Earthquake Prediction: A MLP & SVM Comparison

Daniel van den Akker*
Responsible Professor: Elvin Isuf†
Supervisors: Mohammad Sabbaqi, Maosheng Yang‡
EEMCS, Delft University of Technology, The Netherlands

January 23, 2022

Abstract

Multi-Layer Perceptron and Support Vector Machine have both been widely used in machine learning. In this research paper, these models have been applied to binary classification on an individual time series basis. The goal was to see whether they can predict earthquakes, using earthquakes measured at specific stations across New Zealand. As it turns out, both models serve as satisfactory classifiers. However, their performances are dependent on the stations the data was accumulated from.

1 Introduction

An earthquake is a natural disaster that is caused by a shift or displacement of tectonic plates. This sudden shift results in the release of seismic energy which subsequently shakes the surface of the earth. The violent vibrations of the surface of the earth lead to damage and loss of human lives. Being able to predict an earthquake could help towards hazard prevention. However, this tends to be difficult due to the inaccessible location of the hypocenter of the earthquake, which is often tens of kilometers underground, and the abruptness at which an earthquake can occur. A prediction should give information on the magnitude, location or time of an earthquake. Of the three, the time estimation of a model falls under short-term prediction (less than 1 year), intermediate term prediction (1 to 10 years) or long term prediction (10 to 100 years) [1].

In the past, various methods of earthquake prediction have been proposed. The first one is trend-based prediction, which uses statistical modeling to make estimations. This method does not perform well on an individual case basis and only predicts long-term trends. The second method is precursor-based [2]. This method looks at geophysical properties like the climate, radon gas emissions and even the behaviour of animals. However, this method depends on the occurrence of specific precursors even though in practice such precursors usually occur without any subsequent seismic events, or are hard to detect [3]. Therefore, this technique does not lead to satisfactory results and will not be used in this research paper.

*d.a.vandenakker@student.tudelft.nl
†e.isufi@tudelft.nl
‡m.sabbaqi@tudelft.nl, m.yang-2@tudelft.nl

Instead, this paper will apply a third method for earthquake prediction: machine learning. In recent years, machine learning has proven to be a useful tool for dealing with data-intensive, non-linear and complex problems [4]. Machine learning can act as an unbiased tool to find patterns over insurmountable amounts of data.

1.1 Literature Review

In the past, there has been research into earthquake prediction using neural networks. A widely accepted method for precursor-based prediction uses parameters such as: longitude, latitude, magnitude and depth. For instance [5], constructs a Multi-Layer Perceptron (MLP) for earthquake prediction using a precursor method, which takes as parameters the time between earthquake events, can predict the magnitude, location and date of an earthquake. Another trend-based research by [6] does a comparison between two models, Feed-Forward Neural Networks (FFNN) Long Short-Term Memory (LSTM). Their research has shown that the R2 score yielded by the LSTM is 59% higher than for the FFNN. Moreover, the research paper [7], constructs an LSTM which discovers the spatio-temporal correlations among earthquake occurrences and takes advantage of these correlations to make accurate earthquake predictions.

A different approach was taken by [8], which uses Fingerprint And Similarity Thresholding (FAST), which creates a digital fingerprint of a seismic waveform of an earthquake and extracts key features. These distinct features allow for efficient comparison of seismic waveforms over large amounts of data.

Other research papers also use seismic waveforms as a precursor parameter. For example, [9], uses a Convolutional Neural Network (CNN) for earthquake detection and prediction using a single seismic waveform. This research has shown promising results and has outperformed other algorithms such as the previously mentioned FAST approach. A significant difference with this paper, is that the data collected was a single waveform from a single location, whereas this paper collected data of multiple waveforms from multiple locations.

1.2 Research question

The research question is: "How does a MLP compare with a SVM when operating on individual time series in predicting earthquakes?". The input will be seismic waveforms from different locations. The paper will compare two model (SVM and MLP) using known evaluation metrics.

This paper will also provide a baseline for other research papers which will research do into other machine learning models or consider waveforms from channels/locations.

2 Methodology

As reviewed in section 1.1, there are two main methods for earthquake prediction: trend-based and precursor-based. In this research paper, we take a precursor-based approach. The precursor being 30 seconds of seismic waveform. In order to carry out the experiments three concepts are introduced. In section 2.1, we discuss how the data has been collected and what preprocess steps were taken. Next, section 2.2, we take a look into the machine learning models MLP and SVM. Lastly, in section 2.3, we describe what metrics are utilized to evaluate each model.

2.1 Data Preparation

The setting of the dataset is New Zealand, specifically defined by the geographic location of the bounding box coordinates: latitude -47.749 to -33.779 and longitude 166.104 to 178.990. The time range of the dataset is between January 2016 and July 2021. The dataset has been procured using a web service of the International Federation of Digital Seismograph Networks (FDSN) [10]. The aggregated dataset consists of two types of data: stations and earthquake events. Stations are facilities where the vertical velocity of seismic activity is recorded using weak-motion sensors. The index and metadata of 58 stations can be found in Table 2 in the appendix. The number of earthquake events that are within the time frame and the bounding box is just over 122,000. The locations of these stations and these earthquake events are visualized in Figure 1.

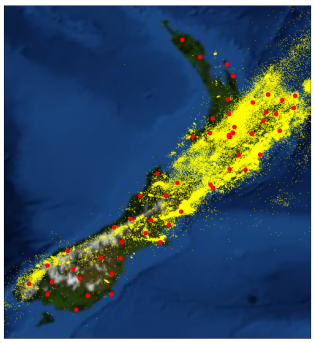


Figure 1: Map of New Zealand. Each red dot represents a station and each yellow dot represents an earthquake.

We are interested in two types or classes of seismic waveforms: non-earthquake also called normal waveforms, and earthquake waveforms. The predicate of the research question is that the seismic activity preceding a seismic event can act as a precursor. For the purposes of this research paper a time length of 30 seconds was chosen, as in [11]. This time period was deemed appropriate since a longer period would result in more irrelevant data and a shorter time period would risk having too little data to make a meaningful analysis. Figure 2 displays an unprocessed waveform consisting of three parts: the waveform during the earthquake, 30 seconds before the earthquake and the normal waveform. The sample rate of the raw seismic waveform is 100HZ. In order to reduce the training time of the models and reduce the susceptibility of the seismic data to noise, the waveforms were downsampled to 50HZ. The amplitude of each earthquake is normalized to values between 1 and -1. Figure 3 demonstrates a normalized waveform being downsampled.

2.1.1 Earthquake Seismic Activity

In this section, the operation performed on the earthquake events and the associated waveforms are discussed. Two properties of an earthquake are magnitude and depth, the distribution of which can be found in Figure 4. In order to make the earthquake data more identifiable, outliers were removed. Namely, earthquake with a magnitude less than one or more than three, and earthquakes with hypocenters[1] more than 200 kilometers deep. After

[1] `https://en.wikipedia.org/wiki/Hypocenter`

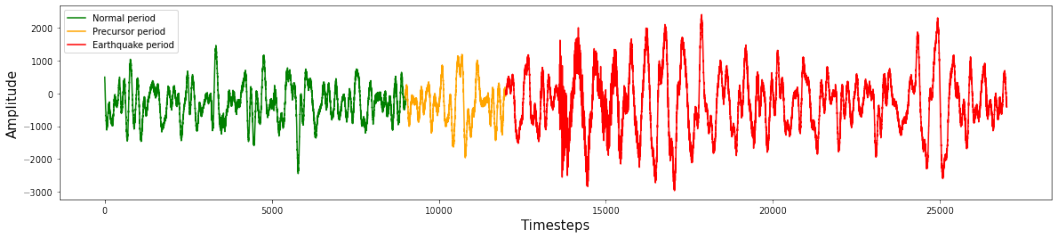


Figure 2: Waveform representation of an earthquake with ID "smi:nz.org.geonet/2019p006920" and station ID "NZ.WEL.10.HHZ"

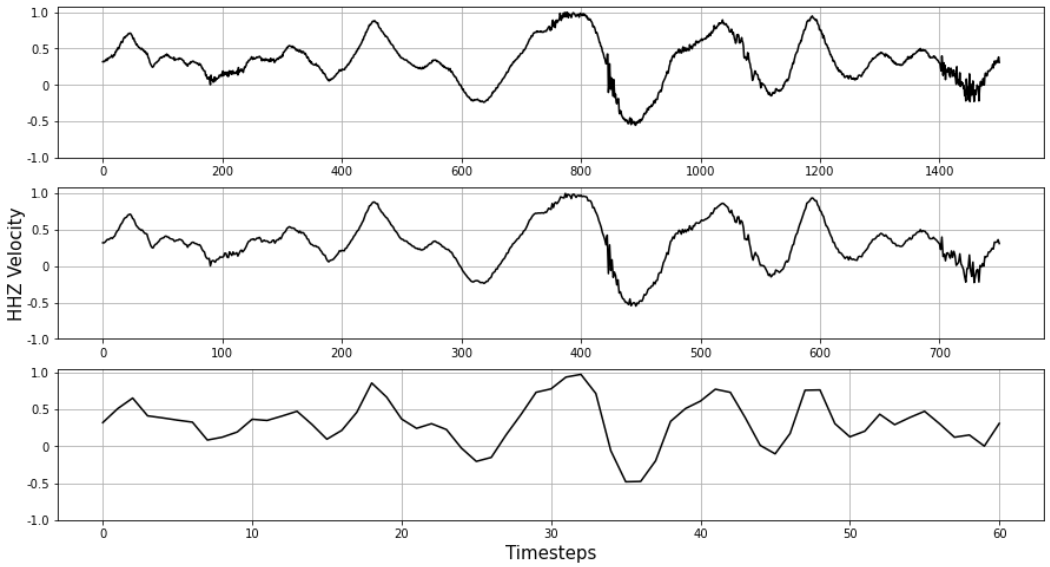


Figure 3: Representation of a normalized seismic waveform at frequencies: 100HZ (top), 50HZ (middle), 2HZ (bottom)

this filtering, 106,000 earthquakes remain. The next step is to divide each earthquake to its nearest station. The distances from the earthquakes to the stations were calculated using the great-circle distance[2]. The distribution of distances of earthquakes to their nearest station can be seen in Figure 5. Earthquakes with a great-circle distance greater than 270 kilometers from the designated station were removed. The upper distance limit guarantees a good signal-to-noise ratio for small magnitude events at high frequencies [12].

The number of features or length of each sample is calculated by multiplying the frequency by the time length (30 seconds). According to [13], to ensure that the dataset is reliable, the sample size should increase with the number of measurements/features as well as the inherent error rate of the classifier. In order to curate to experiments done using the highest frequency, a size of 1000 samples per class was chosen for a total of 2000 samples. The distribution of earthquakes per station is seen in Figure 6. The total number of waveforms considered is

[2] `https://en.wikipedia.org/wiki/Great-circle_distance`

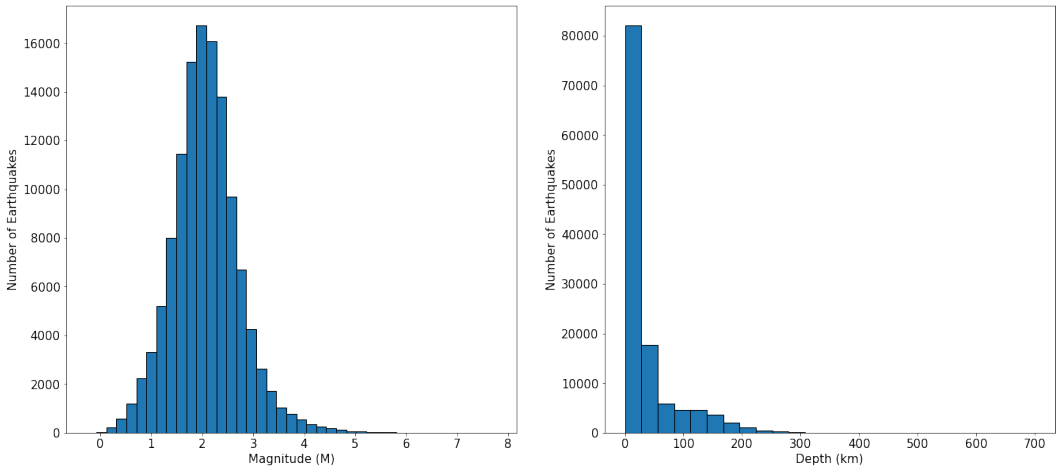


Figure 4: Distribution of magnitude (left) and depth (right) of earthquakes. N = 106623 earthquakes

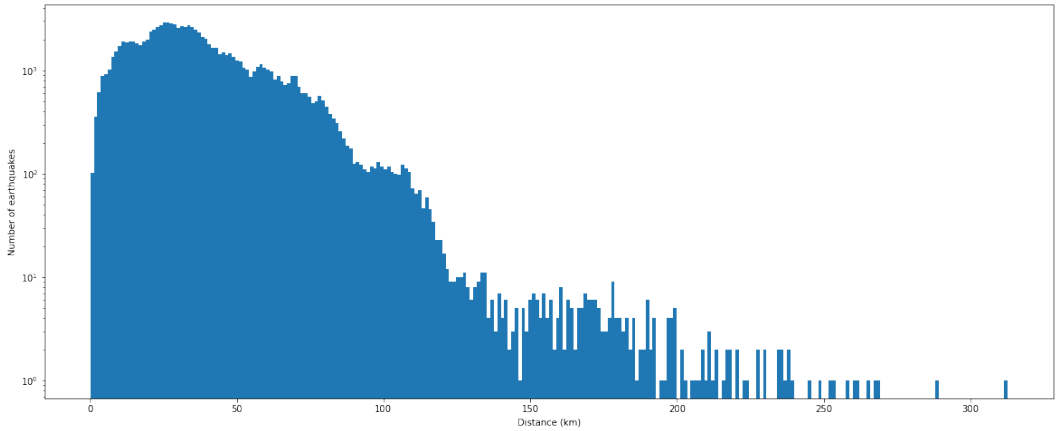


Figure 5: Distance distribution of distances of each earthquake to the nearest station of 58 stations. N = 106623 earthquakes

2.1.2 Non-Earthquake Seismic Activity

The non-earthquake seismic activity also called normal seismic activity are 30 seconds of seismic waveform that do not precede an earthquake. In order to create a balance dataset, the same number of normal seismic activity have been found as the number of earthquake seismic activity per station. A balance dataset means that each class is represented equally in the dataset. An imbalanced dataset often results in skewed models; one class is favoured over another.

The premise of the research question is that seismic activity before an earthquake can act as a classifier for the occurrence of an earthquake. This means that in order to find a classifier for absence of an earthquake we must find time periods that are the furthest away from any earthquake. For this purpose an algorithm was deployed which maximizes the time

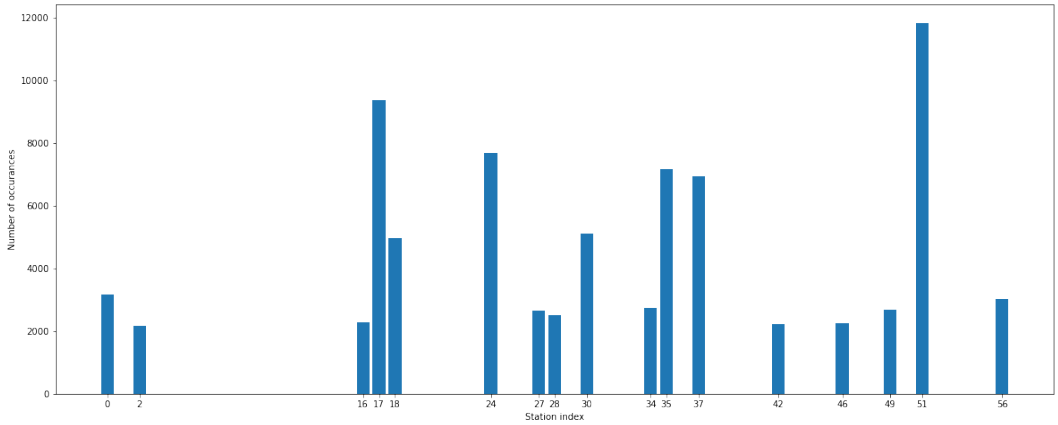


Figure 6: Earthquake distribution to the nearest station of 58 stations. N = 104731 earthquakes

span, or buffer, between the normal seismic activity periods and the times that earthquakes occur. Stations with fewer designated earthquakes, require less samples of normal seismic activity. Therefore, the buffer size is inversely proportional to the sample size.

2.2 Machine Learning Models

This paper evaluates how two deep learning models accurately predict the occurrence of an earthquake in a time-series manner: the Multi-layer Perceptron and the Support Vector Machine. Both models can be used for binary classification tasks, which includes the current earthquake prediction problem. Therefore, each sample in the dataset described in section 2.1, was labeled 1 for earthquake seismic waveform and 0 for normal seismic activity. The differences and implementation of each model are detailed below.

2.2.1 Multi-layer Perceptron

MLP is a type of Feed-Forward Neural Network (FFNN) that consists of at least three types of layers: input, hidden and output. Each layer consists of a set of neurons, which are each connected to all of the neurons in the preceding layer. The depth of a MLP is brought by the inclusion of multiple hidden layers. Each layer, except for the input layer, uses a non-linear activation function. This activation function and the inclusion of multiple hidden layers allows for distinction of data that is not linearly separable. The topology of a simple MLP is shown in Figure 7. Each neuron also has an associated weight. During training, a MLP learns by adjusting the weights through a process called backpropagation.

As mentioned in the previous section, the number of features of a sample is equal to the the frequency of the waveform multiplied by 30 seconds. For example, a waveform of 50HZ has 1500 features. This is equal to the number of neurons in the input layer. The output layer always has only one neuron.

In this research, the MLP was implemented using the PyTorch library. This implementation included several techniques to prevent overfitting (when a model corresponds too closely to a specific dataset, making it not generalizable enough). First of all, the regularization technique was applied. This involves adding a penalty to the loss function in the

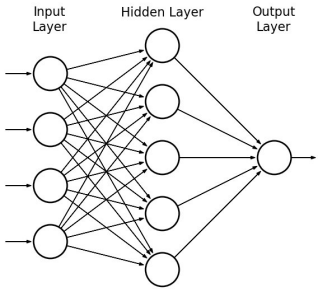


Figure 7: Topology of a simple MLP with one hidden layer. Image adapted from [14].

case of larger weights. There are two types of regularization: L1 and L2. The former uses the sum of weights while the latter uses the sum of the squared weights. Both types have been implemented in this research. Secondly, the dropout technique was used. This disables or "drops out" randomly selected neurons that effect the output of the model negatively. Thirdly, batch normalization was used as a third technique to speed up the training process. This technique standardizes the weights in each layer per batch, making them more stable and easy to process. For optimization both Adam and Stochastic Gradient Descent (SGD) were considered as hyperparameters. Lastly, the implemented model divides the dataset into 70%, 20% and 10%, for training, validation and testing respectively.

2.2.2 Support Vector Machine

The Support Vector Machine contains supervised learning methods in order to classify data, analyze data regression and detect outliers. Moreover, this model creates a hyperplane to separate the classes and often tries to maximize the distance from the hyperplane to the nearest data point of any class. Moreover, there are several kernel functions that map the dataset into a higher dimension. Of the many different functions, three were used in this experiment. The linear, the polynomial and the RBF (Radial Basis Function or Gaussian Radial Basis Function) kernel functions all separate classes in different manners to bring the hyperplane to a higher dimension. See Figures 8, 9, 10 for their mapping equations.

$$K(x_1, x_2) = (x_1 * x_2)$$

Figure 8: Equation for the linear kernel function.

$$K(x_1, x_2) = (x_1 * x_2 + 1)\,\mathrm{d}$$

Figure 9: Equation for the polynomial kernel function.

This model was implemented using the sklearn library. The samples used were divided into groups of 80% and 20% for training and testing respectively.

$$K(x_i, x_j) = exp(-\sigma \left|\left|x_i - x_j\right|\right|^2)$$

Figure 10: Equation for the RBF kernel function

2.3 Evaluation Metrics

To evaluate the models described in 2.2, evaluation matrices accuracy, precision, recall and F1-score have been used in this study. These matrices were calculated using the results of the confusion matrix:

- True Positive (TP): An earthquake occurred and an earthquake was predicted.
- False Positive (FP): No earthquake occurred and an earthquake was predicted.
- True Negative (TN): No earthquake occurred and no earthquake was predicted.
- False Negative (FN): An earthquake occurred and no earthquake was predicted.

Accuracy (1) describes the ratio of correct predictions to all predictions made. In other words, this metric looks at the correctly predicted occurrence and absence of an earthquake.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

Precision (2) measures the ratio of correctly predicted earthquakes to all predicted earthquakes. In contrast, Recall, also known as Sensitivity, (3) is used to calculate the ratio of correctly predicted earthquakes to all occurred earthquakes.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Lastly, F1-score (4) can be considered as the mean between Precision and Recall. This metric gives a more stable outcome, which facilitates a comparison between models.

$$F1\text{-}score = \frac{2TP}{2TP + FP + FN} \tag{4}$$

3 Results

This section will report the results of the experiments using the evaluation metrics described in section 2.3.

The performance of the SVM per station, for the kernels function Linear, Polynomial and RBF, are found in Figures 11, 12, 13, 14. The figures show that RBF has the highest accuracy and F1-score for all stations, and has the highest precision and recall for most stations. Polynomial kernel function has the second highest accuracy for all stations and the highest for precision, recall and F1-score. Linear kernel function has the lowest performance of most stations.

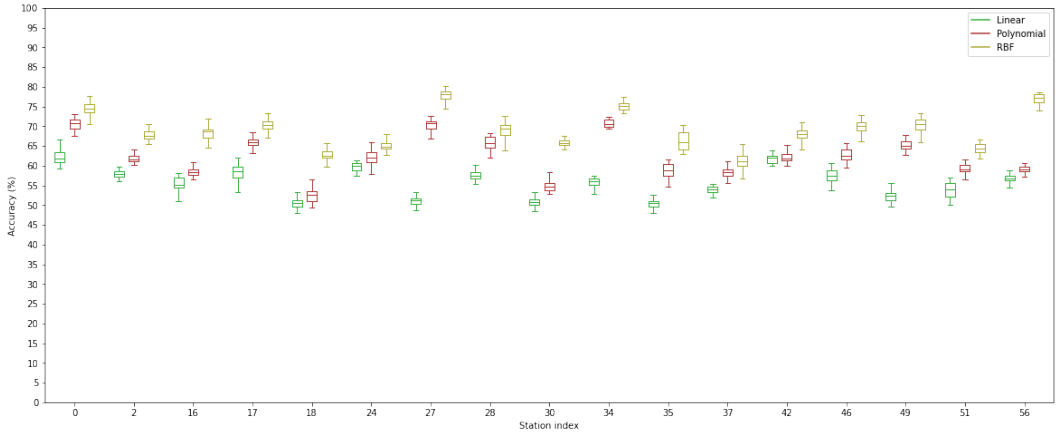


Figure 11: Boxplots of the accuracy of each kernel, for each station. The number of iterations (per boxplot) is 22.

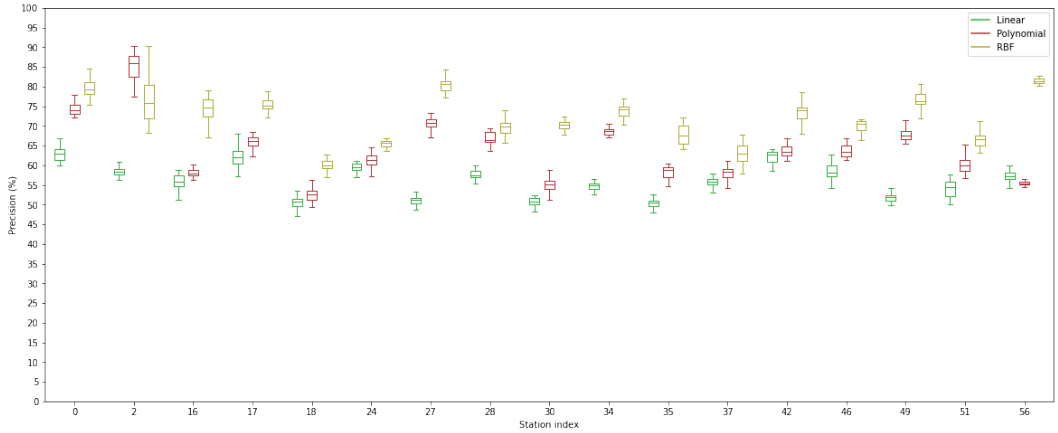


Figure 12: Boxplots of the precision of each kernel, for each station. The number of iterations (per boxplot) is 22.

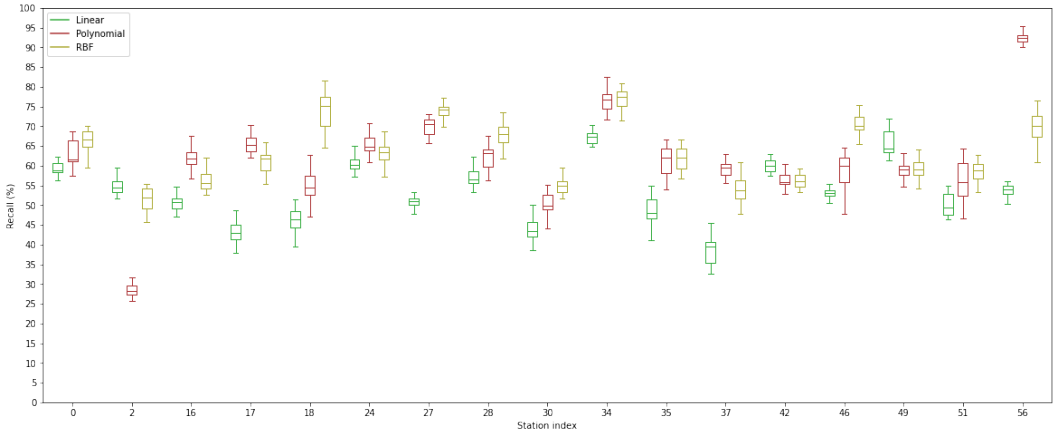


Figure 13: Boxplots of the recall of each kernel, for each station. The number of iterations (per boxplot) is 22.

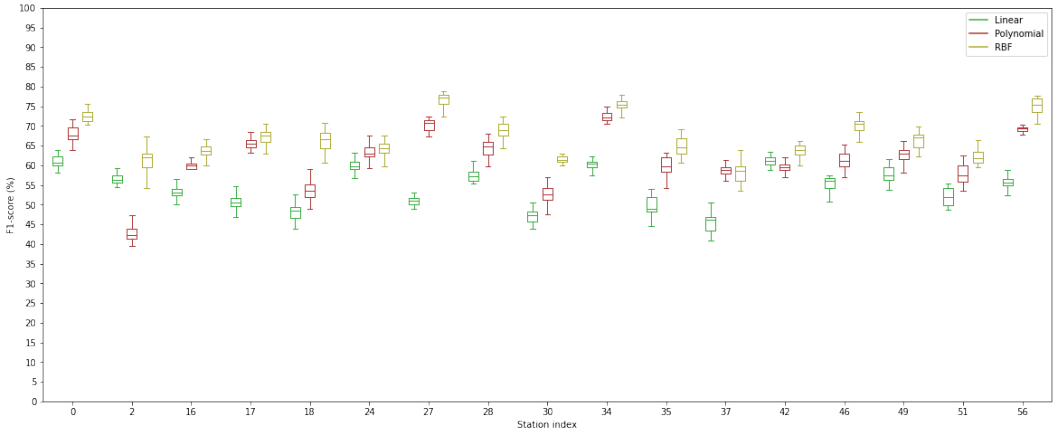
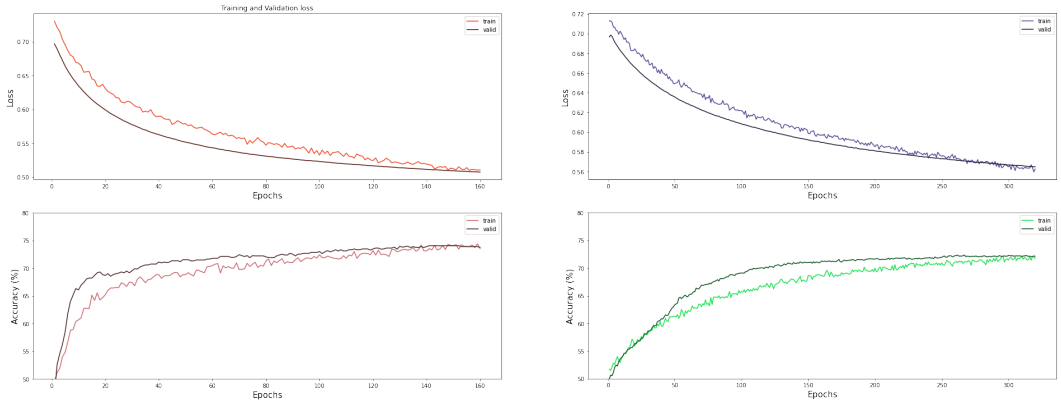


Figure 14: Boxplots of the F1-score of each kernel, for each station. The number of iterations (per boxplot) is 22.

The MLP model has been trained for each station using the hyperparameters in Table 1. The optimizer Adam was used, as there was no significant difference in accuracy between optimizer GCD and Adam. Similarly, L2 regularization was optimal for training every station. The training losses and accuracies of station with indices 0 and 17, can be found in Figure 15. All of the learning curves of all of the station can be found in Figures 20, 21 and 22 in the Appendix. As a result the trained MLP was tested and evaluated using known matrices, see Figures 16, 17, 18 and 19.

Station index	Epochs	Batch size	Learning Rate	Dropout	Number of hidden Layers	Number of neurons per layer
0	160	256	1e-6	0.1	2	[1024, 128]
2	120	256	1e-4	0.8	1	[128]
16	130	128	1e-4	0.9	1	[128]
17	320	1024	1e-6	0.3	4	[1024, 512, 256, 64]
18	400	128	1e-6	0.8	1	[1112]
24	140	512	1e-4	0.6	1	[128]
27	140	512	1e-5	0.9	1	[512]
28	100	512	1e-5	0.8	1	[512]
30	180	512	4e-4	0.9	2	[1024, 68]
34	140	128	1e-3	0.9	2	[1024, 508]
35	100	256	1e-2	0.7	3	[1024, 512, 88]
37	120	512	8e-5	0.88	2	[1024, 512]
42	160	128	4e-5	0.9	1	[512]
46	120	128	1e-3	0.85	2	[1024, 512]
49	120	128	2e-5	0.74	1	[128]
51	120	1024	2e-6	0.45	2	[1146, 886]
56	260	1024	2.5e-6	0.5	2	[1026, 484]

Table 1: MLP hyperparameters for each station.



(a) Station index = 0

(b) Station index = 17

Figure 15: The training and validation, losses and accuracies during training of MLP for stations with indices 0 and 1.

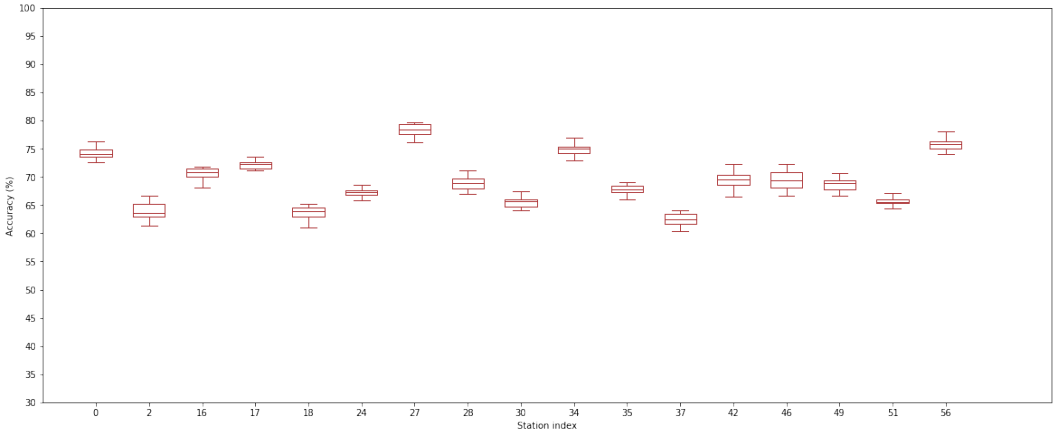


Figure 16: Boxplots of the accuracy of the MLP, for each station. The number of iterations (per boxplot) is 20.

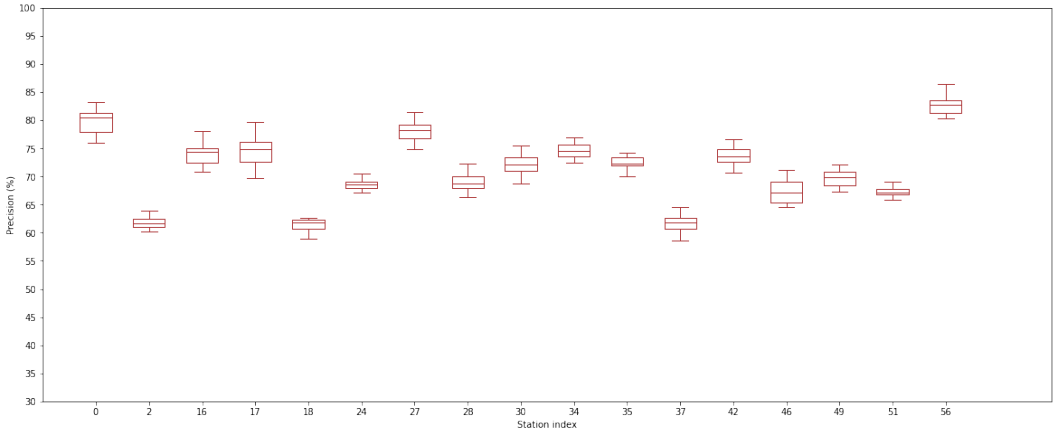


Figure 17: Boxplots of the precision of the MLP, for each station. The number of iterations (per boxplot) is 20.

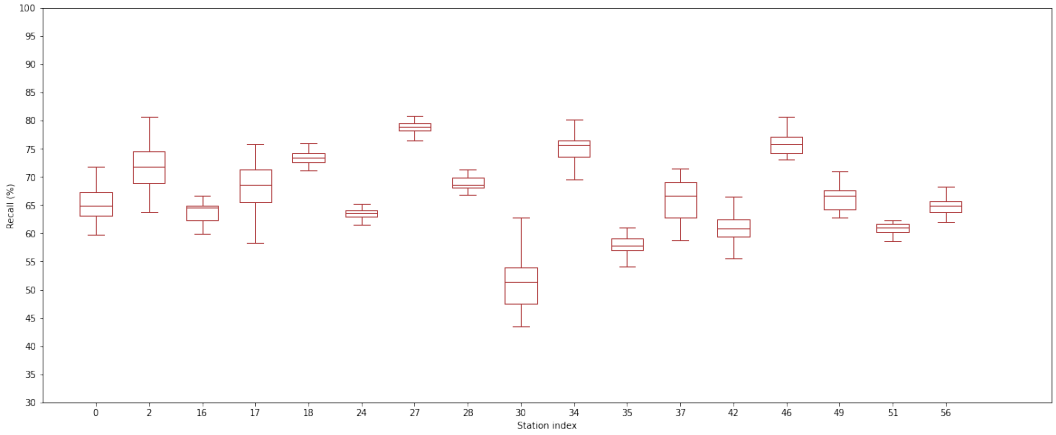


Figure 18: Boxplots of the recall of the MLP, for each station. The number of iterations (per boxplot) is 20.

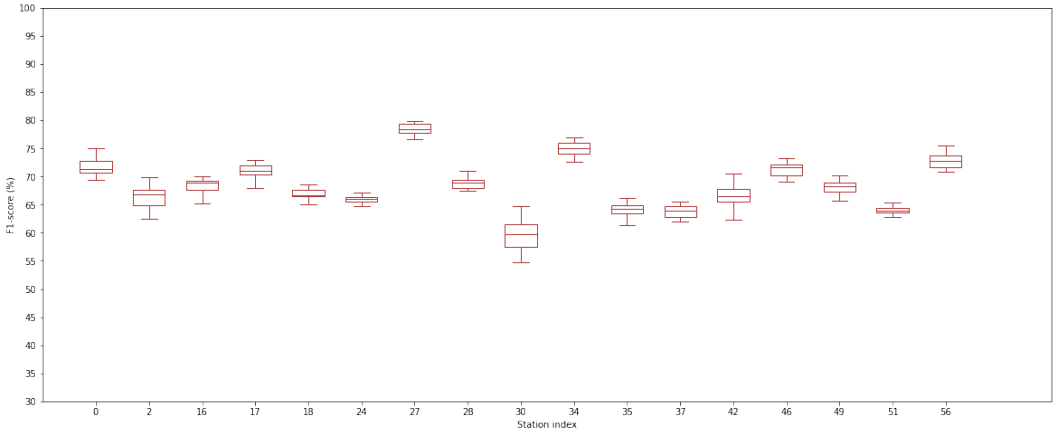


Figure 19: Boxplots of the F1-score of the MLP, for each station. The number of iterations (per boxplot) is 20.

4 Responsible Research

Time series classification of earthquakes is a developing field and an accurate and generalizable earthquake predictor could have great benefits towards hazard prevention. The proposed models can predict earthquakes 30 seconds before they occur at a specific station. This allows for more insight towards local seismic activity and could provide more time for safety procedures, therefore potentially saving lives. Additionally, this research could be a stepping stone towards learning more about the patterns of seismic activity. This is useful, as more knowledge on this topic could mean that the time window for earthquake prediction could be elongated.

That being said, there are some considerations that need to be taken into account. Firstly, earthquake prediction models could create a false sense of security. The possible

occurrence of earthquakes is reduced to a binary classification task, thereby oversimplifying a complex reality. As a result, these predictions can never be presented as certain. Moreover, currently the models give a binary result (an earthquake is about to happen or not), but do not provide a probability score. However, they could be modified to do so. This would bring about a moral question: what is the probability threshold to start taking action against an earthquake? Do you start preparing at a chance of 60 per cent? Or 90?

Nevertheless, this research should be easily reproducible. The data about earthquakes in New Zealand is still available and easily accessible as it is open source. Moreover, the data collection and analysis methods were described in section 2.1. Here, the specific decisions made to manipulate the data set can also be found. The used models can be recreated using the instructions in section 2.2 and tools from the PyTorch and the sklearn libaries. Finally, in section 3 the hyperparameters are stated for the MLP model. Therefore, anyone with background in machine learning should be able to replicate this research and come up with similar results.

5 Conclusion

This section provides conclusions based on the results from section 3.

In this study, MLP and SVM were implemented to classify earthquake prediction problems in a time series basis. The precursor parameter was seismic activity of 30 seconds measured at individual stations in New Zealand. The results showed satisfactory performances of MLP and SVM. It can therefore be concluded that both models can be applied as an earthquake prediction tools on a time-series basis. Moreover, looking at the general results shows that one model is not favored over the other. It is only when looking at specific stations, that a marginal difference in performance can be found. However, the kernel function that performed best as RBF for the SVM model. This reveals an interesting characteristic of seismic waveforms, namely: they are similar to other seismic waveforms of different stations.

6 Discussion and Future Work

In this section, the conclusions will be contextualized, whilst also considering their contribution to the academic field, their limitations and the possibilities for future work.

As discussed above, both MLP and SVM are relatively good at predicting earthquakes, although the performances are largely station-specific. Therefore, this research provides an valuable contribution to the academic field, as these models have not been used for classifying earthquake prediction problems before in a time series basis.

However, there are also some limitations to this study. Firstly, only frequency of 50HZ was used. This could affect the performance of one of the models. A lower frequency means fewer features, which could incorrectly favor one model over the other. Secondly, the only type of seismic activity considered in this experiment is vertical velocity. This is in contrast to [9], which considers the velocity in three dimensional space. Additionally, seismic activity is always vulnerable to noise caused by environmental factors. This could compromise the dataset. Fourthly, the sensor used to measure seismic activity in New Zealand is a weak motion sensor. This means it is subject to picking up seismic activity beyond the scope of this study, as they took place outside of New Zealand.

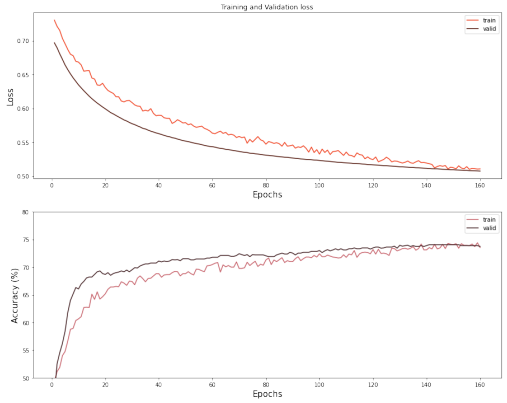
Moreover, the predictions in this study are bound to specific stations, despite the fact that having input from different stations could benefit the quality of the predictions. Lastly, the predictions only forecast whether an earthquake is about to occur, and not the magnitude or specific location. Therefore, these two limitations are appropriate starting points for future research, as having this information would benefit hazard prevention.

A Appendix

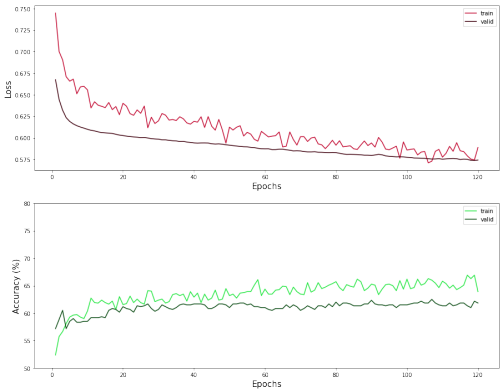
index	station	latitude	longitude	site name
0	BFZ	-40.679647	176.246245	Birch Farm
1	BHW	-41.408231	174.871115	Baring Head
2	BKZ	-39.165666	176.492544	Black Stump Farm
3	COVZ	-39.199914	175.542402	Chateau Observatory
4	CVZ	-44.383180	171.006160	Cave
5	DCZ	-45.464713	167.153533	Deep Cove
6	DSZ	-41.744961	171.804614	Denniston North
7	EAZ	-45.231053	169.308253	Earnscleugh
8	ETVZ	-39.135700	175.710600	East Tongariro
9	FWVZ	-39.254945	175.552952	Far West
10	GRZ	-36.250200	175.457800	Great Barrier Island
11	GVZ	-42.967365	173.034750	Greta Valley
12	HAZ	-37.756100	177.782600	Te Kaha
13	HIZ	-38.512929	174.855686	Hauiti
14	INZ	-42.724500	171.444100	Inchbonnie
15	JCZ	-44.073210	168.785473	Jackson Bay
16	KHEZ	-39.294200	174.014500	Kahui Hut
17	KHZ	-42.415980	173.538970	Kahutara
18	KNZ	-39.021755	177.673669	Kokohu
19	KUZ	-36.745229	175.720873	Kuaotunu
20	LBZ	-44.385553	170.184420	Lake Benmore
21	LTZ	-42.781667	172.271035	Lake Taylor Station
22	MLZ	-45.366544	168.118407	Mavora Lakes
23	MQZ	-43.706082	172.653766	McQueen's Valley
24	MRZ	-40.660545	175.578527	Mangatainoka River
25	MSZ	-44.673334	167.926399	Milford Sound
26	MWZ	-38.334001	177.527779	Matawai
27	MXZ	-37.562259	178.306631	Matakaoa Point
28	NNZ	-41.217103	173.379477	Nelson
29	ODZ	-45.043982	170.644622	Otahua Downs
30	OPRZ	-37.844300	176.554929	Ohinepanea
31	OPZ	-45.884356	170.597767	Otago Peninsula
32	OUZ	-35.219689	173.596133	Omahuta
33	OXZ	-43.325900	172.038300	Oxford
34	PUZ	-38.071548	178.257209	Puketiti
35	PXZ	-40.030644	176.862145	Pawanui
36	QRZ	-40.825522	172.529148	Quartz Range

37	RATZ	-38.866498	175.772176	Rangitukua
38	RPZ	-43.714608	171.053865	Rata Peaks
39	RTZ	-38.615440	176.980518	Ruatahuna
40	SYZ	-46.536890	169.138823	Scrubby Hill
41	THZ	-41.762474	172.905218	Top House
42	TLZ	-38.329400	175.538000	Tolley Road
43	TMVZ	-39.115610	175.704064	Te Maari
44	TOZ	-37.730956	175.501847	Tahuroa Road
45	TRVZ	-39.298816	175.547822	Turoa
46	TSZ	-40.058553	175.961124	Takapari Road
47	TUZ	-45.953975	169.631143	Tuapeka
48	URZ	-38.259249	177.110894	Urewera
49	VRZ	-39.124341	174.758453	Vera Road
50	WCZ	-35.938642	174.345043	Waipu Caves
51	WEL	-41.284048	174.768184	Wellington
52	WHVZ	-39.282500	175.588600	Whangaehu Hut
53	WHZ	-45.892428	167.947031	Wether Hill Road
54	WIZ	-37.524511	177.189302	White Island
55	WKZ	-44.827021	169.017562	Wanaka
56	WSRZ	-37.518110	177.177805	White Island Summit
57	WVZ	-43.074350	170.736754	Waitaha Valley

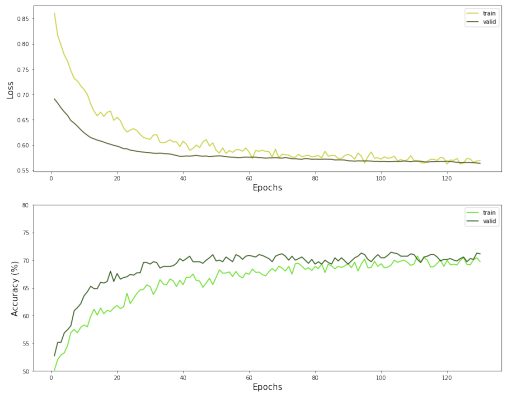
Table 2: Stations for the New Zealand earthquake dataset.



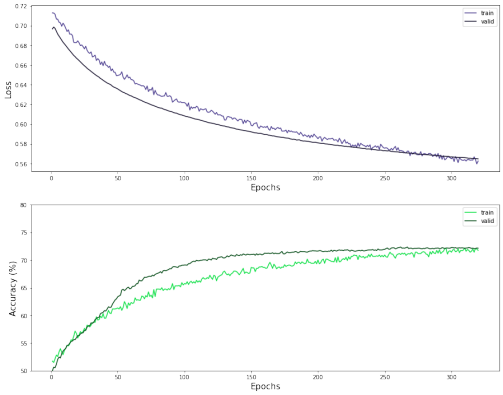
(a) Station index = 0



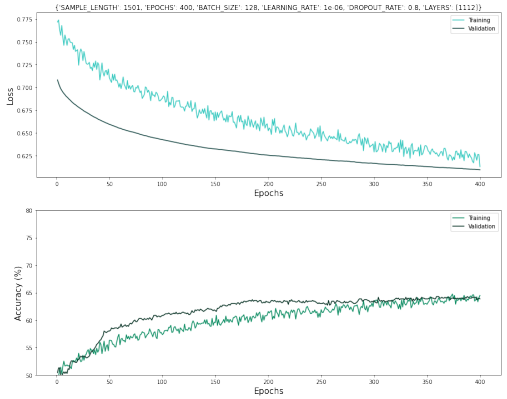
(b) Station index = 2



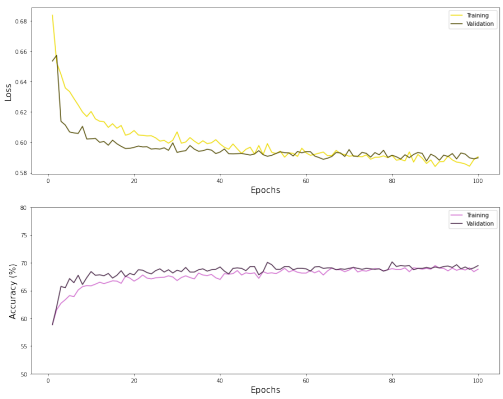
(c) Station index = 16



(d) Station index = 17

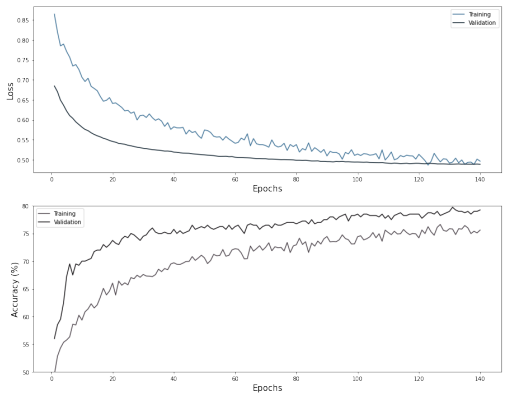


(e) Station index = 18

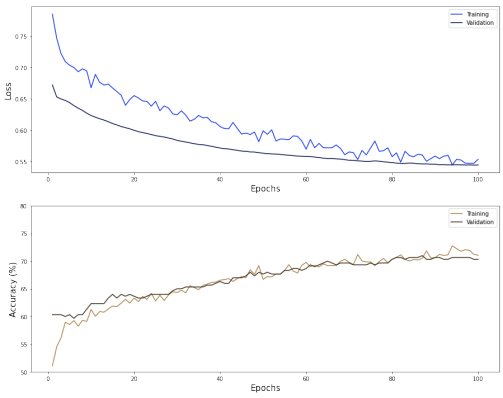


(f) Station index = 24

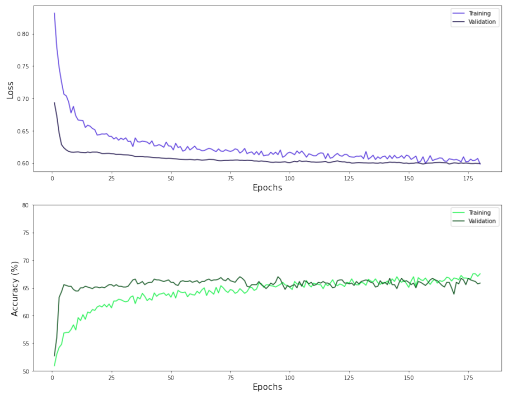
Figure 20: The training and validation, losses and accuracies during training of MLP for stations with indices: 0, 2, 16, 17, 18 and 24.



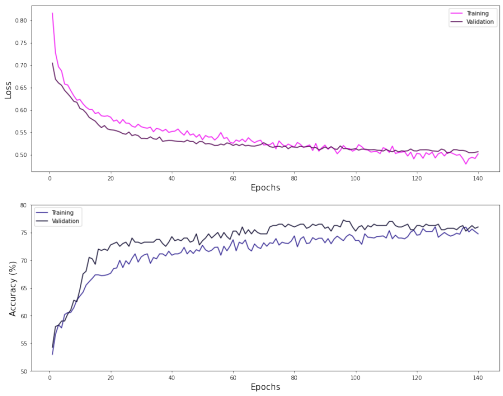
(a) Station index = 27



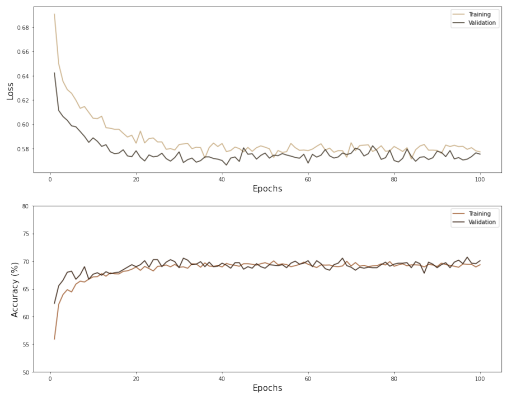
(b) Station index = 28



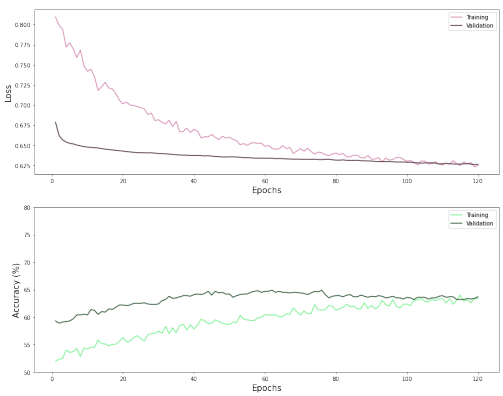
(c) Station index = 30



(d) Station index = 34

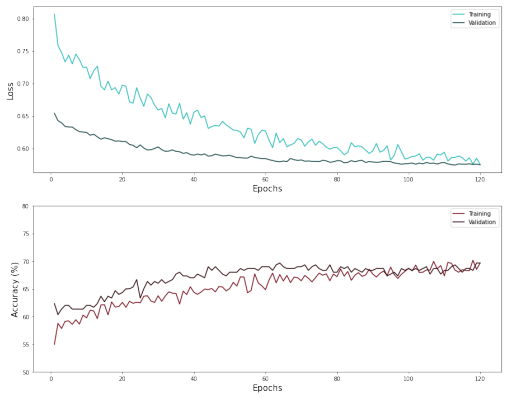


(e) Station index = 35

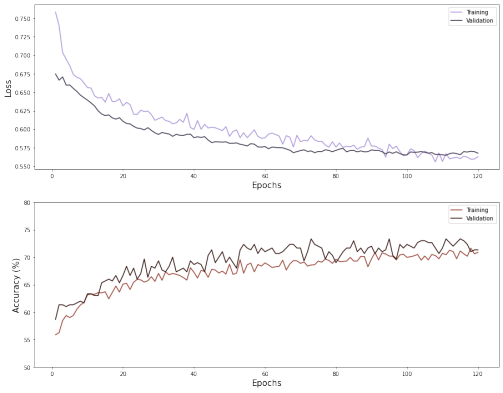


(f) Station index = 37

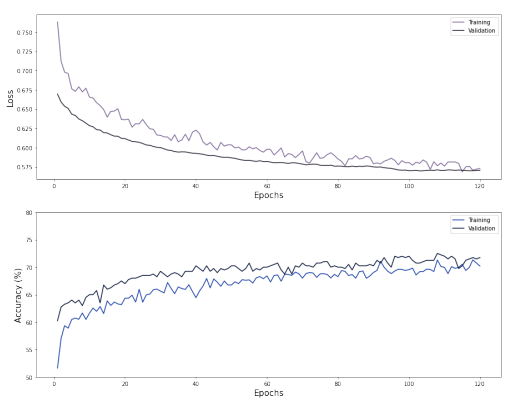
Figure 21: The training and validation, losses and accuracies during training of MLP for stations with indices: 27, 28, 30, 34,35 and 37.



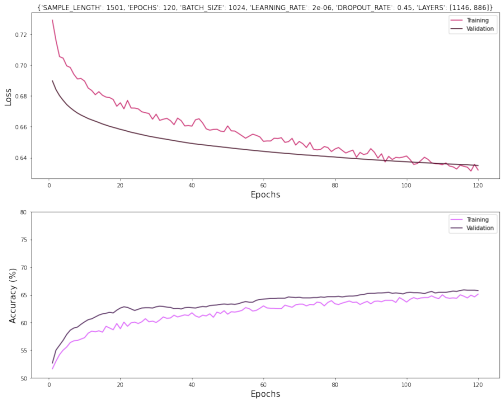
(a) Station index = 42



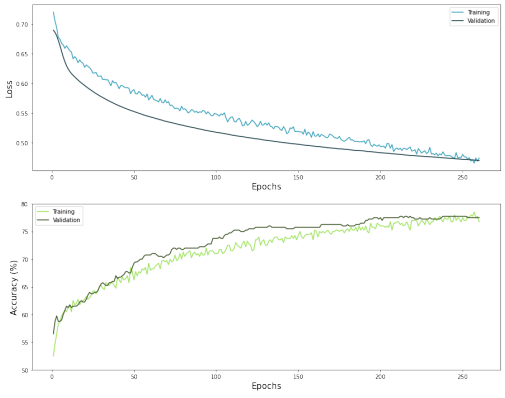
(b) Station index = 46



(c) Station index = 49



(d) Station index = 51



(e) Station index = 56

Figure 22: The training and validation, losses and accuracies during training of MLP for stations with indices: 42, 46, 49, 51 and 56.

References

[1] M. Hayakawa, "Earthquakes and eq prediction", pp. 1–17, Jul. 2015. DOI: 10.1002/9781118770368.ch1.

[2] M. Wyss and D. C. Booth, "The IASPEI procedure for the evaluation of earthquake precursors", *Geophysical Journal International*, vol. 131, no. 3, pp. 423–424, Dec. 1997, ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1997.tb06587.x. eprint: https://academic.oup.com/gji/article-pdf/131/3/423/6097657/131-3-423.pdf. [Online]. Available: https://doi.org/10.1111/j.1365-246X.1997.tb06587.x.

[3] M. Yousefzadeh, S. A. Hosseini, and M. Farnaghi, "Spatiotemporally explicit earthquake prediction using deep neural network", *Soil Dynamics and Earthquake Engineering*, vol. 144, p. 106 663, 2021, ISSN: 0267-7261. DOI: https://doi.org/10.1016/j.soildyn.2021.106663. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0267726121000853.

[4] M. Wyss and D. C. Booth, "The IASPEI procedure for the evaluation of earthquake precursors", *Geophysical Journal International*, vol. 131, no. 3, pp. 423–424, Dec. 1997, ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.1997.tb06587.x. eprint: https://academic.oup.com/gji/article-pdf/131/3/423/6097657/131-3-423.pdf. [Online]. Available: https://doi.org/10.1111/j.1365-246X.1997.tb06587.x.

[5] A. Bhatia, S. Pasari, and A. Mehta, "Earthquake forecasting using artificial neural networks", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2018.

[6] T. Bhandarkar, V. K, N. Satish, S. Sridhar, R. Sivakumar, and S. Ghosh, "Earthquake trend prediction using long short-term memory rnn", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, p. 1304, Apr. 2019. DOI: 10.11591/ijece.v9i2.pp1304-1312.

[7] Q. Wang, Y. Guo, L. Yu, and P. Li, "Earthquake prediction based on spatio-temporal data mining: An lstm network approach", *IEEE Transactions on Emerging Topics in Computing*, vol. 8, no. 1, pp. 148–158, 2020. DOI: 10.1109/TETC.2017.2699169.

[8] C. E. Yoon, O. O'Reilly, K. J. Bergen, and G. C. Beroza, "Earthquake detection through computationally efficient similarity search", *Science Advances*, vol. 1, no. 11, e1501057, 2015. DOI: 10.1126/sciadv.1501057. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.1501057. [Online]. Available: https://www.science.org/doi/abs/10.1126/sciadv.1501057.

[9] T. Perol, M. Gharbi, and M. Denolle, "Convolutional neural network for earthquake detection and location", *Science Advances*, vol. 4, no. 2, e1700578, 2018. DOI: 10.1126/sciadv.1700578. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.1700578. [Online]. Available: https://www.science.org/doi/abs/10.1126/sciadv.1700578.

[10] "Fdsn webservice for new zealand". (), [Online]. Available: https://www.geonet.org.nz/data/tools/FDSN. Accessed: 17.01.2022.

[11] G. Mazzola, "Graph-time convolutional neural networks", *MSc thesis, TU Delft*, 2020.

[12] "Automatic classification of seismic events within a regional seismograph network", *Computers & Geosciences*, vol. 87, pp. 22–30, 2016, ISSN: 0098-3004. DOI: https://doi.org/10.1016/j.cageo.2015.11.006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098300415300832.

[13] S. J. Raudys, A. K. Jain, *et al.*, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners", *IEEE Transactions on pattern analysis and machine intelligence*, vol. 13, no. 3, pp. 252–264, 1991.

[14] A. Ortiz, F. Bonnin-Pascual, E. Garcia-Fidalgo, and J. Company-Corcoles, "Vision-based corrosion detection assisted by a micro-aerial vehicle in a vessel inspection application", *Sensors*, vol. 16, p. 2118, Dec. 2016. DOI: 10.3390/s16122118.