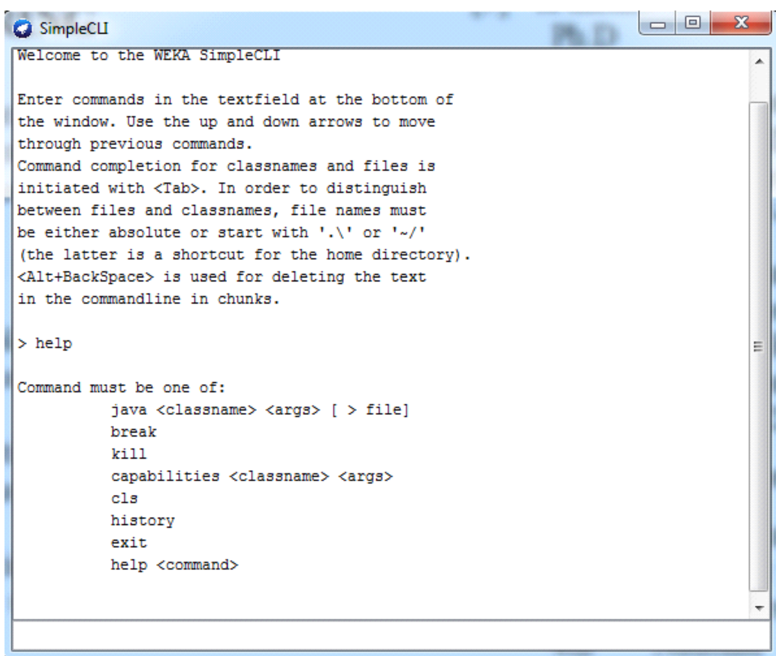

*Figure14 :- **Simple CLI(Sharma et al., 2012)***

•Algorithms

1.2.1. Naïve Bayes:-

In data mining, the Bayes method can be used as one of the classification methods. **(Dash, 2013)** The Bayesian learning framework allows dissimilar data to be combined in a model that can easily interpret and accommodate missing data. **(N. K. Kasabov, 2013)** Naïve Bayes is an extension of the Bayes theorem. Naïve Bayes assumes independence of attributes. The assumption that has been made may not be strictly exact during the classification of some data sets. **(Dash, 2013)**The assumption made is that important or dominant variables are independent given, that the class and the variables have a finite number of values. The Naïve Bayes classifier method is probabilistic in nature as the parameter values (e.g. feature probability distributions and prior class probabilities) have been approximated with regard to the relative frequencies of the training database. **(Bassis, Esposito, & Morabito, 2015)** Naïve Bayes is also known as simple Bayes, independence Bayes and idiot's Bayes. Naïve Bayes classifier method can easily be applied to huge data sets as they can be constructed without complicated iterative parameter estimation schemes. The Naïve Bayes classifier method can be easy to interpret and it is also robust. **(Dash, 2013)**

•J48 (Decision Trees):-

Decision trees are able to model complex non-linear decision boundaries. A large tree can be constructed and then pruned to reduce the cost-complexity criterion. Hence, the required tree is easy to interpret and can provide information that are related to the data structure. Small perturbations in the training data set or the construction procedure might result in huge changes in the classifier method for prediction. **(Webb, 2003)** Decision Trees are a non-parametric, supervised learning method used for both regression and classification. The goal is creating a model, formed by several leaves and a root based on the conditions, to categorize the set of rules and to select the branch to predict the value of the target variable by learning simple decision rules extracted from various data features. The pathway from root to leaf represents classification rules and the outcome of the leaf node can be obtained after the evaluation of all the conditions beside the path. The Decision Tree classifier method classifies new instances by sorting in a top-down manner, i.e. from root to the leaf node. A particular node located in the tree is used to specify the rule of some attribute and the branch that descends from that node corresponds to the value of the attribute. The measures consist of an average quantity of information for a particular event. It is robust in nature and small trees are preferred to large trees. The leafiness or the depth of the tree are the free parameters of the machine learning technique and they have to be adjusted to receive the maximum classification performance in the validation set to avoid over-fitting of the data. **(Lara-Cueva, Benítez, Carrera, Ruiz, & Rojo-Álvarez, 2016)**

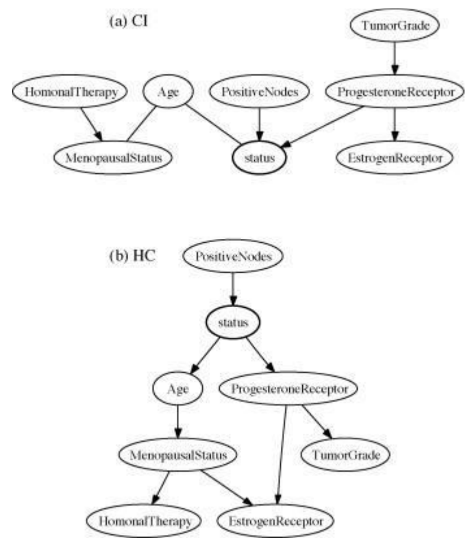

*Figure 15:- **Decision Tree(Dash, 2013)***

The J48 decision tree classifier is a simple C4.5 decision tree classifier. This classifier creates a binary decision tree and this approach is useful in the classification approach. This technique constructs a model for the desired problem and once the tree has been built, it is then applied to each row or tuple in the database that will result in classification of that tuple. The J48 classifier builds the tree by ignoring the missing values, i.e. the value specified by that item can be predicted from what is known about the attributed values for other records. The main idea is to divide the data into different ranges based on the attribute values for that item in the training sample. J48 allows classification wither through the decision trees or through the rules generated from them as shown in the figure above (Figure 15).**(Patil & Sherekar, 2013)**

- *Support Vector Machine (SVM, SMO):-*

Support Vector Machines were first proposed by Vapnik and his group at AT&T Bell laboratories. SVM optimizes the hyper plane positioning to reach the maximum distance from all of the data items located on both sides of the plane as shown in the figure below (Figure 16). **(N Kasabov, 2007)**

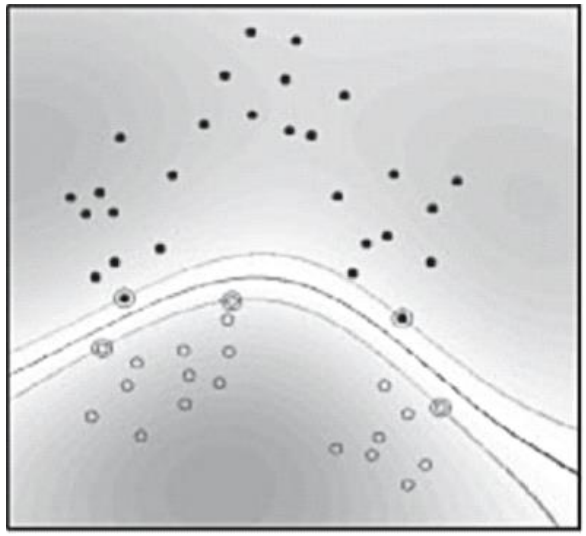

*Figure 16:- **Support Vector Machine (SVM) hyperplane(Nikola Kasabov, 2007)***

A SVM model includes a set of vectors that have been described by the kernel function that verges on border areas between samples that fit in different classes. These SVM models are good classification models, but the models find it difficult to adapt to new data and the extracted knowledge is also limited.**(Nikola Kasabov, 2007)** Support vector machines can be used for multi-class problems either by constructing linear discriminant functions or by using the binary classifier in a one-on-one situation or one–on-all situations simultaneously. Support Vector Machines consist of a class of algorithms that signify the decision boundary in pattern recognition problems, usually in terms of small training samples. Support Vector Machines define nonlinear functions implicitly, through description of the kernel function, whereas the radial basis function defines nonlinear function explicitly. Support vector machines can be also used to test separability. It may be important for high-dimensional data sets because the classes might be separable because of finite size training data sets. The classifier that has the possibility to achieve the linear separability on a training data set does not have to have a good generalization performance, but there may be insufficient data for characterizing the distributions. The degrees of freedom for the SVM model are the parameters of the kernel, the regularization parameter and the choice of kernel. For most kernels, it is quite possible to locate the parameters which show the separable classes. However, this does not seem to be a sensible strategy as it leads to over-fitting and poor generalization of the data. Support Vector Machines are implemented in high-dimensional spaces where the classification problem consists of the misclassification rate, and also the training data is representative of the test conditions **(Webb,**

2003). The SMO algorithm used in WEKA is a function-based support vector machine algorithm based on sequential minimal optimization. **(Reynolds, Kontostathis, & Edwards, 2011)** Training the support vector machine involves the use of large quadratic programming optimization problems (QP). SMO cuts this large QP problem into small bits of QP problems. These small problems are solved in an analytical manner, avoiding the time consuming numerical QP optimization. The amount of memory for SMO optimization is linear to the training size, thus allowing SMO to handle large data sets. As the computation of matrix is avoided, the SMO scales somewhere between the linear and quadratic according to the training set size for some applications, whereas the standard SVM algorithm scales somewhere between linear and cubic which are related to the training set size. SMO does not require any extra matrix storage. In place of the numerical quadratic programming (QP) that the old SVM algorithms use as the inner loop, SMO uses analytic QP steps. Therefore, large SVM training problems can be fitted inside the memory of the personal workstation. As no matrix algorithms are used in SMO, the SMO is less vulnerable to numerical precision problems. **(Platt, 1998)**

- *Nearest Neighbours (IBk):-*

K-Nearest Neighbours are supervised learning algorithms that can be used for classifying data samples that have the nearest training samples in a multi-dimensional feature space. The K-Nearest Neighbours algorithm is described as follows:-

- A set of pair features (e.g. ……) are well-defined, specifying each data point and each data points is then identified by class labels

- A distance measure is selected (e.g. Euclidean Distance or the Manhattan Distance) that will be used to measure the similarity of the data points that are based on all the features.

- The K-Nearest Neighbours are then found for the target data point by analyzing the similarity and then using the majority voting rule that will help in determining which class the target data point belongs to. **(Liang, 2014)**

The disadvantage of the K-Nearest Neighbours is that the resulting estimate isn't a true probability density, as the integral over all the available spaces diverges. The other disadvantage is that all training methods have to retained, thus leading to problems of computer storage and requiring huge amounts of processing for evaluating the densities of new values. **(Bishop, 1995)**

IBk is a type of K-Nearest Neighbour classifier which uses a similar distance metric. The users can specify the number of nearest neighbours in an explicit manner, either in the object editor or automatically by the process of leave-one-out cross-validation. Different search algorithms are used to speed up the process of finding the nearest neighbours. The linear search is the default search, but the other options included are ball trees, cover trees and KD trees. The parameter of the search method is known as the distance function. This includes the Euclidean distance, Minkowski distance, and Manhattan distance. Predictions that arrive from more than one neighbour are weighted according to the distance from the test instance and then two different formulas are implemented for converting the distance into weights. The training instances can be restricted by using the window size option. By the addition of new training instances, the old ones get detached, maintaining the number of training instances to the required size. **(Vijayarani & Muthulakshmi, 2013)**

- *Multi-Layer Perceptron (MLP):-*

Neural Networks aim to simulate the networks of biological neurons. One of the commonly used neural networks is the Multi-Layer Perceptron (MLP). The information is passed in a forward motion (feed-forward network), in one direction from the input layer (predictors) to the output layer (response variables). **(N. K. Kasabov, 2013)** The disadvantages of using Multi-Layer Perceptron is the local minima problem, inability to adapt to new data unless trained on old data, difficulty in rule extraction, and long training times when applied to large data sets. **(N Kasabov, 2007)** It is only the modification of the standard linear perceptron that uses three or more layers of neurons with non-linear activation functions. The model of every neuron consists of the non-linear activation function **(Dash, 2013).**

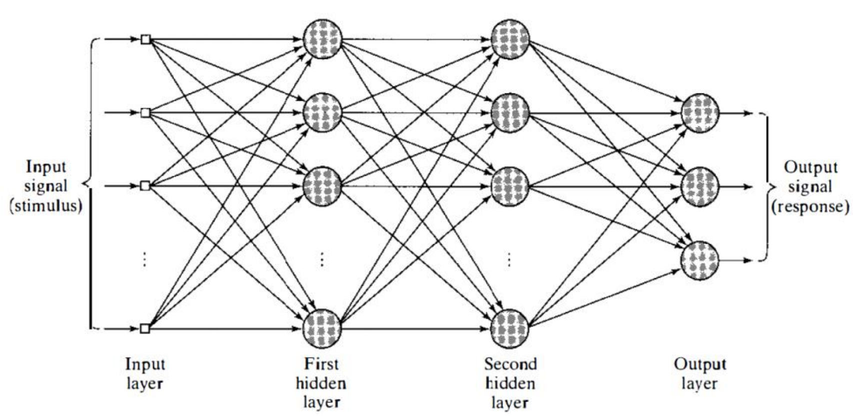

*Figure 17 :- **Architecture of Multi-Layer Perceptron(HAYKIN, 1994)***

The three distinct characteristics of Multi-Layer Perceptron are:-

- In the Multi-Layer Perceptron network, the model of each neuron consists of “nonlinearity” at the output end. The nonlinearity present is a smooth form which means it is able to be differentiated everywhere. The most common type of nonlinearity used here is sigmoidal nonlinearity, which is defined by the equation

$$y_j = \frac{1}{1+\exp(-v_j)} \tag{1}$$

Where y_j is the output of the neuron and is the input activity level of the neuron j. This is important, to ensure the input-output relation will not be reduced to a single layer perceptron.

- The Multi-Layer Perceptron Network consists of one or more layers of hidden neurons that may not be a part of the input or output and are not visible. These neurons facilitate the network to learn the various complex tasks by the extraction of meaningful features from the input vectors or patterns.

- The Multi-Layer Perceptron Network shows high-degree connectivity, through the synapses of the network. To make a change in the connectivity, a change in the population of synaptic weights or connections is required.

The flow of signals inside the Multi-Layer Perceptron network is in a forward direction on a layer-to-layer basis from left to right. Two types of signals are present, which are the Function Signals and the Error Signals. A Function Signal is an input signal which originates at the input end of the network. The signal is known as a function signal because it is assumed that the signal performs a function at the output end of the network and, when the signal passes through the layer, the signal is calculated as a function of the associated weights as well as the inputs given to that neuron. An Error Signal is a signal that originates at the output end of the network and it propagates in a backward manner. This signal is known as an error signal as it involves an error-dependent function.

The input neurons originate from the input layer of the network and the output neurons are located at the output end of the network. The hidden layers constitute the rest of the neurons which are neither a part of the input nor output unit. The hidden layer is set to perform two main functions, which are:-

- •Computation of Function Signals located at the output of the neuron, are stated as continuous non-linear function of the input neurons and the associated synaptic weights of that neuron.

- •Computation of an immediate estimate of gradient vector (i.e. the error surface gradients with respect to connected weights of the neuron) needed for the backward flow through the network.

The neurons from the input layer are then fed into the first layer of the hidden network which is made up of sensory units (source nodes); this resulting output is then fed into the next layer of the hidden network and so on, until it ends at the output of the Multi-Layer Perceptron network.**(HAYKIN, 1994)**

•Evolving Connectionist Systems:-

Evolving Connectionist Systems are an adaptive, incremental learning and knowledge representation system that evolves their functionality and structure. The core of the system includes a connectionist architecture which consists of neurons (information processing units) and their connections.

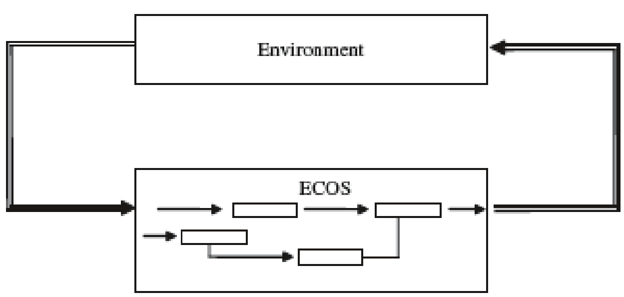

*Figure 18 :-**A simple ECOS Diagram(N Kasabov, 2007)***

Evolving Connectionist Systems evolve in time through interaction with the environment, i.e. they adjust their structure in reference to their environment as shown in the figure above (Figure 18). The main parts of the Evolving Connectionist Systems are given below:-

- •**Presentation Part: -** This part helps in the perception and filtering of the input information. The number of features or inputs varies from example to example.

•**Representation and Memory Part: -** This part stores the pattern information. This part is an evolving, multi-modular structure of NN modules that have been organized in spatially distributed groups.

•**Higher Level Decision Part: -** This part includes numerous modules, with each module making decisions based on a particular problem. The modules receive feedback from the environment and then make a decision based on the adaptation and functioning of the Evolving Connectionist Systems.

•**Action Part: -** The output from the decision modules is fed into the action part module and then the information is passed to the environment.

•**Self-analysis and the rule extraction modules: -** This module is used to extract abstract information that has been compressed from the representation module as well as from the decision module in numerous forms of abstract associations, rules etc.

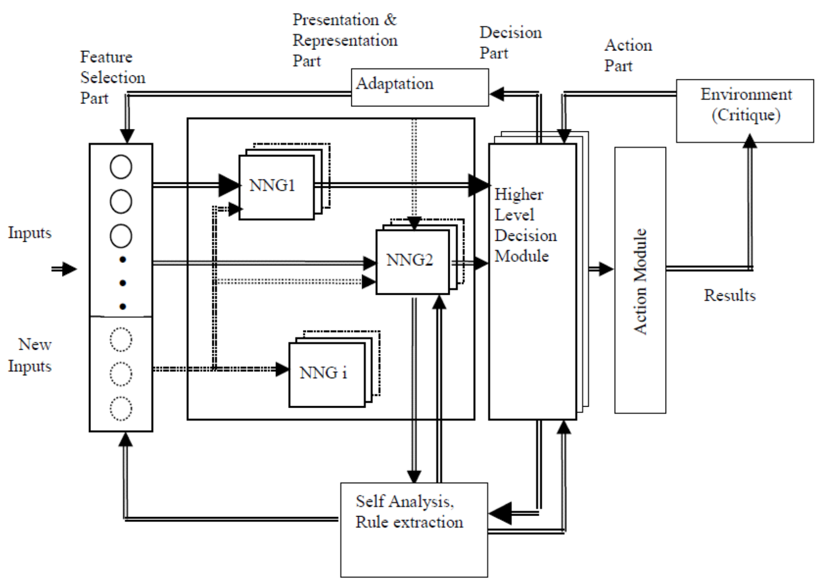

*Figure 19:-**Complex Diagram of Evolving Connectionist Systems(N. K. Kasabov, 1998)***

Evolving Connectionist Systems are a mesh of nodes or neurons with less connections between them that have either been pre-defined with prior knowledge or ‘genetic’ information. The internal connectionist structure consists of modules that are connected with these connections. Slowly, through the process of self-organization, the system becomes more and more wired. This then helps the network to store various exemplars or patterns from training examples. The node is then created and designated to represent individual examples if they are different from the previous examples (depending on the level of differentiation set based on dynamically changing parameters). **(N. K. Kasabov, 1998)**

2.1 Fuzzy Neural Networks

In 1965, fuzzy sets were introduced by Zadeh, to generalize the traditional membership of the element of the set from the binary {0,1} to the value in the interval [0,1]. **(Keller & Tahani, 1992)**

For the pattern classification problem, the construction of fuzzy rules from numerical data is done in two phases: a fuzzy partition of the pattern space and the identification of a fuzzy rule for every fuzzy subspace. If the fuzzy partition is too fine for the number of large fuzzy subspaces, then, some of the fuzzy rules cannot be constructed due to a lack of data points at the corresponding fuzzy subspaces. If the partition of the fuzzy subspace is too coarse for the number of small fuzzy subspaces, then the power of classification for the generated fuzzy rules is low. Therefore, for fuzzy rules that do not have a data point, it is difficult to create a subspace. **(Ishibuchi, Nozaki, & Tanaka, 1992)**

Fuzzy logic is a method that represents uncertainty by the distribution of the consequents and the antecedents of the rule. Numerous mechanisms can be used to transmit an uncertain input to the consequent. Thus, this leads to a computational burden for the system. The neural network has the ability to learn as well as extrapolate the complex relationships between the consequents and the antecedents for rules that consist of single and conjunctive antecedent clauses. For a natural mechanism to conflict resolution and to increase rule storage efficiency, multiple compatible rules can be deposited in a complex structure. **(Keller & Tahani, 1992)**

The fuzzy neural network can process both linguistic information from the human experts and numerical information from the measuring instruments. Finite parameter collection cannot determine the general fuzzy number. **(Ishibuchi, Nozaki, & Tanaka, 1992)** The fuzzy and the neural approaches are equal in time but differ only in the chosen approximator structure. Hence, to bridge the gap between the neural and fuzzy approaches, fuzzy neural networks are used. **(Diao & Passino, 2002)**

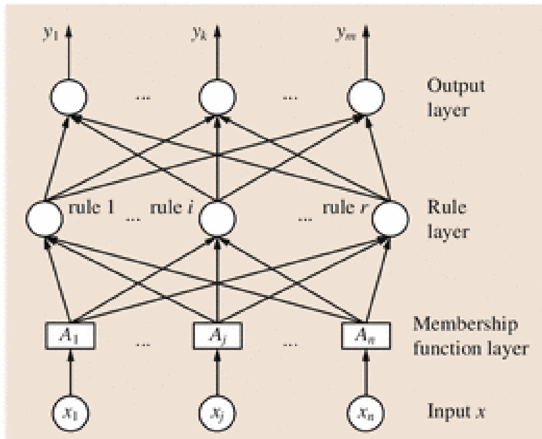

*Figure 20 :- **Structure of the Fuzzy Neural Network (FNN) (Kasabov, 2013)***

The general structure of the fuzzy neural network (FNN) is shown in the above figure (Figure 19). The Fuzzy Neural Network consists of four layers, which are the input layer, membership function (MF) layer, rule-based layer and the output layer.

The incoming input data can either be categorical or numerical. The membership function layer generates the membership functions for the numerical inputs, which is the conversion of numerical values to categorical values. Every rule node is then connected to every membership function node as well as the output node. A product function is performed on the inputs at the rule node. Between the rule layer and the input layer, the membership functions act as fuzzy weights. The tuning of the connection links between the output layer, membership function layer and the rule layer is done during the learning process. The concrete output of the network is produced in the output layer when every node receives inputs from the rule node which, in turn, is connected to the output node. **(Kasabov, 2013)**

2.2. NeuCom

NeuCom is a self-programmable, reasoning and learning computer environment based o connectionist (Neurocomputing) modules. The NeuCom software learns from the data, thus evolving new connectionist modules. These modules then can adapt to new data incoming in an online, incremental, lifelong learning mode. These modules are able to extract meaningful rules that help in discovering and understanding new knowledge in the respective applications. This NeuCom software is based on Evolving Connectionist Systems (ECOS), shown in the figure (Figure 21) given below.

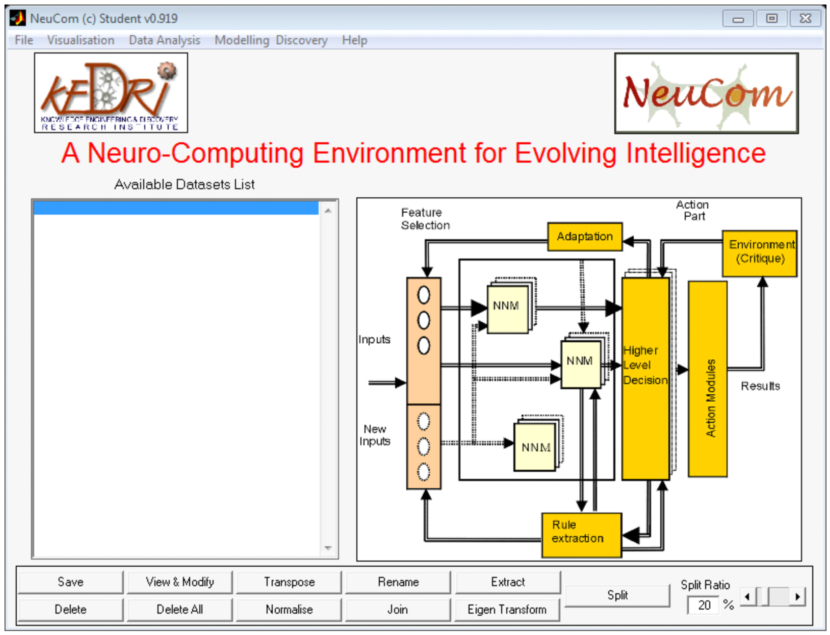

*Figure 21:- **The NeuCom Software***

The figure (Figure 22) shown below is an illustration of the DENFIS in the NeuCom software.

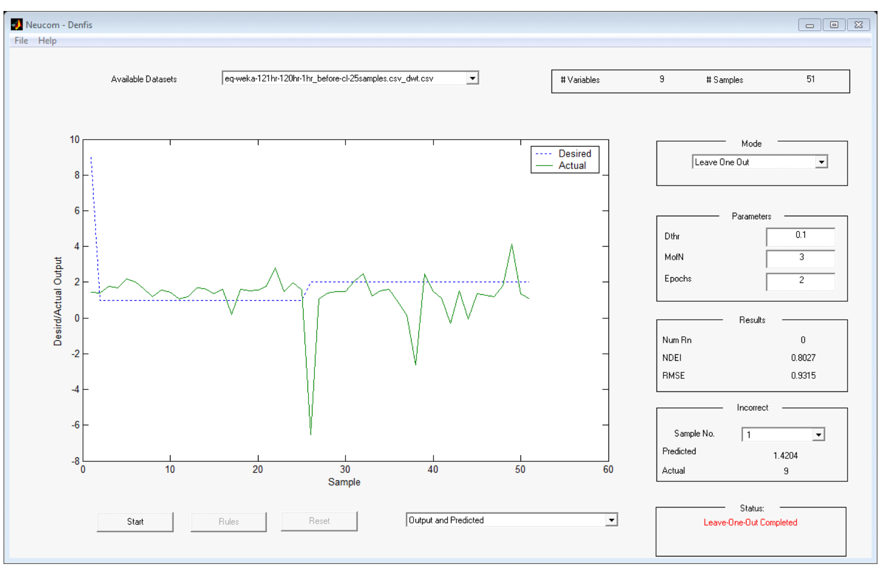

*Figure 22:- **DENFIS in NeuCom***

The figure (Figure 23) shown below is an illustration of the EFuNN in the NeuCom software.

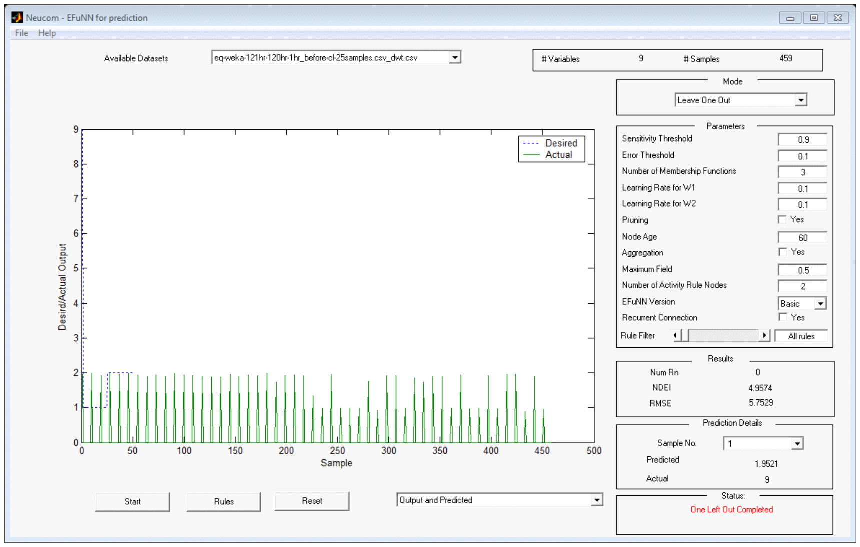

*Figure 23 :-**EFuNN in NeuCom***

The NeuCom software is used to solve complex problems. The problems can be classification, clustering, prediction, data mining, adaptive control and pattern discovery from databases that are present in a dynamic, multi-dimensional and changing data environment. The real-world applications areas include Medicine, Engineering, Arts and Design, Business, Education, Bio-Informatics and Science.

The NeuCom software can be used as both a DSS development environment and also as a decision support system. This software can be used as a decision support system (DSS), in which the users can specify the task and also define the data to be used to obtain a solution, or can be used as a DSS development environment to build sophisticated problems. The former case requires end users that have never programmed, but there are databases available for them to make a decision based on existing human knowledge and data. In the latter case, the users are professional system developers who can develop the environment for various applications. **(KEDRI, 2012)**

•*EFuNN (Evolving Fuzzy Neural Networks):-*

The fuzzy neural networks are connectionist structures interpreted in terms of fuzzy rules. The fuzzy neural networks include neural network characteristics such as training, adaptation, recall and so on. **(N Kasabov, 2007)** The evolving fuzzy neural network model (EFuNN) was introduced to create connectionist structures in an ECOS architecture. These are fuzzy logic systems that include five-layer structures shown in the figure (Figure 23) below. Connections and nodes are connected/created as the data examples are presented. A short-term memory layer can be used as an option through the feedback connection from the rule node layer (also known as case node layer). For the temporal relationships of the data to be memorized, we can make use of the feedback connection layer.

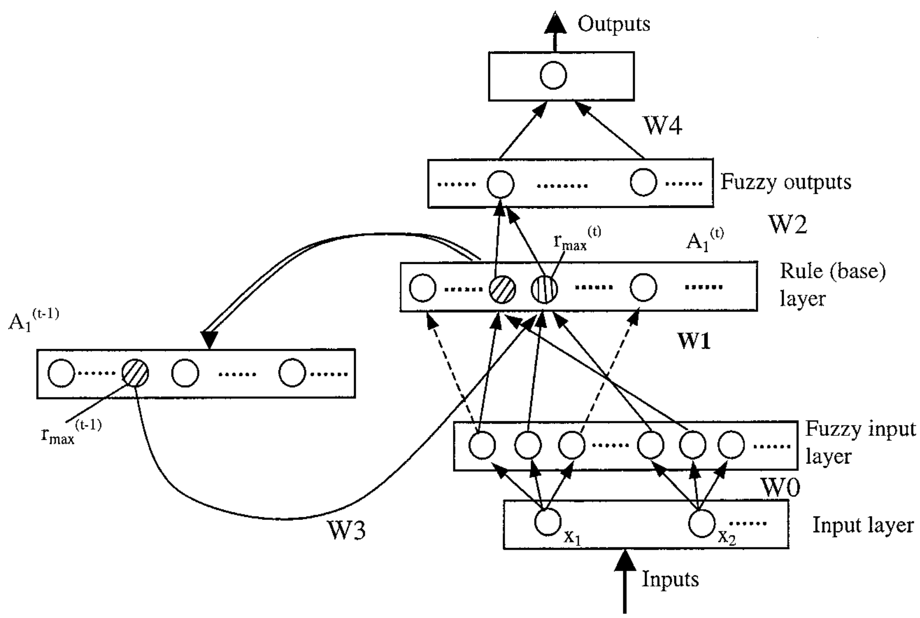

*Figure 24:- **Structure of EFuNN(N Kasabov, 2007)***

The Five-Layer structure is described as follows:-

•The first layer is the input layer which includes the input variables.

•The second layer includes the fuzzy inputs or the fuzzy input neurons as nodes that represent fuzzy quantization of every input variable space. There are different membership functions (MF) that can be applied to these neurons (Gaussian, triangular etc.). For example, two fuzzy neurons can be used to represent values such as “large” and “small”. The type and number of membership functions to be used can be modified in a dynamic manner. The function of the fuzzy input node is to transfer the input values into membership degrees as they belong to the MF.

•The third layer consists of rule nodes that can evolve through unsupervised and/or supervised learning. These nodes represent the prototypes of input-output data associations that are graphically represented as the associations of the hyperspheres from either fuzzy input and/or fuzzy output spaces. The linear activation function is used in this layer. Every rule node is defined by the two vectors of connection weights and , the latter being attuned to the output error through supervised learning, and the former being attuned to the similarity measure within the local area of the problem space through unsupervised learning.

•The fourth layer is the representation of the fuzzy quantization of output variables. The membership degrees in this layer are calculated by the combination of the saturated linear activation function and the weighted sum input function. The weighted output vector that consists of these membership degrees is associated with the input vector that belongs to each of the output membership functions.

•The fifth layer represents the real values of the variables of the output. **(N. K. Kasabov & Song, 2002)**

2.2.2 DENFIS (Dynamic Evolving Neuro-Fuzzy Inference System):-

Neuro-fuzzy inference systems include an inference method and a set of rules that are combined or embodied with the connectionist structure for better adaptation. The evolving neuro-fuzzy inference systems are systems where the inference mechanism and knowledge both change in time and evolve as more examples are being presented into the system. In this model, the knowledge is represented as both statistical features and fuzzy rules that are learned either in an online or offline manner, for a lifelong learning mode. **(N Kasabov, 2007)** DENFIS, which is used as both the offline and online model, uses the Takagi-Sugeno type inference engine. The inference engine used in DENFIS consists of 8fuzzy rules as shown in the equation below.

$$\left\{\begin{array}{c} If\ x_1\ is\ R_{11}\ and\ x_2\ is\ R_{12} and\ ...\ and\ x_q\ is\ R_q\ then\ y\ is\ f_1(x_1 x_2\ ....\ x_q) \\ If\ x_1\ is\ R_{21}\ and\ x_2\ is\ R_{22} and\ ...\ and\ x_q\ is\ R_{2q}\ then\ y\ is\ f_2(x_1 x_2\ ....\ x_q) \\ ......\ . \\ If\ x_1\ is\ R_{m1}\ and\ x_2\ is\ R_{m2} and\ ...\ and\ x_q\ is\ R_{mq}\ then\ y\ is\ f_m(x_1 x_2\ ....\ x_q) \end{array}\right\} \quad (2)$$

Where, i=1, 2….m, j=1, 2….q are fuzzy propositions as antecedents form fuzzy rules; , j=1,2,…..q are antecedent variables that have been defined over the universes of discourse , j=1, 2….q and , i=1,2….m; j=1,2….q, are the fuzzy sets that have been defined by each of their membership functions , i= 1,2….m; j=1,2……q. In the consequent part, is a consequent variable and the polynomial functions, i=1, 2….m are employed.

In the DENFIS offline and online models, all the membership functions are triangular-type functions that depend upon the parameters specified by the following equation:

$$\mu_x = mf(x,a,b,c) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a \le x \le b \\ \frac{c-x}{c-b} & b \le x \le c \\ 0 & c \le x \end{cases} \quad (3)$$

Where is the value of the cluster center on the dimension, and; d=1.2-2; the threshold value,, is a clustering parameter.

If the consequent functions are the crisp constants i.e. then we call such systems zero-order Takagi-Sugeno type fuzzy inference systems. The system will be known as a first-order Takagi-Sugeno type fuzzy inference system as are the linear functions. However, if these functions are not linear functions, then they are called high-order Takagi-Sugeno fuzzy inference systems. **(N. K. Kasabov & Song, 2002)**

Chapter 4:- Introduction to Spiking Neural Networks and NeuCube

•Spiking Neural Networks

•What is a Neural Network?

A neural network is defined as a machine that is aims to model the way in which the brain performs a function of interest or a particular task. The neural network is mainly implemented either by using electric components or simulation software present on the workstation. Neural network is also a massive parallel distributed processor that can store exponential knowledge and it resembles the brain in two ways:-

- Knowledge acquired by the network is through a learning process.
- Connection strength between the neurons are known as "synaptic weights" that are used to store knowledge.**(HAYKIN, 1994)**

A typical neural network works in a forward manner from the input layer to the output layer, with a hidden layer in between, as illustrated in figure (Figure 25) given below.

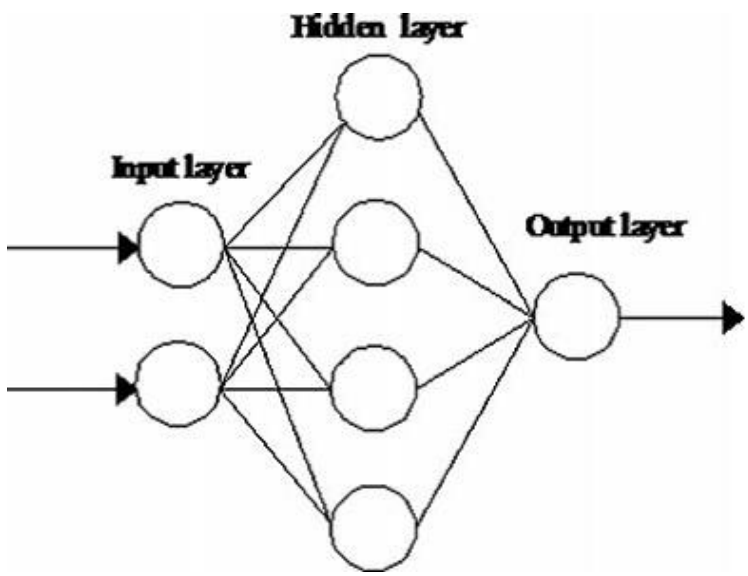

*Figure 25:-**A typical Neural Network(Zamani, Sorbi, & Safavi, 2013)***

•Artificial Neural Networks

Artificial Neural Networks physical cellular systems that can store, utilize and acquire the experiment knowledge. This knowledge is in the form of mappings or stable states embedded in the neural networks, which can later be recalled on the presentation of cues. **(Fullér, 1995)**

Artificial Neural Networks are defined as parallel computational models comprising processing units that are both interconnected and adaptive. The artificial neural networks are a fine-grained, parallel implementation of dynamic systems or nonlinear static systems. The most important feature of artificial neural networks is the adaptive nature that replaces "programming" with "learning by example" in solving problems. Computational models with this feature help the user to solve problems in areas where there is little or no understanding of the problem, given the training data is readily available. The other key feature is intrinsic parallel architecture which allows fast computation of solutions where either the networks are implemented on customized hardware or implemented on parallel digital computers.

The two major functions of Artificial Neural Networks are:-

- Learning: - Learning is the process of feeding new examples to the artificial neural network and then changing the connection weights.
- Recall: - Recall is the process of feeding new examples through the artificial neural network and then examining the output (reaction).

Most of the algorithms known in the area of artificial neural networks have been subject to the concept of Hebb. The concept of Hebb states that in unsupervised learning, the synaptic weights are increased if both the source and the destination neurons have been simultaneously activated. The equation is given as:-

$$w_{ij}(t+1) = w_{ij}(t) + co_i o_j \qquad (4)$$

Where, is weight of the connections between and neurons at the moment t, and and are output signals of neurons and at the similar moment. is the adjusted weight for the next time moment (t+1).

The learning system in artificial neural networks (S, W, P, F, L, and J) is well-defined as it consists of structure S, variable weights W, learning procedure L, function F, parameters P and goal function J. The system starts learning when the system optimists its structure and function while observing events from the problem space. Through the learning process, the system is able to develop its reaction to observed events and thus capture useful information that later can be represented as knowledge.**(N. K. Kasabov, 2013)**

Artificial neural networks are feasible for a wide range of application problems. The problem areas include speech synthesis and recognition, function approximation, pattern classification, image compression, forecasting and prediction, associative memory, adaptive inferences between complex physical systems and humans, clustering, nonlinear system modelling and control and combinatorial optimization. **(Hassoun, 1995)**

Neural networks have been applied in many domains of time-series predictions. Time-series data is a set of observations made in chronological order. The data source of time-series data always includes some amount of noise. The presence of noise requires clarification of time-series data. Filtering of such data is essential. Therefore, neural networks are used, as they have the ability to extract meaning from imprecise or complicated data that is difficult for other computer techniques or humans to understand. The pre-processing of raw data is another procedure of neural networks to help in refining the data into a compatible and easier format. **(Ali, 2014)**

- Spiking Neural Networks

Spiking Neural Networks (SNN) are defined as the third generation of Artificial Neural Networks. In these networks, spiking neurons send and receive information based on the timing of spikes (events) instead of the spike rate. Previous studies have gathered evidence that fewer spiking neurons are required for some problem-solving solutions compared to the neurons of previous generations. The spiking neural network models are considered to be more powerful when compared with previous neural network models. The simplest model of the spiking neuron assumes that the neuron fires when the threshold has reached a certain threshold. One more assumption is based on the fact that the output bit of the spiking neural network is given by either the non-firing or the firing of the specified output neuron during the specified time window. In spiking neural networks, the level of realism is increased by the neurons in a neural simulation, and the neurons also integrate the concept of time. **(Maass, 1997)** The main difference between the spiking neural network and the linear perceptron network is the action potential generated at the simulation time. The shortest distance between the two spikes defines the absolute refractory period of the neuron, followed by the relative refractoriness from where it becomes difficult to generate a spike. For the naturally bursting behaviour, the spikes are generated in a stereotypical manner by neurons, followed by repetitive single spikes, whereas the neurons with chattering behaviour fire stereotypical bursts of closed space spikes. The neurons with fast spiking behaviour fire periodic trains of action potentials with high frequency and without any adaptation, whereas for low-threshold spiking, neurons fire high-frequency spike trains with spike frequency adaptation. To achieve the desired output of the spiking neuron, the synaptic weights have to be adjusted. This step is related the training phase of the spiking neuron. **(Vazquez & Garro, 2015)**

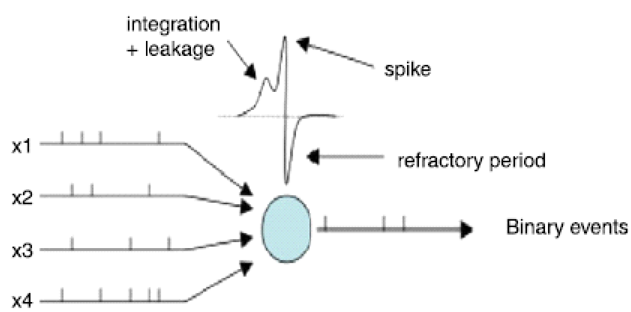

*Figure 26:- **Structure of Leaky-Integrate and Fire Model(N. Kasabov, 2014)***

The most significant problem is predicting the occurrence of abnormal situations in a multi-variate time-series environment. Hence, it is significant to predict when perilous variables reach undesirable changes. Classification of multi-variate time-series data is a challenging problem for traditional machine learning algorithms as they cannot consider the time component in the model building process. The main issue in temporal classification is where one variable's fluctuation regularly causes change in the other variable after a certain amount of time, i.e. either with delays or time lags. Hence, we enhance the method of traditional classification with the temporal component to help it to find the appropriate delay or time lag for every variable during the different stages of model building.

•NeuCube

The NeuCube is a 3D evolving neuro-genetic brain cube that consists of spiking neurons based on an approximate map of the functional and structural areas of interest in the human or animal brain. The gene information included in the NeuCube is of the form of gene regulatory networks related to the spiking neuronal parameters of interest. Different types of spatio-temporal data can be used to train the NeuCube, including: EEG, video-, sound- and image data, fMRI, complex multi-modal data and time-series data. Potential applications for the NeuCube are Brain Computer Interfaces; Neuro-Prosthetics and Neuro- Rehabilitation; fMRI-, EEG-and multi-modal brain data modelling and pattern recognition, modelling brain diseases; emotional and cognitive robots. Analyzing the internal structure of the NeuCube model can initiate new hypotheses on various spatio-temporal pathways. Brain-inspired Spiking Neural Networks have the potential to learn spatio-temporal data through the use of trains of spikes (i.e. binary temporal events) that have been transmitted between spatially located synapses and neurons. Both spatial and temporal information can be encoded in a spiking neural network as the locations of neurons and synapses, and neurons and time with regard to their spiking activity. **(N. Kasabov, 2012)**

Architecture:-

The important idea is to create a multi-modular integrated system that includes different modules, which consist of different genetic parameters and neuronal types that interact in a specific way with different parts of the brain and also with the different functions of interest (e.g. sound recognition: - motor-control, vision: - sensory information processing). Hence, the whole system works in an integrated mode for pattern recognition. The NeuCube architecture has a specific structure that includes a set of algorithms depending on various problems. Potential applications for the NeuCube are Brain Computer Interfaces; Neuro-Prosthetics and Neuro-Rehabilitation; fMRI-, EEG-and multi-modal brain data modelling and pattern recognition, modelling brain diseases; emotional and cognitive robots; prediction of ecological and geophysical data (i.e. seismic data, remote sensing data, aphid population data).

The neural model that NeuCube uses is the Leaky Integrate and Fired model (LIF). The parameters used are Drift, Mod, refractory Period, Threshold Firing. The gene regulatory network plays an important role in optimizing the parameters The most relevant parameters in the network are those that are involve in the learning processes, STDP the rate, in the deSNNs the mode value and the drifting rate (which is the amount to increase or decrease the synaptic weights between the reservoir and the output neurons), the LIF parameters specially the action potential threshold value and the leak value (the amount of membrane potential that the neuron lose when no spike arrives). The neurons are connected through a small world connectivity algorithm. Each neuron is connected to its nearest neighbors given a distance, e.g. $d<2.5$. All neurons can be pre and post synaptic neurons except input neurons, which are presynaptic neurons the connection weights are randomly initialized between [-1,1] for neurons in the reservoir, for input neurons is [-2,2] . Negative values indicate inhibitory connections and positive excitatory connections. We initialize 80% positive and 20% negative.

The synapses change (learning) according to the STDP rule, i.e. a connection between two neurons (presynaptic and postsynaptic) is strengthened according to the time between the spike emitted by the presynaptic neuron and the spike emitted by the postsynaptic neuron. If the presynaptic neurons fires before the postsynaptic neuron, then the synapse (connection) are strengthen; on the other hand, if the postsynaptic neuron fires before the presynaptic neuron, then the synapse are weaken. The amount of increment of decrease depends on an exponential which length depends on a time window

$$\mathrm{W(x)} = \mathrm{A}\exp\left(\frac{-x}{\tau}\right) \qquad for\ x > 0$$

$$\mathrm{W(x)} = -\mathrm{A}\exp\left(\frac{x}{\tau}\right) \quad for\ x < 0 \tag{5}$$

The final output of the network depends on the application and the training time, we cannot define the final output, even for the connections of the output neurons.

In the MLP you can simulate temporal information making a recursive neural network (ANNEX).The static data are in the forms of vector whereas the temporal data are in the matrix forms. So to feed the temporal data into these static algorithms, we concatenate the matrix data to make it into a single vector data.

The schematic diagram of NeuCube architecture is shown in the figure below (Figure 27).

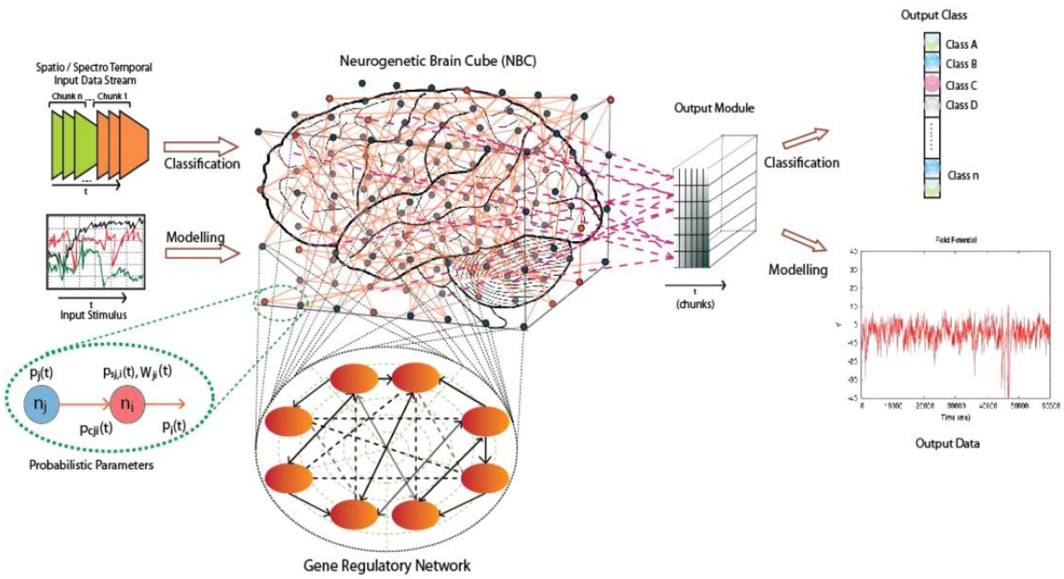

*Figure 27:- **Schematic diagram of the NeuCube software architecture(N. K. Kasabov, 2014)***

The NeuCube consists of the following modules:-

- Input Information Encoding Module
- Input Mapping Module
- 3D SNN Module (NeuCube Module)
- Output/Classification Module
- Gene Regulatory Network Module (GRN) (optional)
- Parameter Optimisation Module
- Visualization and Knowledge Extraction Module

The input module helps in transferring the input data into trains of spikes when the required data is processed. EEG, fMRI and the other brain data are directly transferred into the NeuCube module.

The NeuCube software is an approximate map of the study of relevant brain regions, along with its generic information, as 3D neuronal spiking architecture. Initially, the NeuCube can also include connections between the different areas of the brain. There are two types of spatio-temporal data that can be used as input into the NeuCube, which are:-

- Spatio-temporal data is used for measuring the activity of the brain (e.g. fMRI, EEG) when definite stimuli are present. This data is then entered into the corresponding spatially located areas of the NeuCube.

- Direct stimuli data or the time-series data, which will be first encoded into the spike trains and then fed as input into the input module.

The NeuCube architecture is three-tier architecture, where the gene regulatory network module (GRN) is at the lowest level, the NeuCube module at the middle level and the classification (evaluation) module at the highest level. In the NeuCube, the parameters of the neurons are controlled by the gene regulatory network (GRN). These genes are also affected by the activity of the spiking neurons. The neurons that are present in the NeuCube module are connected to the neurons in the output module in a two-way manner for the state of the cube to be interpreted/classified/recognized in the output module, and thus the result can have an effect on the further activity of the neurons in the cube as feedback.

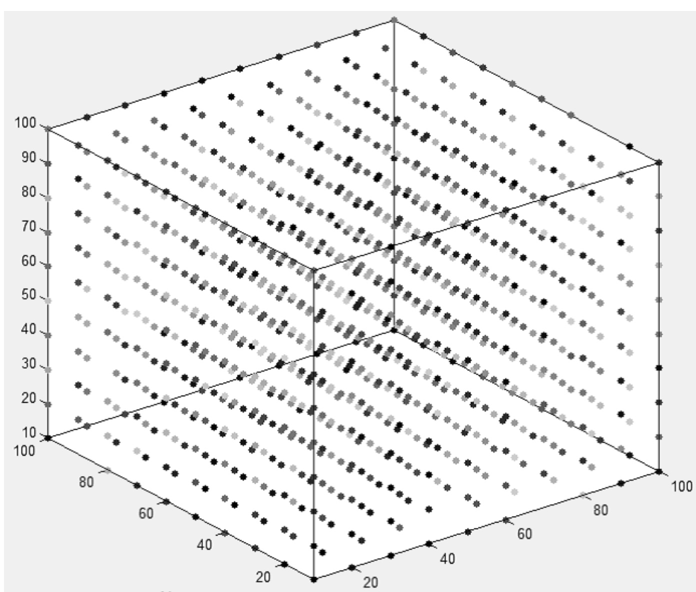

*Figure 28:-**A spiking neural network reservoir (SNNr) of 1000 neurons(Tu et al., 2014)***

The spiking neural network of 1000 neurons (10*10*10) is shown in the figure above (Figure 28). The size of the spiking neural network in the NeuCube can vary depending upon the data and the prediction task.

There is a two-stage learning process in the NeuCube. Firstly, there is the unsupervised learning which makes the spiking neural networks learn the spatio-temporal relations from input data by adjusting the respective connection weights in the spiking neural networks. Secondly, the supervised learning intends to learn the class information that is associated with every training data sample. The modelling process of the data in the NeuCube takes place in five steps, namely data encoding, initialization of the reservoir, unsupervised learning of the reservoir, supervised learning of the reservoir, and testing of the new sample. **(Tu et al., 2014)**

Learning:-

New neurons and connections are made from the new incoming data despite NeuCube having an initial structure. The NeuCube's functionality and structure evolve in time and with incoming data**. (N. Kasabov, 2012)** the learning process of the NeuCube occurs in two ways:-

- Unsupervised Learning: -

This stage is anticipated to encode the "hidden" spatio-temporal relationships from the input data into the neural connection weights. As per the Hebbian learning rule, if interaction between the two input neurons is persistent, then the connection between the neurons will be strengthened. The reservoir is trained using the spike-timing dependent plasticity (STDP) learning rule. The rule states that if neuron fires before neuron, then the connection weights from neuron to neuron will increase, whereas the connection weights from neuron to neuron will decrease. This confirms that the time difference for the input spiking trains that encode the temporal patterns of the original output signals will be captured by the asymmetrical connection weights, and the neuron firing state is the reservoir.

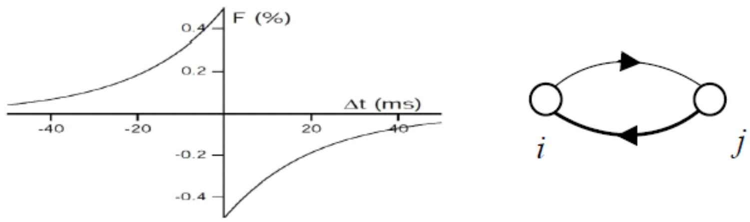

*Figure 29:- **STDP Learning Rule(Tu et al., 2014)***

In spiking neural networks, when the neuron fires, it emits a spike and then its potential is reset to 0. Every neuron that is connected to this firing neuron will receive a spike and its potential increases in relations to the connection weights of the current firing neuron. The potential of each of the neurons includes a small rate of constant leakage over time unless it becomes 0. As the learning process is finished, the connection weights included in the reservoir encode the temporal relationships from the input spatio-temporal data. **(Tu et al., 2014)**

- Supervised Learning: -

In this learning stage, we train the classifier using the class label information associated with the training samples. The Dynamic Evolving Spiking Neural Network is used as the classifier in this learning stage. The deSNN is used because it emphasizes the importance of the first spike which is similar to the biological system, and it is also computationally efficient. For every training sample, an output neuron is created and is connected with every neuron in the reservoir. The initial connection weights are set at zero. The establishment of the connection weights is done by the Rank-Order learning rule (RO). The potential of the neuron at time is calculated using the following equation

$$\mathrm{P(i,t)} = \sum mod^{order(j)} w_{j,i} \qquad (6)$$

Where, mod is the modulation factor and order (j) is arrival order of the spikes to the connection, is amongst all the spikes from all the connection to the neuron. A higher priority is specified for the first incoming spike at the output neuron by this learning rule. As the initial spike in the reservoir gets excited, the connection weights are set by the following equation

$$w_{j,i} = mod^{order(j)} \qquad (7)$$

After the arrival of first spikes, the connection weights are modified to the firing state of the resultant neuron in the reservoir. When the neuron fires, a spike is emitted towards all the output neurons that have a connection with it, and thus the connection weights between this neuron and the resultant output neurons strengthen, or else the connection weights weaken and the potential of the output of the neuron leaks as time elapses. As the potential of the output neuron exceeds a certain threshold, the spike is emitted. When the training is finished, the connection weights between the reservoir neurons and the output neurons encode both the support neurons and the spike order of the training sample. **(Tu et al., 2014)**

Experiment and Classification of Seismic Data:-

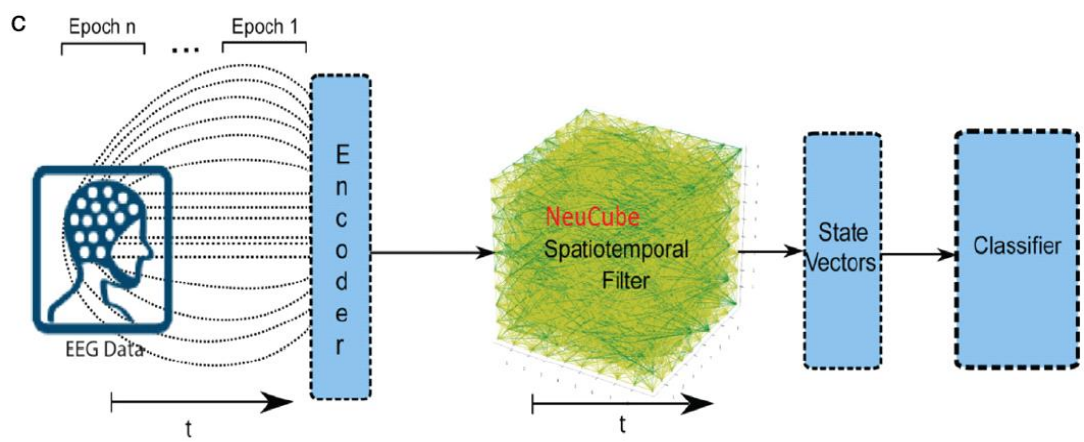

*Figure 30:- **Experimental Setting for the NeuCube model(N. K. Kasabov, 2014)***

Classification of the Seismic Data:-

The above figure (Figure 30) shows the classification process of the EEG NeuCube model. We classify the seismic data in a similar manner. The process is as follows:-

- Seismic data in the form of 100*24 matrix is prepared and then the determination of the label names for the output classes is done.

- After setting up the data, the input data is encoded using various encoding algorithms and then fed into the NeuCube module.

- Unsupervised training is done in the NeuCube using the STDP algorithm.

- After the training is finished, the connections between neurons can be visualized.

- The second phase consists of training the classifier.
- Visualization of the test result can then be viewed in order to understand the data.

For unsupervised learning, the NeuCube uses the Spike Timing Dependent Plasticity (STDP) algorithm and for supervised learning, the NeuCube uses the Dynamic Evolving Spiking Neural Network (deSNN) algorithm.

Dynamic evolving SNN (deSNNs):-

The synaptic weight parameter is dynamic in nature and changes with each pre-synaptic spike due to the two short-term synaptic plasticity processes, which are depression and facilitation. This empowers the neural networks capability to carry out computations on the spatio-temporal data.

The disadvantage of the eSNN learning is that the connection weights of every synapse are adjusted only once in the model based on the first rank order. This is not efficient for complex spectro- and spatio-temporal data. For the implementation of the dynamic synapses, the connection weights arriving over time have to be further refined on the spike, which leads to spike time learning (STDP).**(N. Kasabov et al., 2013)**

Chapter 5:- The Proposed Methodology

•Proposed Methodology Framework

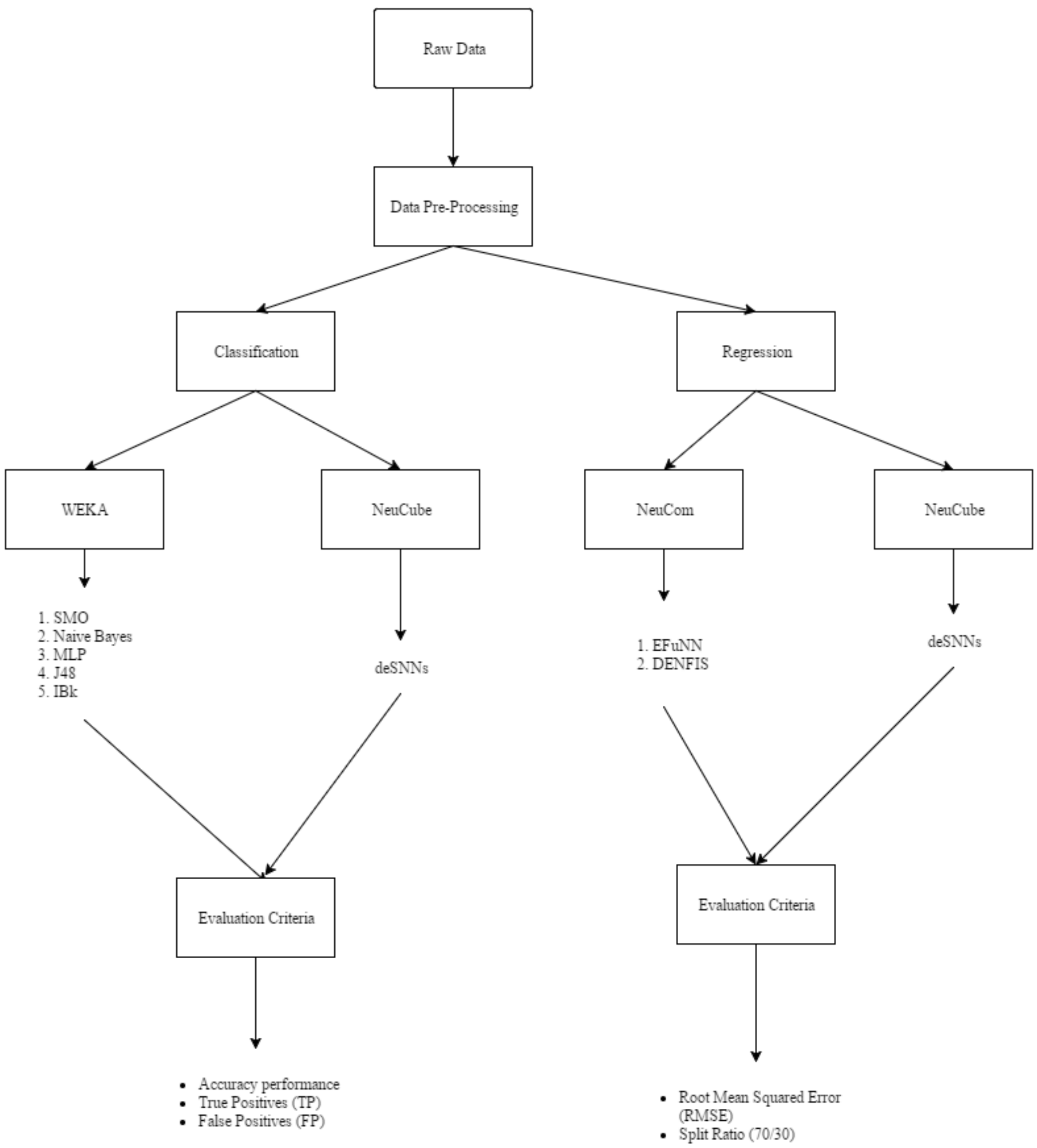

*Figure 31:- **The proposed Methodology Framework***

As shown in the above figure (Figure 32), the first step was the collection of raw data, which is the seismic data. The seismic data set used in this study was collected from the GeoNet website (www.geonet.org.nz). After the data was collected, we moved to the next process of data pre-processing, one of the important processes. After the data pre-processing, we moved onto the next phase, which is divided into two processes that are:-

- Classification
- Regression

•Classification:-

Classification is to group the entities (e.g. the magnitudes of the earthquakes, specimens or patients) and it also works with the labelled data where the class values are represented as features of the data set.

In this study, the comparative analysis of the classification of the earthquake (seismic data) was done by using two software programme's: WEKA and NeuCube. In WEKA the following data-mining algorithms were used: SMO (Support Vector Machines), J48 (Decision Trees), Naïve Bayes, IBk (K-Nearest Neighbours) and MLP (Feed-Forward Neural Network), whereas the NeuCube used Spiking Neural Networks. SVM classifier algorithm was used in this study as it is used for multi-class problems either by using a binary classifier or a c linear discriminant function. **(Webb, 2003)** J-48 algorithm was used in this study because a model can be created with leaves and a root based on the conditions or requirements for categorizing a set of rules to be able to predict the value of the target variable and to classify new instances in a top-down approach as it uses the non-parametric supervised learning method. **(Lara-Cueva et al., 2016)** Naive Bayes algorithm was used in this study as the classifier assumes the independence of attributes and it does not use complicated iterative parameter estimation schemes. **(Dash, 2013)** IBk is a supervised learning K-Nearest Neighbours algorithm which uses the linear search. The parameter for this search method is the Euclidean distance. It classifies the data samples in a multi-dimensional space according to its distance from other training samples. **(Vijayarani & Muthulakshmi, 2013)** Multi-Layer Perceptron is a feed-forward neural network which can learn and classify new incoming data based on the trained data which then can be useful to extract the features. **(N. K. Kasabov, 2013)**

Detection of an event is not adequate on its own, so detection of an event in real-time is necessary. Hence, to make an intelligible comparative analysis of the seismic data, we identify the characteristics of the events that need to be analyzed. Therefore, the detection of an event is established based on the results of the following analysis:-

- Spatial analysis used to infer the location
- Temporal analysis used for real-time detection of the event.

In WEKA we classify the results based on the following:-

- ***True Positives (TP):-*** Events detected by the system and confirmed by the required organization.
- ***False Positives (FP):-*** Events detected by the system and not confirmed by the organization.
- ***False Negatives (FN):-*** Events reported by the organization but not detected by the system.
- ***True Negatives (TN):-*** Events that did not occur and have not been detected by the system.
- ***Precision:-*** Precision is defined as the ratio of the appropriately detected events to the total number of detected events:

$$\text{Precision}=\frac{TP}{TP+FP} \tag{8}$$

- ***Recall:-*** Recall is defined as the ratio of the appropriately detected events to the total number of occurring events:

$$\text{Recall}=\frac{TP}{TP+FN} \tag{9}$$

•***F-Measure/F-Score:-*** F-Measure/F-Score is defined as the harmonic mean between Precision and Recall:

$$\text{F-Score}=2 * \frac{Precision*Recall}{Precision+Recall} \tag{10}$$

(Avvenuti, Cresci, La Polla, Marchetti, & Tesconi, 2014)

•Regression:

In Regression, a set of feature values is used to predict certain values (e.g. predicting the occurrence of earthquakes).

In this study, the comparative analysis of the regression of the earthquake (seismic data) was done by the two software programmes used: NeuCom and NeuCube. The fuzzy neural network algorithms used in NeuCom are EFuNN (Evolving Fuzzy Neural Networks) and DENFIS (Dynamic Evolving Fuzzy Inference System) whereas NeuCube used Spiking Neural Networks. The EFuNN network algorithm is used in this study because it consists of a feedback connection layer which can help in learning the temporal relationships of the data, and can also extract features or knowledge rules for further experimentation. **(N. K. Kasabov & Song, 2002)** DENFIS network algorithm is used in this study, with new incoming data, this algorithm evolves and adapts itself to the inference mechanism and knowledge. Also, the knowledge represented in this network is either fuzzy rules or statistical features that can be learnt in an offline or online mode.**(N Kasabov, 2007)**

The classification and the regression experiments done in NeuCube used the deSNNs algorithm.

After performing the classification and the regression experiments on the seismic data, we performed comparative analysis based on the evaluation criteria set. The evaluation criteria for the classification of the seismic data is the Accuracy Performance, True Positives and False Positives. The evaluation criteria for the regression of the seismic data is the Root Mean Squared Error (RMSE).

While performing the experiment for classification, the split ratio was set as 70% of data for training and 30% of data for testing and we used the leave one out cross-validation for the regression experiment.

Root Mean Squared Error (RMSE):-

The Root Mean Squared Error (RMSE) is used to measure the difference between the observed and predicted values. The value of RMSE is between 0 and infinity, where, 0 is the perfect fit. The equation for RMSE is given as:

$$\text{RMSE}=\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2/n} \qquad (11)$$

Where, n is the number of observations and and are the desired and the computed output.**(Zamani et al., 2013)**

- **Why the Proposed Methodology?**

We have used this proposed methodology to understand the way in which data mining has become important in certain applications based on time-series data. We have tried to compare a relative standard technology known as traditional machine learning algorithms, with the evolving connectionist systems based on fuzzy neural networks and lastly, the NeuCube which is based on spiking neural networks which are the future of data mining. Prediction of an earthquake, which is a natural calamity, is very important due to their regular occurrence in New Zealand. Neural networks are useful in solving some pattern recognition problems which are difficult to formalise. Hence, we are trying to determine under which conditions the neural network can try to predict or detect the presence of a seismic event. The movement to spiking neural networks was because of its capability to learn from new examples presented to it. **(Romeo, 1994)** The dynamic nature of the seismic data set provides motivation for the use of different classifier strategies for the ability to learn, adapt or modify decisions. **(Orozco-Alzate, Acosta-Muñoz, & Londoño-Bonilla, 2012)** We have used a very small data set sample from Canterbury as the area is very prone to the occurrence of earthquakes. This proposed methodology will help us to understand which algorithm provides better results regarding the prediction of earthquakes and, with further research and collaboration with the Geological Department of New Zealand, we can probably have a model that will or may be able to predict the occurrence of earthquakes in the near future.. The NeuCube model used in this study is in its preliminary stages from the point of view of prediction of seismic events.

- **How will this methodology benefit other users?**

There is a huge gap between the deployment of custom solutions and the research achievements stated in the literature at geology research laboratories. **(Orozco-Alzate et al., 2012)** Some standard machine learning algortihms and software packages have been used in this methodology that can help the researchers of geological and ecological science to further their development in researching on the occurrence of various types of earthquakes. The standard algorithms help in the understanding and comaprison of the new software "NeuCube" with the machine learning algorithms for the researches to choose on which software they can use.

Chapter 6:- Experiment Results and Analysis

•Experiment Data set

The experiment is conducted with the data set taken from the New Zealand National Seismograph Network. The data set that has been collected is of the earthquakes that have occurred since 2010.

*Figure 31 :- **The New Zealand National Seismograph Network(GeoNet, n.d.-b)***

The table below (Table 1) consists of 24 samples of seismic data collected from the Canterbury area in New Zealand. The data consists of 12 low-magnitude earthquakes and 12 high-magnitude earthquakes, and has been extracted or collected from the GeoNet website (www.geonet.org.nz), which is a collaboration between the Earthquake Commission and GNS Science. Yes, the preprocessing was applied before the data was fed into the NeuCube. The features used in this data are of numerical in nature. Classification process classifies the data into high magnitude and low magnitude earthquakes and the regression process gives the output as the magnitude of the earthquake. The encoding that was used in the NeuCube was the Address Event Representation encoding method. In NeuCube we used 4 channels with 100 features and the data in Weka consists of 400 features.

Public ID	Date	Magnitude	Depth (km)
3366146	September 3 2010	7.1	11
3450113	January 19 2011	5.1	9
3468575	February 21 2011	6.3	5
3474093	March 5 2011	5.0	10
3497857	April 16 2011	5.3	9
3505099	April 29 2011	5.2	11
3525624	June 5 2011	5.5	9
3528810	June 13 2011	5.9	9
3591999	October 9 2011	5.6	8
3631359	December 23 2011	5.8	10
2015p012816	January 5 2015	6.0	5
2015p305812	April 24 2015	6.2	52

*Table 1:- **The seismic data set has been collected from the region of Canterbury, New Zealand since 2010 and shows 24 earthquakes in total,12 low-level and 12 high-level earthquakes (shown above) taken from the website www.geonet.org.nz***

The table below (Table 2) consists of 25 samples of seismic data collected from the Canterbury area, New Zealand, and shows 12 low-magnitude earthquakes and 13 high-magnitude earthquakes. The data has been extracted or collected from the GeoNet website (www.geonet.org.nz), which is a collaboration between the Earthquake Commission and GNS Science.

Public ID	Date	Magnitude	Depth (km)
3366146	September 3 2010	7.1	11
3450113	January 19 2011	5.1	9
3468575	February 21 2011	6.3	5
3474093	March 5 2011	5.0	10
3497857	April 16 2011	5.3	9
3505099	April 29 2011	5.2	11
3525624	June 5 2011	5.5	9
3528810	June 13 2011	5.9	9
3591999	October 9 2011	5.6	8
3631359	December 23 2011	5.8	10
2015p012816	January 5 2015	6.0	5
2015p305812	April 24 2015	6.2	52
2016p118944	February 14 2016	5.7	8

*Table 2 :-**The seismic data set has been collected from the region of Canterbury, New Zealand since 2010 and shows 25 earthquakes in total, 12 low-level and 13 high-level earthquakes (shown above) taken from the website www.geonet.org.nz***

- Results
- Classification

Here, the comparative analysis for the classification of seismic data is based on the classification accuracy, the F-Score, True Positives and False Positives between the NeuCube software (Spiking Neural Network) and WEKA software (Traditional machine Learning Algorithms). The algorithms shown in the table below (Table 3) have been used due to their similarity to the linear classifiers.

There were 10 runs performed for each algorithm. The standard deviation has been calculated for each run and the average standard deivation has been taken across 10 runs.

The seismic data set used for this experiment consists of 25 samples: 12 low-magnitude earthquakes and 13 high-magnitude earthquakes with 1Hz frequency and 100 data time points. The results for the classification of the seismic data (one hour before the earthquake occurs) are shown in the table below (Table 3).

Method	Performance				
	Other parameters	Accuracy (%)	F-Score	TP	FP
SMO(SVM)(WEKA)	SVM Kernel: Polynomial Degree	49.14±17.53	0.515	5	7
Naïve Bayes (WEKA)	Use Kernel Estimator: False	53.71±17.15	0.520	6	6
J48(Decision Trees) (WEKA)	Pruning Confidence:0.25, Number of Folds:3	49.71±19.39	0.599	8	4
IBk(k-Nearest Neighbours) (WEKA)	Nearest Neighbour Search Algorithm: Linear NN Search	37.138±15.97	0.421	3	9
MLP (WEKA)	Number of Hidden Units:3, Number of Training Cycles:500	45.71±17.00	0.520	6	6
NeuCube	Mod:0.8,Refractory Time:6,STDP:0.001, Drift:0.005,Potential Leak Rate:0.002, Firing Threshold:0.4	80	0.791	9	3

*Table 3:- **The comparative results of the 25 samples (one hour before the earthquake occurs) with 1 Hz frequency and 100 data points between WEKA and NeuCube including the parameters used.***

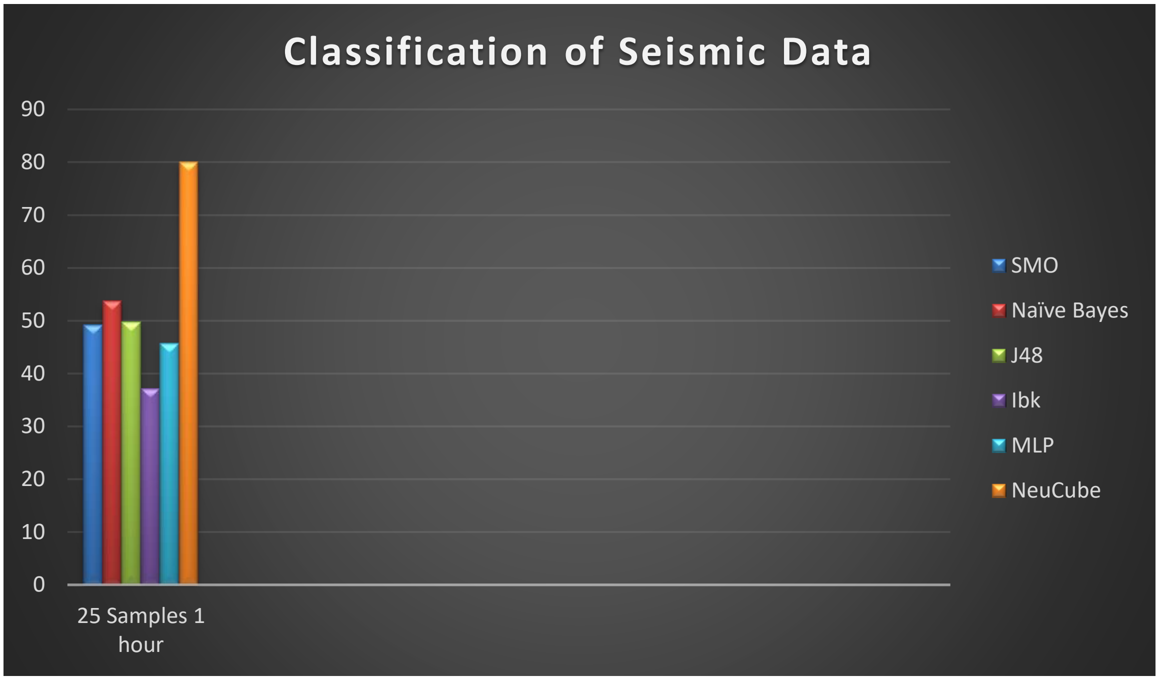

*Figure 32:-**The comparative results of the 25 samples (one hour before the earthquake occurs) with 1 Hz frequency and 100 data points between WEKA and NeuCube.***

The seismic data set used for this experiment consists of 24 samples: 12 low-magnitude earthquakes and 12 high-magnitude earthquakes with 1Hz frequency and 100 data time points. The results for the classification of the seismic data (one hour before the earthquake occurs) are shown in the table below (Table 4).

Method	Performance				
	Other parameters	Accuracy (%)	F-Score	TP	FP
SMO(SVM)(WEKA)	SVM Kernel: Polynomial Degree	54.16 ± 19.75	0.541	6	6
Naïve Bayes (WEKA)	Use Kernel Estimator: False	61.90 ± 13.10	0.541	6	6
J48(Decision Trees) (WEKA)	Pruning Confidence:0.25, Number of Folds:3	52.97 ± 18.11	0.541	7	5
IBk(k-Nearest Neighbours) (WEKA)	Nearest Neighbour Search Algorithm: Linear NN Search	45.23 ± 16.67	0.365	3	9
MLP (WEKA)	Number of Hidden Units:3, Number of Training Cycles:500	55.95 ± 17.83	0.583	7	5
NeuCube	Mod:0.903,Refractory Time:6,STDP:0.01, Drift:0.0001,Potential Leak Rate:0.001, Firing Threshold:0.4	91.67	0.92	11	1

Table 4: ***The comparative results of the 24 samples (one hour before the earthquake occurs) with 1 Hz frequency and 100 data points between WEKA and NeuCube including the parameters used.***

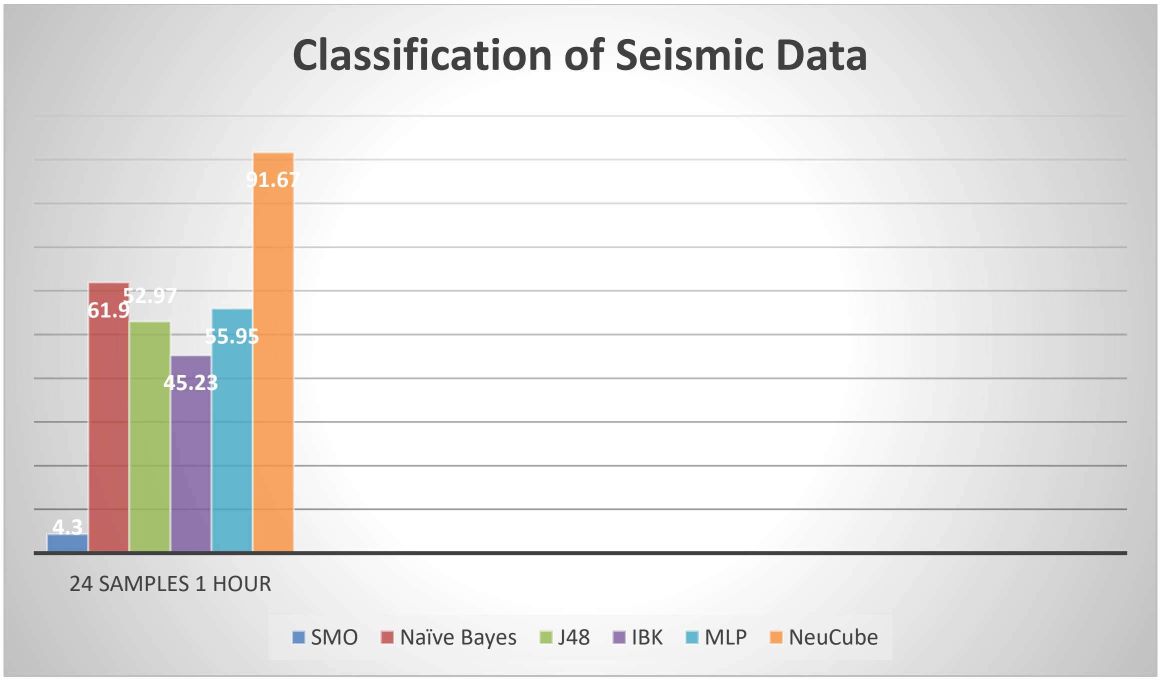

*Figure 33:-**The comparative results of the 24 samples (one hour before the earthquake occurs) with 1 Hz frequency and 100 data points between WEKA and NeuCube.***

The seismic data set used for this experiment consists of 24 samples: 12 low-magnitude earthquakes and 12 high-magnitude earthquakes with 1Hz frequency and 100 data time points. The results for the classification of the seismic data (six hours before the earthquake occurs) are shown in the below table (Table 5).

Method	Performance				
	Other parameters	Accuracy (%)	F-Score	TP	FP
SMO(SVM)(WEKA)	SVM Kernel: Polynomial Degree	49.995±16.31	0.500	6	6
Naïve Bayes (WEKA)	Use Kernel Estimator: False	58.92±9.71	0.580	6	6
J48(Decision Trees) (WEKA)	Pruning Confidence:0.25, Number of Folds:3	52.97±18.59	0.748	10	2
IBk(k-Nearest Neighbours) (WEKA)	Nearest Neighbour Search Algorithm: Linear NN Search	46.42±15.90	0.497	5	7
MLP (WEKA)	Number of Hidden Units:3, Number of Training Cycles:500	51.78±15.65	0.541	6	6
NeuCube	Mod:0.93,Refractory Time:6,STDP:0.01, Drift:0.005,Potential Leak Rate:0.002, Firing Threshold:0.3	83	0.83	10	3

Table 5: ***The comparative results of the 24 samples (six hours before the earthquake occurs) with 1 Hz frequency and 100 data points between WEKA and NeuCube including the parameters used.***

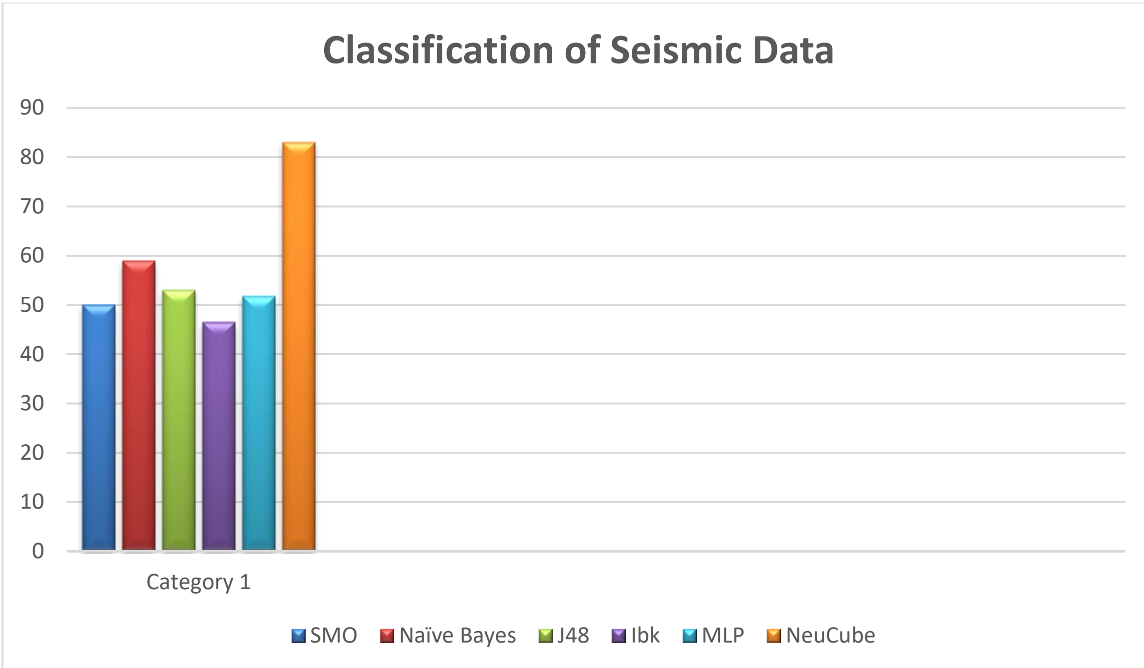

*Figure 34:- **The comparative results of the 24 samples (six hours before the earthquake occurs) with 1 Hz frequency and 100 data points between WEKA and NeuCube.***

- Regression

The seismic data set used for this experiment consists of 24 samples: 12 low-magnitude earthquakes and 12 high-magnitude earthquakes with 1Hz frequency and 100 data time points. The results for the classification of the seismic data (one hour before the earthquake occurs) are shown in the table below (Table 6).

Method	Performance	
	Other Parameters	
EFuNN (NeuCom)	Sensitivity threshold:0.9, Number of Membership Functions:3	1.60
DENFIS(NeuCom)	Distance Threshold:0.5, Number of Membership Functions:12 Number of Epochs:4	0.85
NeuCube	Mod:0.93,Refractory Time:6,STDP:0.01, Drift:0.0005,Potential Leak Rate:0.002, Firing Threshold:0.4	1.51

*Table 6:-**The comparative results of the 24 samples (one hour before the earthquake occurs) with 1 Hz frequency and 100 data points between NeuCom and NeuCube including the parameters used.***

The seismic data set used for this experiment consists of 24 samples: 12 low-magnitude earthquakes and 12 high-magnitude earthquakes with 1Hz frequency and 100 data time points. The results for the classification of the seismic data (six hours before the earthquake occurs) are shown in the table below (Table 7).

Method	Performance	
	Other Parameters	
EFuNN (NeuCom)	Sensitivity threshold:0.9, Number of Membership Functions:3	1.58
DENFIS(NeuCom)	Distance Threshold:0.1, Number of Membership Functions:2 Number of Epochs:4	0.76
NeuCube	Mod:0.93,Refractory Time:6,STDP:0.01, Drift:0.0005,Potential Leak Rate:0.002, Firing Threshold:0.4	1.44

Table :7 ***The comparative results of the 24 samples (six hours before the earthquake occurs) with 1 Hz frequency and 100 data points between NeuCom and NeuCube including the parameters used.***

The seismic data set used for this experiment consists of 25 samples: 12 low-magnitude earthquakes and 13 high-magnitude earthquakes with 1Hz frequency and 100 data time points. The results for the classification of the seismic data (one hour before the earthquake occurs) are shown in the table below (Table 8).

Method	Performance	
	Other Parameters	
EFuNN (NeuCom)	Sensitivity threshold:0.9, Number of Membership Functions:3	1.61
DENFIS(NeuCom)	Distance Threshold:0.9, Number of Membership Functions:2 Number of Epochs:2	0.83
NeuCube	Mod:0.93,Refractory Time:6,STDP:0.01, Drift:0.0005,Potential Leak Rate:0.002, Firing Threshold:0.5	0.44

*Table 8:-**The comparative results of the 25 samples (one hour before the earthquake occurs) with 1 Hz frequency and 100 data points between NeuCom and NeuCube including the parameters used.***

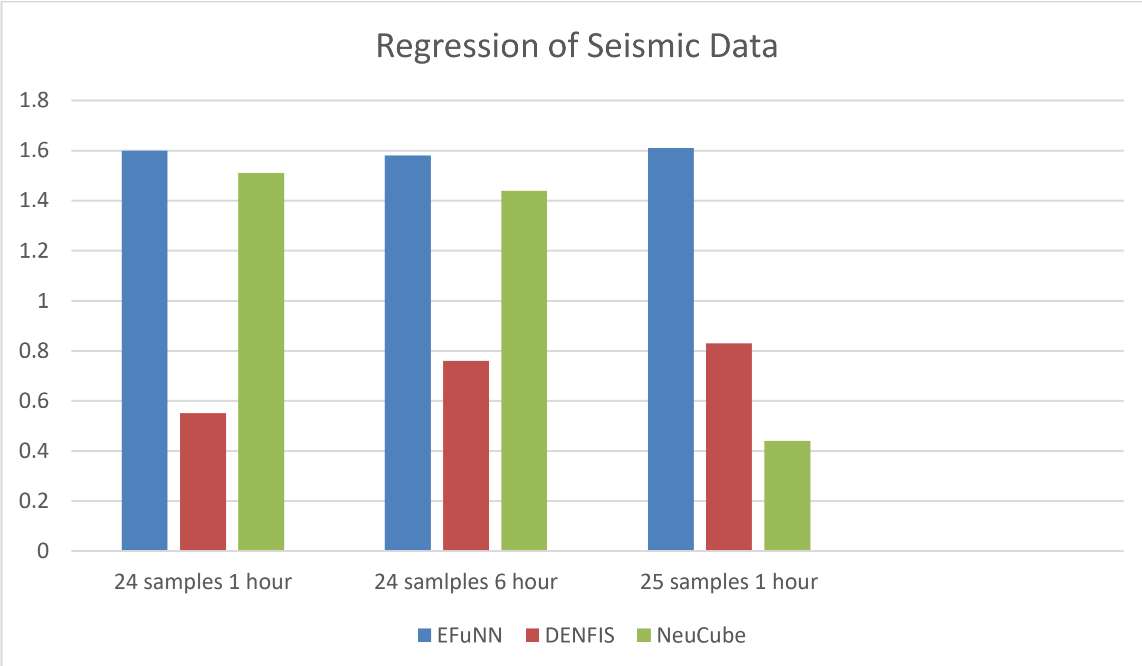

*Figure 35:-**The comparative results of the 25 samples (one hour before the earthquake occurs) with 1 Hz frequency, 24 samples (one hour before the earthquake occurs), 24 samples (six hours before the earthquake occurs) and 100 data points between NeuCom and NeuCube.***

- Visualisation

It is important to have a visualisation system for the exploration of fast data and also for analysing, and offering critical visual analytical information so researchers or scientists can make fast judgements. Raw time-series data requires pre-processing before it is rendered. The changing features hidden in the data can be spatially uncovered by the visualisation of the time-varying component. The ability for interactive detection and exploration of hidden phenomena in data is important. To understand the complex behaviour and also to detect the essential features in the data, visualisation is required. It is necessary for the visualisation software to offer functionality for changing viewpoints, and for modifying colour mappings and classification functions at an interactive rate. Large-field seismic data sets pose various challenges for analysing as well as processing. With the difficulty of 2D technique visualisation, the 3D visualisation delivers important clues through the natural representation of the time-varying information of the time-series data. Visualisation of the various areas of the seismic data divulges the sequence of seismic wave propagation, energy releasing events, attenuation over time, and the initiation.

Insights developed through the visualisation of the seismic data will benefit researchers and seismologists wanting to understand the characteristics of earthquakes. **(Hsieh, Chen, & Ma, 2010)**

The visualisation methods also help analysts to choose complex patterns, generate hypotheses and recommend explanations for more analysis and, lastly, present patterns in an easy manner. The visualisation approaches consist of both generally used information graphics such as maps, tables, histograms, charts and scatter plots and also sophisticated multidimensional visualisation techniques. **(Guo, Gahegan, MacEachren, & Zhou, 2005)**

The exploitation and identification of the profiles of recurring trends and relationships are used to offer knowledge to help in the performance of concurrent multiple time-series data prediction, thus enabling improved accuracy in the prediction of time-series data. In actuality, localised disturbances might be of great significance when they capture the conditions in circumstances when the time-series data is deviating from the norm. The interesting part will be to capture similar deviations from the global trajectory over time in a repeated manner, and for capturing periodic deviations from the norm which are similar in magnitude and shape. These phenomena can only be captured by models built on the data and not tainted by data outside of the phenomena. The repository that consists of the profiles and recurring trends of the phenomena is considered as a knowledge base and is key to performing multiple time-series prediction. **(Widiputra, Pears, & Kasabov, 2011)**

Forecasting and modelling of seismic data (earthquake) is an extremely complex and challenging task as its complexity originates from several sources. Firstly, the temporal scales fluctuate from hundredths of a second to a couple of minutes to help resolve the highest frequencies in relation to the shaking of the ground. Secondly, earthquakes consist of irregular geometrical structures. Thirdly, the soil material properties contained in these structures are very heterogeneous. Fourthly, multiple spatial scales distinguish the data. Fifthly, large- magnitude earthquakes offer non-linear material behaviour and, lastly, the source and geology parameters are observable in an indirect manner, thus introducing uncertainty into the modelling processes.

The problems or challenges which occur in visualisation are:-

- Time Varying Data
- Multiple Variables
- Large Data
- Vector and Displacement Fields
- Unstructured Mesh **(Ma, Stompel, Bielak, Ghattas, & Kim, 2003)**

Consideration should be given to the relationships between the features while displaying certain values in time-series data. Permitting the user to look directly at the data by eliminating unwanted data values will enhance the field view of seismic data, thus providing a synaptic view to help in the global interpretation. This can be done by choosing a specific colour for large-magnitude earthquakes and another specific colour for low-magnitude earthquakes. **(Wolfe & Liu, 1988)**

3D Visualisation of seismic data is perplexing due to the high presence of noise, the dense nature of the data, and the difficulty of manipulating, selecting and identifying the structures. Hence, 3D seismic data interpretation is split into lengthy 2D interpretations of piles of selected cross-sections throughout the data. **(Patel, Bruckner, Viola, & Groller, 2010)**

The various software that has been used for visualisation of the seismic data in the previous studies are as follows:-

- EarthCube
- Voxel Geo 2.0
- GeoViz
- GENESIS
- NEURON
- SpikeNet

NeuVis:-

NeuVis software is a 3D visualisation module for the NeuCube. We feed the trained cube into the NeuVis software. The number of connections and neurons within the 3D structure of the NeuCube requires visualisation beyond the 2D weight/connectivity matrix or the orthographic view of the volume. The NeuVis that is created is a special renderer for the NeuCube data sets which use the GLSL and JOGL (Java Bindings for OpenGL) shaders to render up to 1.5 million neurons and their connections with a steady framerate of about 60 fps. This network can also be simulated in time for showing the spiking neuron activity which also allows the pausing or the slowing down of the simulation. **(Marks, Estevez, & Connor, 2014)**

The screenshot of the NeuVis visualisation software is shown below (Figure 36).

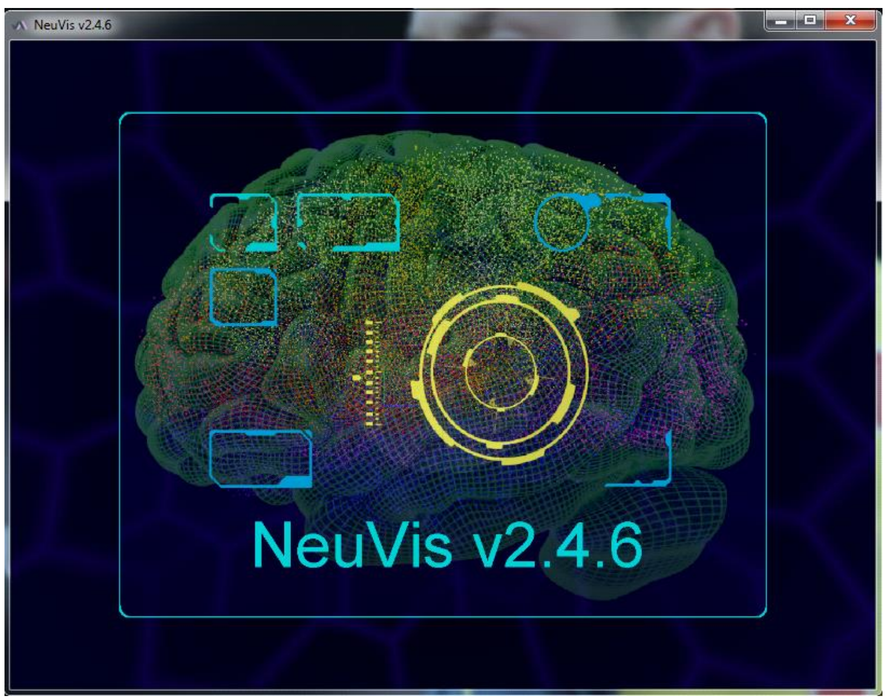

Figure 36:- ***NeuVis-3D Visualisation software***

The interactive mechanisms allow for the playback of the development of the connection weights and spiking patterns throughout its learning period. The visualisation also includes the connection length analysis and the analysing functionality for the connections to show its path. A 3D mouse cursor is available for looking at particular neurons, spiking history and parameters. **(N. Kasabov et al., 2016)**

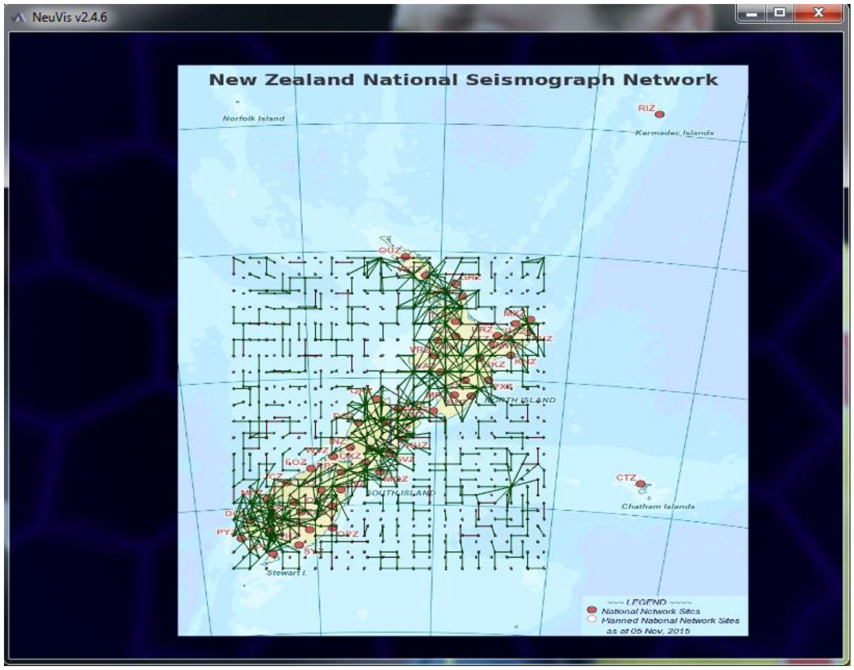

Figure 37:- **The visualisation of the connection between neurons that are based on the seismic data trained in the NeuCube is shown in the NeuVis software.**

The screenshot shown above (Figure 37) illustrates the connection between the neurons based on the seismic data that were trained in the NeuCube. Below the connections we can see the New Zealand National Seismograph Network Map and with time we can see the connections forming.

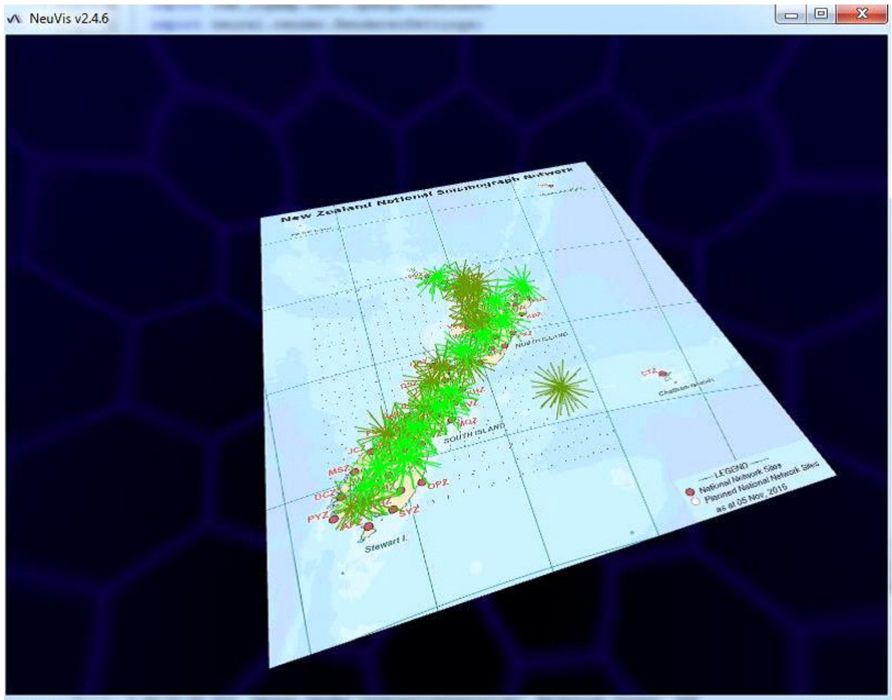

*Figure 38:- **The visualisation of spikes based on the seismic data trained in the NeuCube is shown in the NeuVIs software .The light green are the spikes with strong connections and the dark green are the spikes with weak connections.***

The screenshot shown above (Figure 38) illustrates the spikes based on the seismic data that were trained in the NeuCube. Below the connections we can the New Zealand National Seismograph Network Map. The bright green spikes are the strong connections and the dark green spikes are the weak connections. The group of bright green spikes illustrates that there is high seismic activity. Given the threshold parameter, we can get information on the spike activity in relation to the time.When we look at this visualisation for the real time data, we can visualize the high or the low activities occuring at differeent regions of New Zealand and also when the connections are forming we can get an idea or the starting point for the occurence of earthquakes during the time.

- ## Analysis and Discussion

In this study, the analysis of classification is based on the following Evaluation Criteria:-

- Classification Performance:-

- True Positives :-

- False Positives:-

In this study, for regression we use the following Evaluation Criteria:-

- RMSE error.

According to the preliminary experiments conducted, the results of classification of the 25 sample seismic data set (one hour before the earthquake occurs), show that the NeuCube is able to classify better than traditional machine learning algorithms and the evolving connectionist systems, with an accuracy performance of 80%.

Regarding the conduction of the preliminary experiments, the results of classification of the 24 sample seismic data set (one hour before the earthquake occurs), show that the NeuCube is able to classify better than traditional machine learning algorithms and the evolving connectionist systems, with an accuracy performance of 91.67%.

In relation to the preliminary experiments, the results of classification of the 24 sample seismic data set (six hours before the earthquake occurs), show that the NeuCube is able to classify better than traditional machine learning algorithms and the evolving connectionist systems, with an accuracy performance of 83%.

Referring to the preliminary experiments of regression, the results show that for the 24 sample seismic data set (six hours and one hour before the earthquake occurs), the DENFIS (NeuCom) algorithm predicts better when compared with EFuNN (NeuCom) and the NeuCube.

In connection to the 25 sample seismic data set (one hour before the earthquake occurs), the NeuCube predicts better than DENFIS (NeuCom) and EFuNN (NeuCom)

Effective performance of the evolving spatio-temporal data model depends upon the correct combination of numerous parameters. There is a concern regarding computational scaling with systems such as NeuCube. This is mainly due to the consumption of power, the physical size of the system in certain applications, and the computational power. In relation to the spiking neural networks, a robust information theory supports the implementation and design of the networks. In relation to the evolving spatio-temporal model (NeuCube), the network structure is based on some a-priori knowledge of the data set. Here, the advantage is the representation of the spectral and/or spatial components of the data set sources with time and also in retaining the relationship between the temporal aspect and the spatial and/or spectral aspect of the data. **(N. Kasabov et al., 2016)**

The NeuCube method has the ability to scale larger data sets. But, the computational complexity is based on the storage space, computing capacity and the storage space. The computational complexity can either be linear or exponential.

There are challenging questions such as, how much noise is acceptable? How can we understand and model the transitions based on the external stimuli that triggers the spatio-temporal states? How much capacity has the NeuCube for learning both the temporal and spatial characteristics of the data set? **(N. K. Kasabov, 2014)**

The earthquake is a geological event controlled by the geo-structure locally and it also has an internal relationship with the tectonic unit. It is also understood that the distribution of earthquakes is controlled by the local geo-structure and is uneven, hence a regional seismo tectonic frame can also be considered during data processing. **(Sheng, Mu, Zhang, & Lv, 2015)**

The fuzzy inference is defined as the process of formulating mapping of data from the given input to the output using fuzzy logic. This mapping then helps to provide a basis for the respective decisions to be made. **(Sandhu, Salaria, & Singh, 2008)**

The significance of the outcomes of the different classification algorithms helps in understanding and refining the model with the intention of providing better results. The analysis of the values of the true positive rate and false positive rate helps us to gain an insight to tune the classifiers. **(Orozco-Alzate et al., 2012)**

Chapter 7:- Conclusion and Future Work

This chapter reviews the study and discusses the limitations. There is an outline of the contribution and a brief discussion of future work.

•A review of the Study:-

The thesis presents a case study that uses the seismic data set as spatio-temporal data. The seismic data set has been taken from the GeoNet website (www.geonet.org.nz) relates only to the Christchurch area of New Zealand, and is available publicly.

The thesis starts with a literature review and an introduction to WEKA and the machine learning algorithms used in WEKA. An introduction to Evolving Connectionist Systems (ECOS) and NeuCom is also provided along with a description of EFuNN (Evolving Fuzzy Neural Networks) and DENFIS (Dynamic Evolving Neuro-Fuzzy Inference Systems) algorithms used in NeuCom.

A literature review of Spiking Neural Networks is given along with a description of the software NeuCube that was used.

The proposed methodology framework is given to explain the comparative analysis process used in this study.

The preliminary experiment results and analysis are briefly described as well as some brief discussions in this study.

•Limitations:-

The limitations in this study are:-

- The size of the seismic data set used in this study was relatively small, with only 24 and 25 samples of seismic data set.
- The seismic dataset was taken only for the Christchurch area in New Zealand, and for earthquakes that have occurred since 2010.
- The incomplete understanding of numerous features in relation to the prediction of the earthquakes.
- The unavailability of a real-time seismic dataset.

•Future Work:-

The limitations discussed in the above section can be a starting point for future work. The size of the seismic data to be used can be a large sized data set taken from across the regions of New Zealand and also can be applied to the seismic data from other countries such as Chile, Japan etc.. The availability of real-time seismic data and experimenting with it can help in the understanding of the classification and prediction of earthquakes in a better manner.

The 3D visualisation of spiking neural networks done in this study is the starting point for a new visualisation using virtual reality software. It can be taken further by refinement and also by adding new features. Future work can also consist of research relating to time-series data with noise. Is some noise acceptable to provide a better result or will removal of all noise provide a better result?

References

Ali, A. (2014). *An improved multilayer perceptron based on wavelet approach for physical time series prediction*. Universiti Tun Hussein Onn Malaysia.

Anderson, H., & Webb, T. (1994). New Zealand seismicity: patterns revealed by the upgraded National Seismograph Network. *New Zealand journal of geology and geophysics, 37*(4), 477-493.

Avvenuti, M., Cresci, S., La Polla, M. N., Marchetti, A., & Tesconi, M. (2014). Earthquake emergency management by social sensing*IEEE.* Symposium conducted at the meeting of the Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on

Azam, F., Sharif, M., Yasmin, M., & Mohsin, S. (2014). Artificial intelligence based techniques for earthquake prediction: a review. *Sci Int, 26*(4), 1495-1502.

Baeten, G. J. M., Ferber, R.-G., & Lengeling, R. (2002). Seismic data acquisition and method for spatially filtering seismic data: Google Patents.

Bassis, S., Esposito, A., & Morabito, F. C. (2015). *Advances in Neural Networks: Computational and Theoretical Issues* (Vol. 37): Springer.

BBC GCSE BiteSize. (2014). Diagram of an earthquake.

Bishop, C. M. (1995). *Neural networks for pattern recognition*: Oxford university press.

Cole, M. A., Elliott, R. J., Okubo, T., & Strobl, E. (2013). *Natural disasters and plant survival: The impact of the Kobe earthquake*.

Dash, R. K. (2013). Selection of the best classifier from different datasets using WEKA*ESRSA Publications.* Symposium conducted at the meeting of the International Journal of Engineering Research and Technology

DeMets, C., Gordon, R. G., Argus, D., & Stein, S. (1990). Current plate motions. *Geophysical journal international, 101*(2), 425-478.

Diao, Y., & Passino, K. M. (2002). Adaptive neural/fuzzy control for interpolated nonlinear systems. *IEEE Transactions on Fuzzy Systems, 10*(5), 583-595.

Djarfour, N., Ferahtia, J., Babaia, F., Baddari, K., Said, E.-a., & Farfour, M. (2014). Seismic noise filtering based on Generalized Regression Neural Networks. *Computers & Geosciences, 69*, 1-9.

Earth Science beta. (2014). What tectonic mechanisms cause the North and South Islands of New Zealand to be so geologically different?

Fountalis, I., Bracco, A., Dilkina, B., Dovrolis, C., & Keilholz, S. (2016). {\ delta}-MAPS: From spatio-temporal data to a weighted and lagged network between functional domains. *arXiv preprint arXiv:1602.07249*.

Fullér, R. (1995). Neural fuzzy systems.

GeoNet. (n.d.-a). Earthquake Facts and Statistics.

GeoNet. (n.d.-b). New Zealand National Seismograph Network.

GeoNet. (n.d.-c). *What causes earthquakes in New Zealand?* Retrieved 26,June, 2016, from http://info.geonet.org.nz/display/quake/Earthquake+FAQ

GNS Science. (n.d.-a). *Earthquakes at a Plate Boundary*. Retrieved 25,June, 2016, from http://www.gns.cri.nz/Home/Learning/Science-Topics/Earthquakes/Earthquakes-at-a-Plate-Boundary

GNS Science. (n.d.-b). Where do earthquakes happen in New Zealand?

Guo, D., Gahegan, M., MacEachren, A. M., & Zhou, B. (2005). Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach. *Cartography and Geographic Information Science, 32*(2), 113-132.
Hassoun, M. H. (1995). *Fundamentals of artificial neural networks*: MIT press.
HAYKIN, S. (1994). Neural networks, A comprehensive Foundation.
Hsieh, T.-J., Chen, C.-K., & Ma, K.-L. (2010). Visualizing field-measured seismic data*IEEE.* Symposium conducted at the meeting of the Visualization Symposium (PacificVis), 2010 IEEE Pacific
IRIS, D. A. Data Services Products: EMC-References IRIS EMC References.
Ishibuchi, H., Nozaki, K., & Tanaka, H. (1992). Distributed representation of fuzzy rules and its application to pattern classification. *Fuzzy sets and systems, 52*(1), 21-32.
Kasabov, N. (2007). Evolving Connectionist Systems: The Knowledge Engineering Approach (Evolving Connectionist Systems): Springer London.
Kasabov, N. (2007). Global, local and personalised modeling and pattern discovery in bioinformatics: An integrated approach. *Pattern Recognition Letters, 28*(6), 673-685.
Kasabov, N. (2012). Neucube evospike architecture for spatio-temporal modelling and pattern recognition of brain signals. In *Artificial Neural Networks in Pattern Recognition* (pp. 225-243): Springer.
Kasabov, N. (2014). Brain-like Information Processing for Spatio-Temporal Pattern Recognition. In *Springer Handbook of Bio-/Neuroinformatics* (pp. 813-834): Springer.
Kasabov, N., Dhoble, K., Nuntalid, N., & Indiveri, G. (2013). Dynamic evolving spiking neural networks for on-line spatio-and spectro-temporal pattern recognition. *Neural Networks, 41*, 188-201.
Kasabov, N., Scott, N. M., Tu, E., Marks, S., Sengupta, N., Capecci, E., . . . Hartono, R. (2016). Evolving spatio-temporal data machines based on the NeuCube neuromorphic framework: design methodology and selected applications. *Neural Networks, 78*, 1-14.
Kasabov, N. K. (1998). ECOS: Evolving Connectionist Systems and the ECO Learning Paradigm*Citeseer.* Symposium conducted at the meeting of the Iconip
Kasabov, N. K. (2013). *Springer Handbook of Bio-/neuro-informatics*: Springer Science & Business Media.
Kasabov, N. K. (2014). NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. *Neural Networks, 52*, 62-76.
Kasabov, N. K., & Song, Q. (2002). DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *Fuzzy Systems, IEEE Transactions on, 10*(2), 144-154.
KEDRI. (2012). *NeuCom*. Retrieved 27,June, 2016, from https://kedri.aut.ac.nz/areas-of-expertise/data-mining-and-decision-support-systems/neucom
Keller, J. M., & Tahani, H. (1992). Backpropagation neural networks for fuzzy logic. *Information Sciences, 62*(3), 205-221.
Kumar, A., Walia, V., Singh, S., Bajwa, B. S., Dhar, S., & Yang, T. F. (2013). Earthquake precursory studies at Amritsar Punjab, India using radon measurement techniques. *International Journal of Physical Sciences, 7*(42), 5669-5677.
Lara-Cueva, R., Benítez, D., Carrera, E., Ruiz, M., & Rojo-Álvarez, J. (2016). Feature selection of seismic waveforms for long period event detection at Cotopaxi Volcano. *Journal of Volcanology and Geothermal Research, 316*, 34-49.
Liang, W., Krishnamurthi, R., Kasabov, N., & Feigin, V. (2014). Information methods for predicting risk and outcome of stroke. In *Springer Handbook of Bio-/Neuroinformatics* (pp. 993-1001): Springer.
Lighthill, J. (1996). *A critical review of VAN: earthquake prediction from seismic electrical signals*: World Scientific.
Ma, K.-L., Stompel, A., Bielak, J., Ghattas, O., & Kim, E. J. (2003). Visualizing very large-scale earthquake simulations*IEEE.* Symposium conducted at the meeting of the Supercomputing, 2003 ACM/IEEE Conference

Maass, W. (1997). Networks of spiking neurons: the third generation of neural network models. *Neural Networks, 10*(9), 1659-1671.

Marks, S., Estevez, J. E., & Connor, A. M. (2014). Towards the Holodeck: fully immersive virtual reality visualisation of scientific and engineering data*ACM.* Symposium conducted at the meeting of the Proceedings of the 29th International Conference on Image and Vision Computing New Zealand

Orozco-Alzate, M., Acosta-Muñoz, C., & Londoño-Bonilla, J. M. (2012). *The automated identification of volcanic earthquakes: concepts, applications and challenges*: INTECH Open Access Publisher.

Panakkat, A., & Adeli, H. (2009). Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators. *Computer-Aided Civil and Infrastructure Engineering, 24*(4), 280-292.

Patel, D., Bruckner, S., Viola, I., & Groller, E. M. (2010). Seismic volume visualization for horizon extraction*IEEE.* Symposium conducted at the meeting of the Visualization Symposium (PacificVis), 2010 IEEE Pacific

Patil, T. R., & Sherekar, S. (2013). Performance analysis of Naive Bayes and J48 classification algorithm for data classification. *International Journal of Computer Science and Applications, 6*(2), 256-261.

Platt, J. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines.

Resilience, S. (n.d.). Seismicity in New Zealand.

Reynolds, K., Kontostathis, A., & Edwards, L. (2011). Using machine learning to detect cyberbullying*IEEE.* Symposium conducted at the meeting of the Machine Learning and Applications and Workshops (ICMLA), 2011 10th International Conference on

Romeo, G. (1994). Seismic signals detection and classification using artiricial neural networks. *Annals of Geophysics, 37*(3).

Sandhu, P. S., Salaria, D. S., & Singh, H. (2008). A comparative analysis of fuzzy, neuro-fuzzy and fuzzy-GA based approaches for software reusability evaluation. *Proc of Worlds Academy of Science, Engineering and Technology, 29*.

Sharma, R., Alam, M. A., & Rani, A. (2012). K-means clustering in spatial data mining using weka interface*Citeseer.* Symposium conducted at the meeting of the International conference on advances in communication and computing technologies (ICACACT)

Sheng, J., Mu, D., Zhang, H., & Lv, H. (2015). Seismotectonics Considered Artificial Neural Network Earthquake Predic-tion in Northeast Seismic Region of China. *Open Civil Engineering Journal, 9*, 522-528.

Tu, E., Kasabov, N., Othman, M., Li, Y., Worner, S., Yang, J., & Jia, Z. (2014). Neucube (st) for spatio-temporal data predictive modelling with a case study on ecological data*IEEE.* Symposium conducted at the meeting of the Neural Networks (IJCNN), 2014 International Joint Conference on

Vazquez, R. A., & Garro, B. A. (2015). Training spiking neural models using artificial bee colony. *Computational intelligence and neuroscience, 2015*, 18.

Vijayarani, S., & Muthulakshmi, M. (2013). Comparative analysis of bayes and lazy classification algorithms. *International Journal of Advanced Research in Computer and Communication Engineering, 2*(8), 3118-3124.

Webb, A. R. (2003). *Statistical pattern recognition*: John Wiley & Sons.

Widiputra, H., Pears, R., & Kasabov, N. (2011). Multiple time-series prediction through multiple time-series relationships profiling and clustered recurring trends. In *Advances in Knowledge Discovery and Data Mining* (pp. 161-172): Springer.

Wolfe, R., & Liu, C. (1988). Interactive visualization of 3D seismic data: A volumetric method. *Computer Graphics and Applications, IEEE, 8*(4), 24-30.

Zamani, A., Sorbi, M. R., & Safavi, A. A. (2013). Application of neural network and ANFIS model for earthquake occurrence in Iran. *Earth Science Informatics, 6*(2), 71-85.