Short-term Earthquake Prediction via Recurrent Neural Networks

Xiangyu Du x.du-1@student.tudelft.nl | Computer Science and Engineering, TU Delft

Introduction

- One of the most devastating catastrophes on earth
- Intrinsic nature is random [1] → hard to forecast
- Little research on short term forecast
- Compare the performances of different recurrent neural networks: vanilla RNN, LSTM and Bi-LSTM

Research Question

How do individual time series model compare with each other (vanilla RNN, LSTM, Bi-LSTM)?

Focusing on short term prediction (30 seconds in advance)

Experiment

Step 1: **Data Preprocessing**

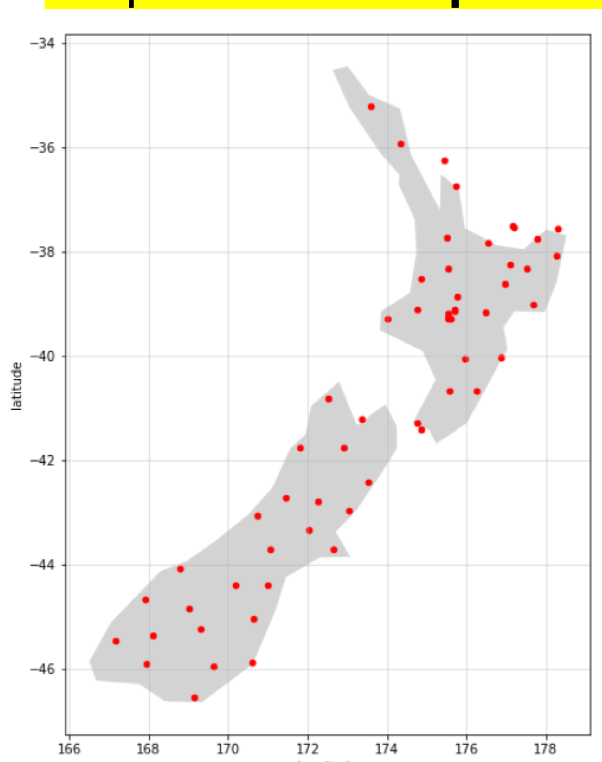


Figure 1: Dots represent 58 stations

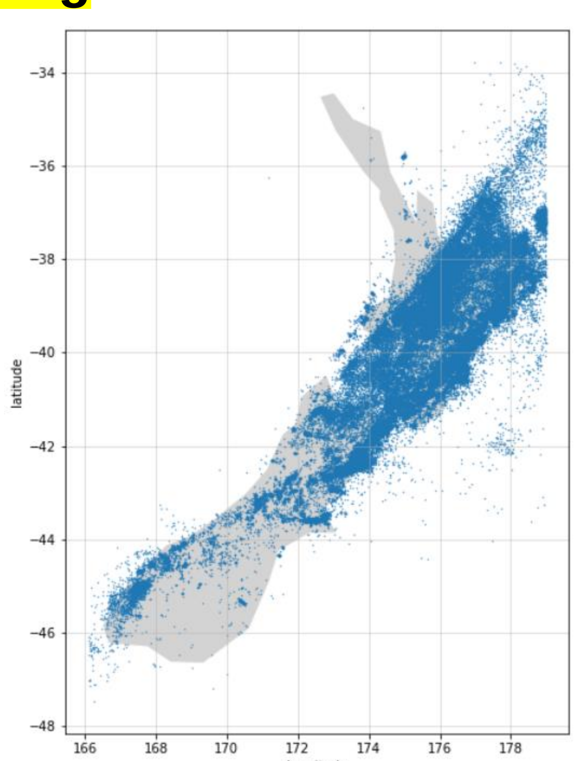


Figure 2: Earthquakes happened in New Zealand between 2016 and 2020

- Earthquakes happened in New Zealand between 2016 and 2020 [2] (123165)
- Filter out data without time, latitude, longitude, magnitude and depth measurements (122465)

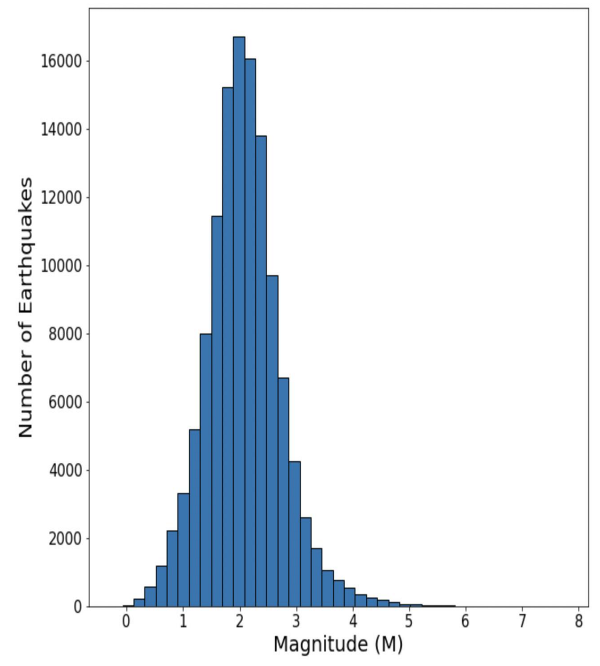


Figure 3: Distribution of Magnitude

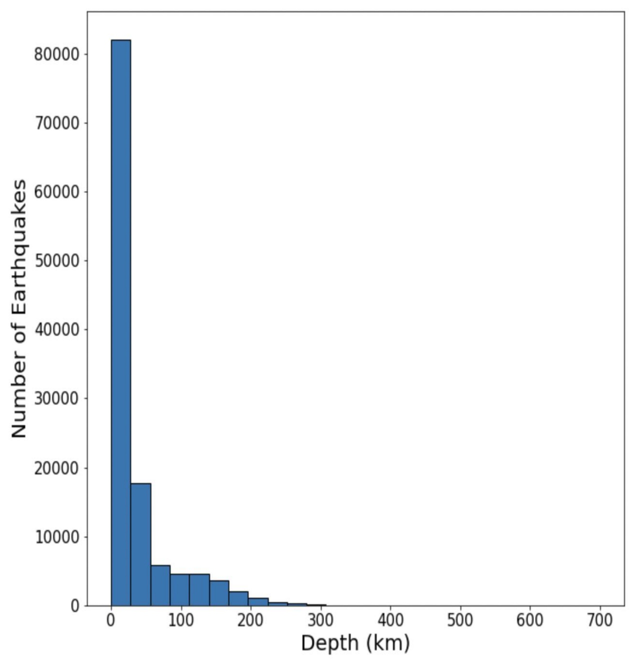


Figure 4: Distribution of Depth

- Filter earthquakes with magnitude between 1 and 3 & with depth less than 200km. (106623)

Figure 5: 100HZ, not normalized seismic waveforms

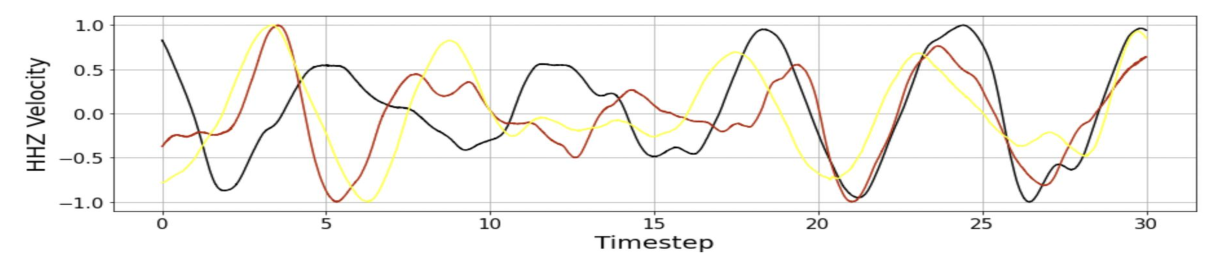


Figure 6: 50HZ, downsampled and normalized seismic waveforms

- Downsample [3] signals

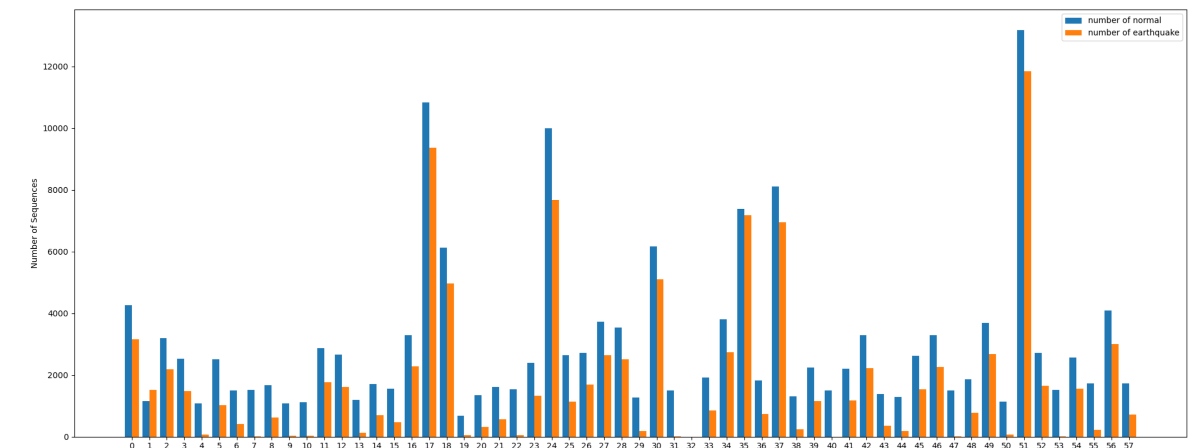


Figure 7: #earthquakes & normal signals assigned to each station

- Signals close to mid timestamp of two earthquakes are classified as normal → **160k normal signals**
- Assign each waveform signal to closest station

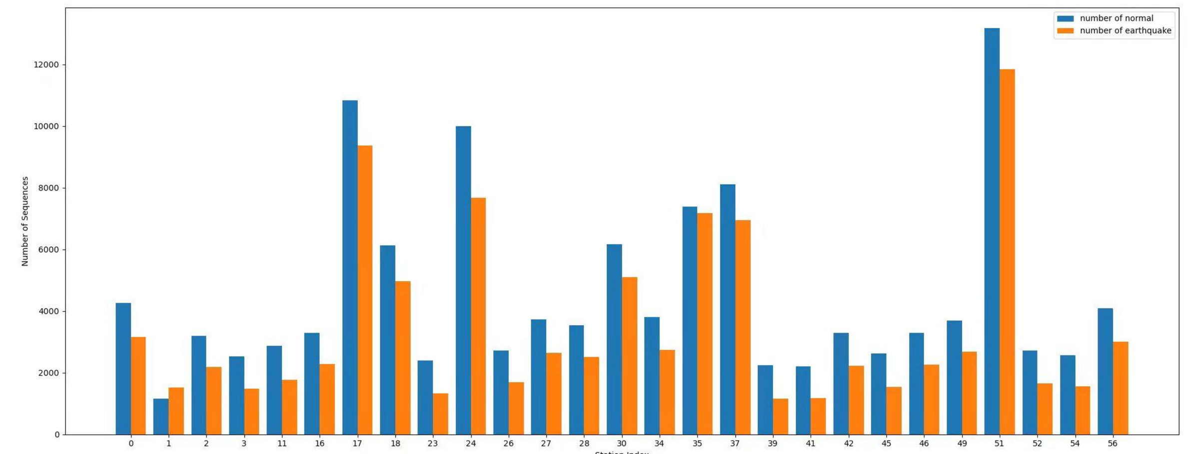


Figure 8: #earthquakes & normal signals assigned to each station after filtering

- Stations with #normal signals > 2 * #earthquake signals are discarded → **27 stations**

Step 2: **Training with Neural Networks**

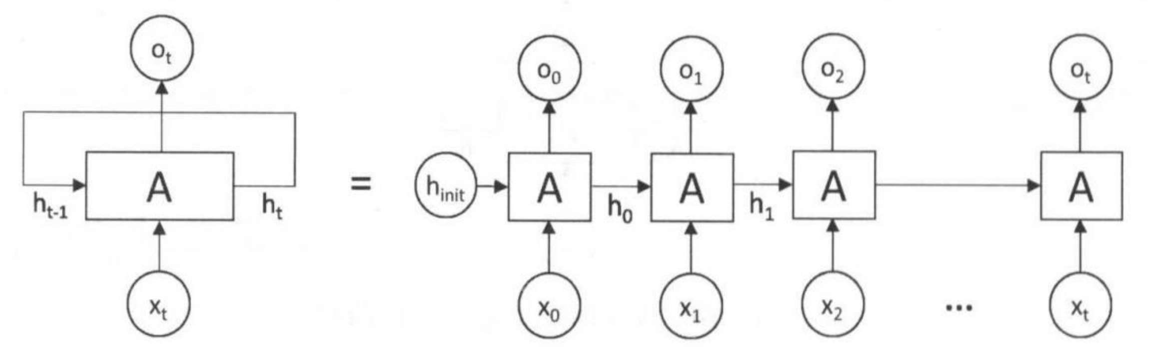


Figure 9: RNN structure [4]

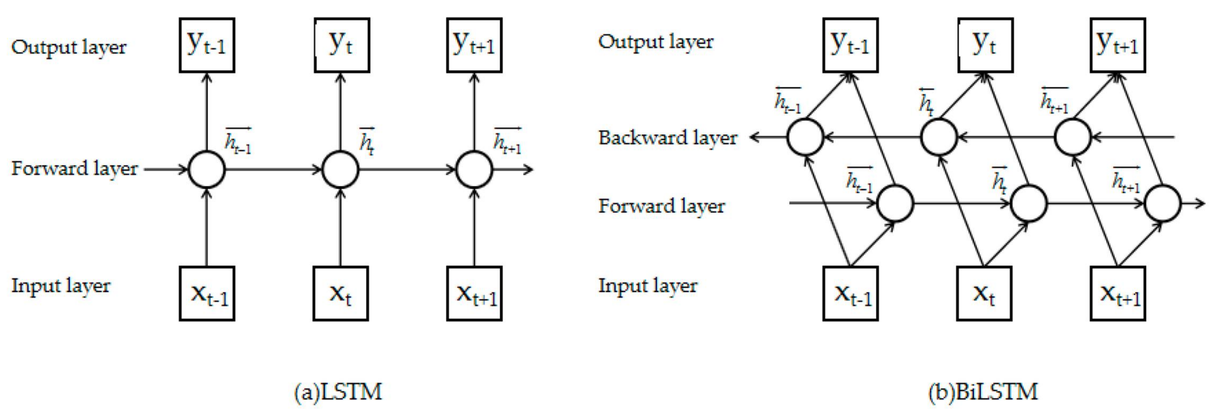


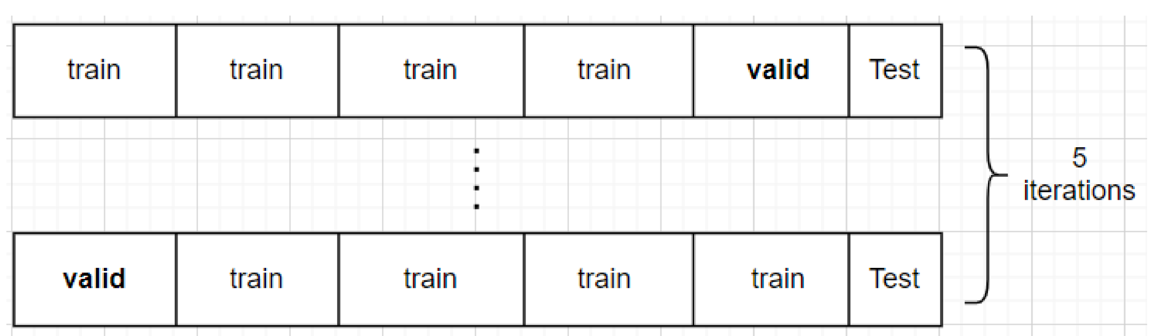
Figure 10: LSTM & Bi-LSTM structure [5]

Architecture:

- An RNN/LSTM/Bidirectional LSTM layer with input size **1** (one station at a time) and **128** neurons. (Two Bi-LSTM layers for Bi-LSTM)
- A Dropout layer (**0.2**)
- A Dense layer
- Activation function: **Sigmoid**
- Batchsize: **64**

Processing:

- **5-fold cross validation**

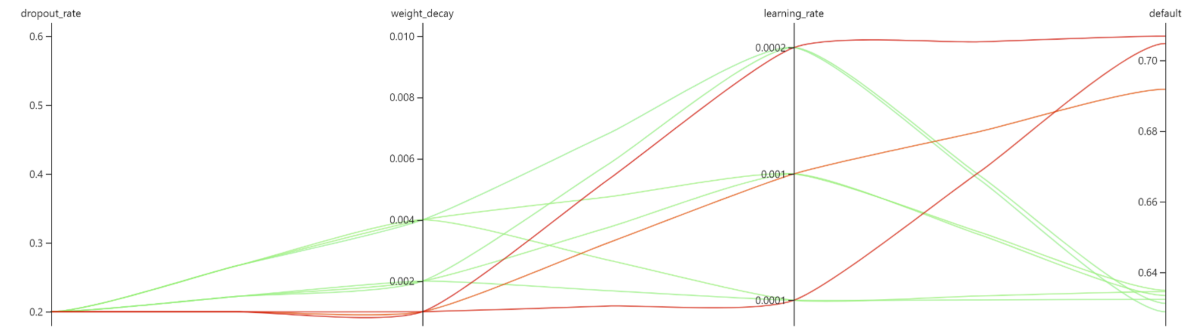


Step 3: **Dealing with Over-fitting**

- 5 have over-fitting → Grid Search

Hyperparameters	Parameter Grid
Dropout Rate	[0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5]
Regularization Parameter (Weight Decay)	[0.002, 0.004, 0.006, 0.008, 0.01]
Learning Rate	[0.00005, 0.0001, 0.00015, 0.0002]

- Result for one network for one station



Introduction

Step 4: **Evaluation: Confusion Matrix & (Weighted) Precision & Recall & F1 score & Accuracy, Boxplot**

	Accuracy	Precision	Recall	F1-score	FN	FP	TN	TP
vanilla RNN	0.655875	0.622033	0.655875	0.594271	296.8182	58.09091	155.8182	82.13636
LSTM	0.664583	0.646016	0.664583	0.604485	298.9364	55.97273	153.0818	84.87273
Bi-LSTM	0.656961	0.606048	0.656961	0.582474	313.6455	41.26364	170.6091	67.34545

Figure 9: Avg result of each metrics

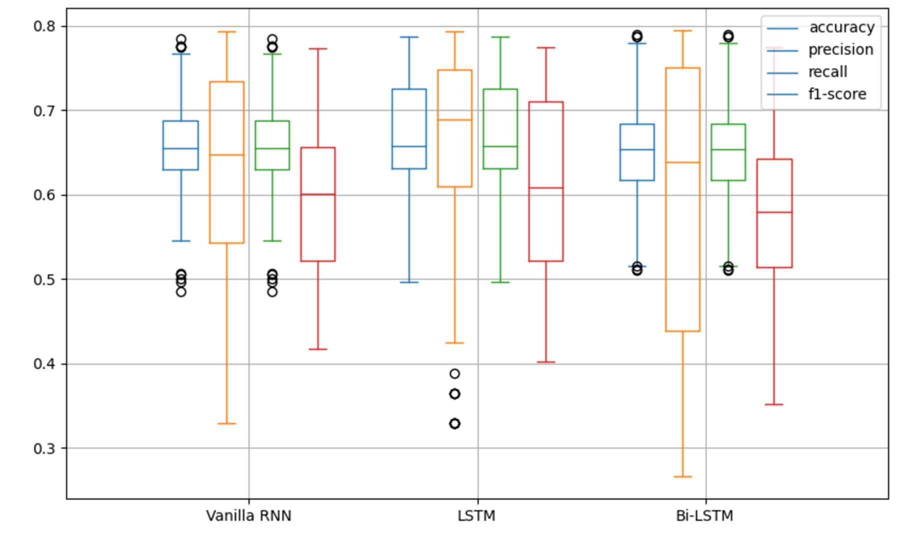


Figure 10: Boxplot of each metric over models

Conclusion

- In general, LSTM performs the best while vanilla RNN performs the worst. Bi-LSTM might suffer from noise in the data.
- All three models prone to correctly classify normal signals rather than earthquake signals.

Future work

- Deepen the layers of networks
- Grid Search on all stations with more parameters and larger ranges
- Try other deep learning models

References

[1] . Bhandarkar et al. "Earthquake trend prediction using long short-term memory RNN". In: International Journal of Electrical and Computer Engineering 9.2 (2019), p. 1304.

[2] https://www.geonet.org.nz/data/tools/FDSN

[3] L. Ruiz, F. Gama, and A. Ribeiro. "Gated Graph Recurrent Neural Networks". In: arXiv (2020), arXiv–2002.

[4] https://www.cnblogs.com/huangyc/p/10366783.html

[5] https://analyticsindiamag.com/complete-guide-to-bidirectional-lstm-with-python-codes/