House Price Prediction in Taipei by Machine Learning Models

Yu-Ren Lin, and Chien-Chang Chen

***Abstract*—House price in Taipei city is a widely discussed issue. Generally, the price is decided by people's understanding from past dealing-price. In the past decades, many researchers pay attention on the study of house price prediction by computer computation. Recently, different machine learning models are adopted to analyze the actual price registration dataset for predicting house price.**

This study examines linear regression, MLP (Multi-layer perceptron), and LSTM (Long Short-Term Memory) models on prediction of the actual price registration dataset. Various parameters and combinations are also test in our experiments. Experimental results show that LSTM deep neural network has better prediction than others. In the selection of optimizer, the Adam function exhibits better than SGD or RMSProp functions. In our limited experiments, single-layer deep neural network model leads to better results than different multi-layer deep neural network models.

***Index Terms*—House Price Prediction, Machine Learning, Multi-layer perceptron, Long Short-Term Memory.**

I. Introduction

Buying-house is the largest consumption in normal people life. As we knew the dealing price is not public for a long time, the dealing price is always determined by agency or seller. Therefore, most people rely on third party effort for avoiding blurred dealing price.

In 2012, Actual Price Registration was implemented by MOI (Ministry of Interior) to force seller price publicly announced on website. Build up the niche between house marketing also helps analysis of future dealing trend and the preparation before implement to taxation actual estate price. Nowadays, the rapid improvement on machine learning techniques and hardware devices makes the usage of applying machine learning techniques to predict being feasible.

In the field of applying artificial intelligence technique to predict house price, the Boston House price prediction competition, provided by the US Kaggle platform, is the most famous example [8]. The competition attracts lots of attendances to try to acquire the best estimation.

In this paper, we will examine important machine learning techniques for finding a best prediction on actual price registration dataset. The dataset is constructed between 2017 and 2019 from MOI. The selected models include linear regression, MLP and LSTM.

Chien-Chang Chen is with the Department of Computer Science and Information Engineering, Tamkang University, Taipei, Taiwan (corresponding author to provide e-mail: ccchen34@mail.tku.edu.tw).

Remaining of the paper is organized as follows. Section 2 reviews related works about this study. Section 3 shows the proposed models. Experimental results are demonstrated at Section 4. Brief conclusions are given in Section 5.

II. Related Works Review

A. Actual Price Registration

MOI in Taiwan aggressively implements actual selling price system in securing publics privacy way. In the actual price registration dataset, there are totally 19 items should be recorded in the actual price registration dataset. In our experiments, part of these items are selected for training the system and predicting the house price.

B. Machine Learning and Deep Learning

Machine learning is one of the important branches of artificial intelligence. Machine learning is good at dealing with four main problems, classification, clustering, regression, and dimensionality reduction. Regression deals with continuous data prediction, as we know house price is continuous in time series, so in terms of machine learning field, it belongs regression issue. Sifei Lu et al [4] introduced a variety of housing price predicting techniques using regression methods.

Deep learning is a branch of machine learning. Deep learning simulates human neural network operation to transact large and complex data for making important decision.

Time series predictions such as house prices, stock prices, and exchange rates are suitable for using algorithms with memory functions. Some study explores the Deep learning technology in housing price prediction, especially the performance of time series [6]. In deep learning algorithms, LSTM (Long Short-Term Memory) performs well in dealing the memory related issue [2, 6].

In this study, we focus on establishing LSTM model with different parameters. Moreover, in our experiments, we also adopt different layered MLP for estimating the house price.

C. Data preprocessing

The dataset in our experiments is extracted from MOI of the record within the period of 2017 to 2019. The data downloaded from website has the format of .csv or .xls. A serious of pre-processing should be applied to the dataset. After applying the following steps, including price unit conversion, new aging field, administrative code, area unit conversion, missing values fill-in, extreme values removal, land and berth transaction data removal, regularization.

Although there are 19 items in the dataset, we empirically adopt total price yuan, building shifting ping, room number, hall, health, berth shifting ping, berth total price yuan, house age, complete year, transaction year.

Figure 1 shows the transaction numbers of different administrative districts in Taipei city. In this figure, Zhongshan District has the highest transaction numbers and Nangang District has the least transaction numbers.

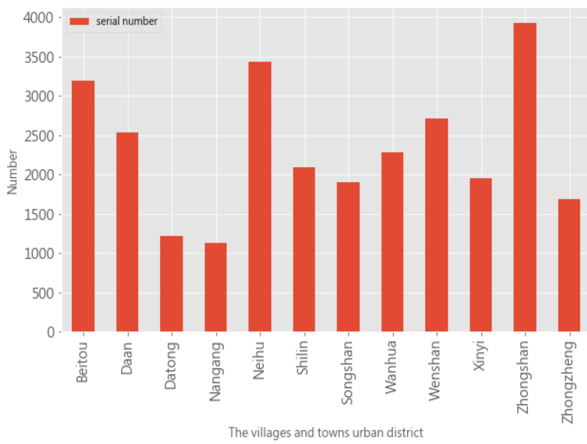


Fig. 1. Transaction numbers in different districts of Taipei city.

Figure 2 shows the correlation coefficient heat map, in which the lighter color shows the higher correlated degree.

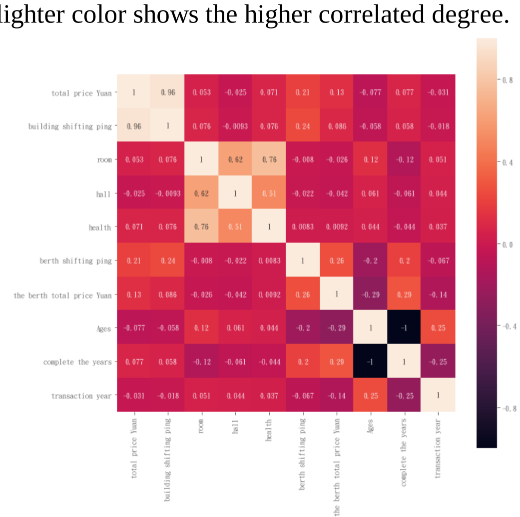


Fig. 2. Correlation coefficient heat map.

The residential elite area Daan has the highest average unit price per ping, while the average unit price of Wenshan and Beitou Districts are relatively low.

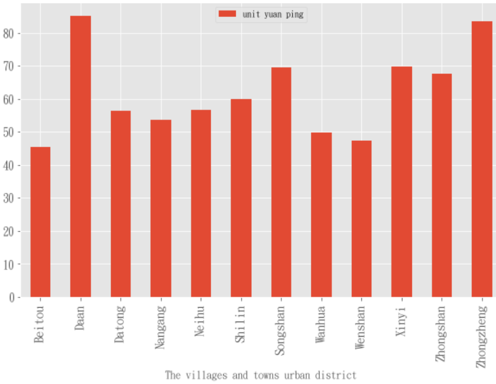


Fig. 3. Average unit price per ping in each administrative district.

III. PROPOSED MODELS

In this section, structures of three different models linear regression, MLP and LSTM are introduced. Section 3.1 introduces the selected parameters, like loss function, Dynamic Time Warping (DTW), and Optimizer.

A. *Parameters Selection*

The mean squared error (MSE) is used as the loss function indicator to evaluate the performance of the neural network. Eq. (1) shows the calculation of MSE.

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_t)^2 \tag{1}$$

Dynamic Time Warping (DTW) is used to measure the similarity between two series data. In this paper, we use the optimized DTW algorithm of deep learning: FastDTW [7] for distinguishing performance of a model. The Euclidian distance calculation (2) is needed for acquiring FastDTW value.

$$\mathrm{d}(x, y) = \sqrt{\sum_{i=1}^{m}(x_i - y_i)^2} \tag{2}$$

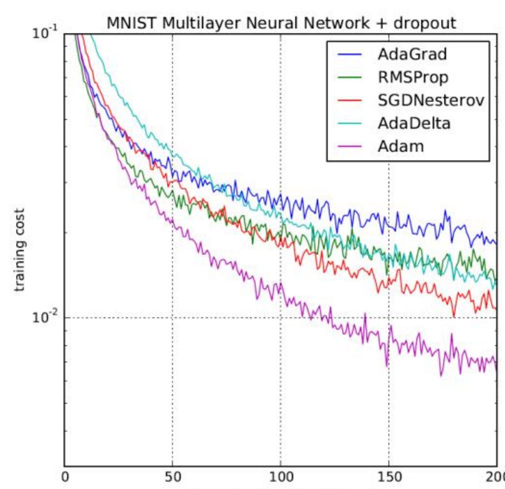


Fig. 4. Optimizer comparison chart.

There are many optimizers can be selected in deep learning model. In our experiments, we choose popularly used SGD, RMSprop, and Adam. Although Figure 4 shows that the Adam is presented as a significant loss in MNIST dataset among these three optimizers [3], the study finds whether this property also existed in house price prediction.

B. *Research Architecture*

This study uses a series step to predict house price. Figure 5 shows the steps of the proposed method, in which the data collection, data preparation, data training, and prediction evaluation. Especially in the training procedure, three different methods like linear regression, MLP and LSTM are trained.

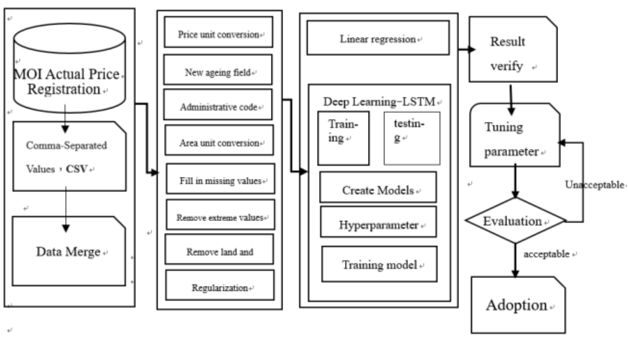


Fig. 5. Research architecture of the paper.

IV. Experimental Results

This section shows experimental results of the proposed scheme and these results are compared through MSE and DTW. Four different experiments are examined as follows.

1. House price prediction by simple and multiple linear regressions.

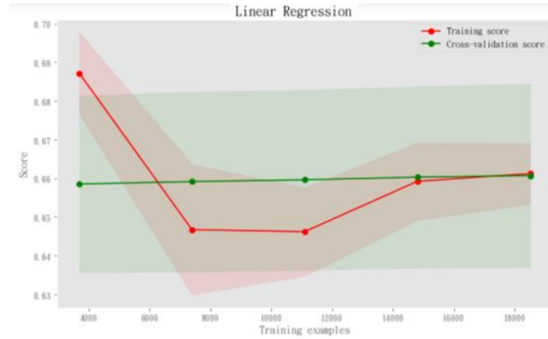
2. House price prediction by single-layered LSTM and different optimizers.

3. House price prediction by multi-layered LSTM.

4. House price prediction by multi-layered LSTM with Consumer Price Index (CPI) feature.

A. *Experiment 1: House price prediction by simple and multiple linear regressions*

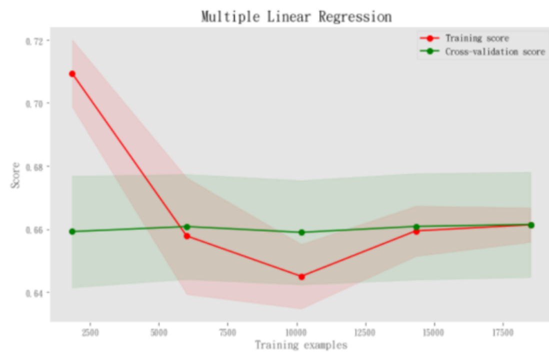
A single feature of building transfer total area under a simple linear regression is illustrated in Figure 6. Multiple linear regression models by 2, 3, and 6 features are shown in Figures 7, 8, and 9, respectively. In these experiments, 2 features are building shifting ping area and age; 3 features are building shifting ping, age and room; 6 features are 3 features and number of health, berth shifting ping and t the berth total price.



MSE: 2582811.35

R Square: 0.689063037761

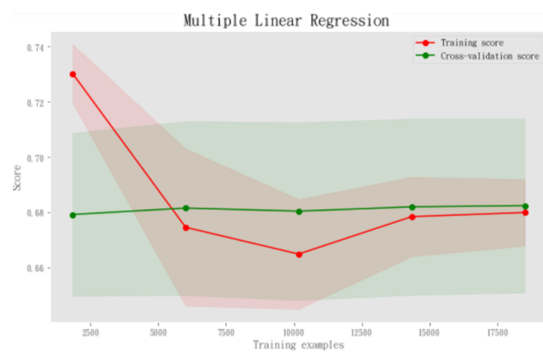
Fig. 6. Simple linear regression training results.



MSE: 2823091.77

R Square: 0.667606606621

Fig. 7. 2 Features' multiple linear regression training results.



MSE: 3304323.80

R Square: 0.644825887619

Fig. 8. 3 Features' multiple linear regression training results.

The above experimental results show that the overall MSE values are both high and the prediction accuracy is low. Among these linear regression models, using 6 features acquires best evaluation results both in MSE and R Square values.

MSE: 2429138.93

R Square: 0.738285031822

Fig. 9. 6 Features' multiple linear regression training results.

B. *Experiment 2: House price prediction by single-layered LSTM and different optimizers*

Features of building shifting ping, Age, room, health, berth shifting ping and the berth total price are applied to LSTM and different optimizers. Combinations of optimizers, batch sizes, and epochs are experimented in this subsection. Figure 10 shows the proposed LSTM model. Figure 11 and 12 show the loss function and fit results of the SGD optimizer, 20 batch size, and 100 epochs, respectively.

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 256)	268288
dropout_1 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 1)	257

Total params: 268,545
Trainable params: 268,545
Non-trainable params: 0

Fig. 10. The proposed LSTM model.

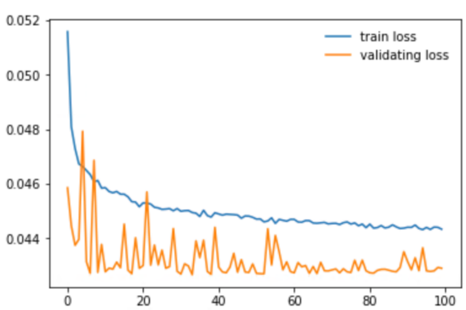
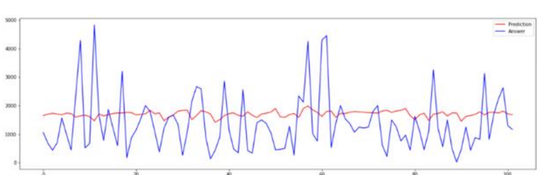


Fig. 11. The Loss Function training state with 100 epochs in the SGD optimizer.



FastDTW Distance:4387470.66

Fig. 12. Fit results for 100 epochs in the SGD optimizer.

Figure 13 and 14 show the loss function and fit results of the SGD optimizer, 20 batch size, and 200 epochs, respectively.

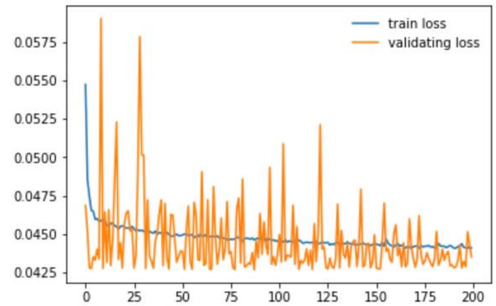
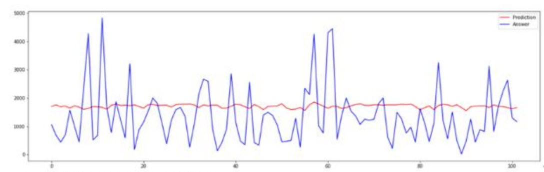


Fig. 13. The Loss Function training state with 200 epochs in the SGD optimizer.



FastDTW Distance: 4425107.73

Fig. 14. Fit results for 200 epochs in the SGD optimizer.

Figure 15 and 16 show the loss function and fit results of the RMSProp optimizer, 20 batch size, and 100 epochs, respectively.

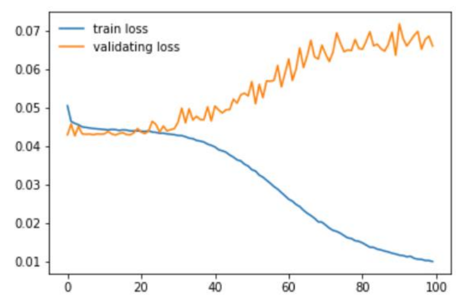
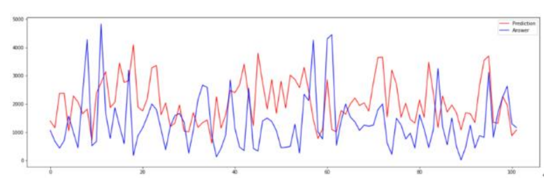


Fig. 15. The Loss Function training state with 100 epochs in the RMSProp optimizer.



FastDTW: 3620341.52

Fig. 16. Fit results for 100 epochs in the RMSProp optimizer.

Figure 17 and 18 show the loss function and fit results of the RMSProp optimizer, 20 batch size, and 200 epochs, respectively.

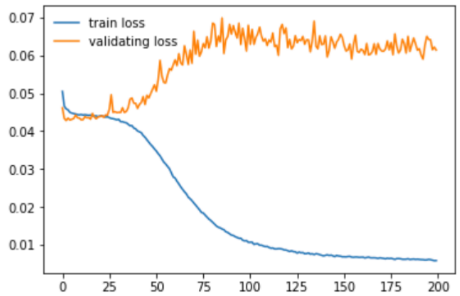


Fig. 17. The Loss Function training state with 200 epochs in the RMSProp optimizer.

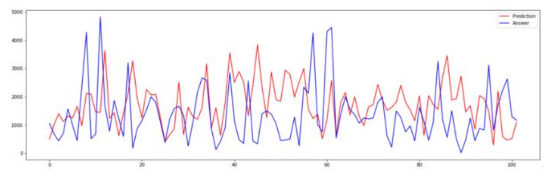

FastDTW: 3620341.52
Fig. 18. Fit results for 200 epochs in the RMSProp optimizer.

Figure 19 and 20 show the loss function and fit results of the Adam optimizer, 20 batch size, and 100 epochs, respectively.

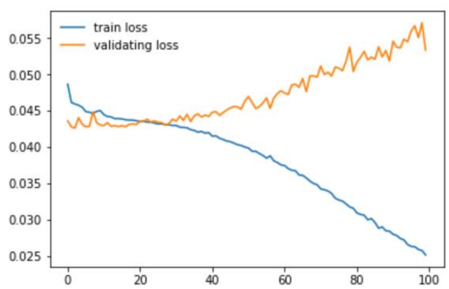


Fig. 19. The Loss Function training state with 100 epochs in the Adam optimizer.

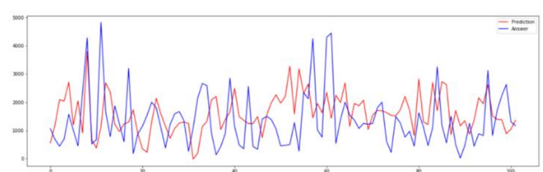

FastDTW: 3500581.28
Fig. 20. Fit results for 100 epochs in the Adam optimizer.

Figure 21 and 22 show the loss function and fit results of the Adam optimizer, 20 batch size, and 200 epochs, respectively.

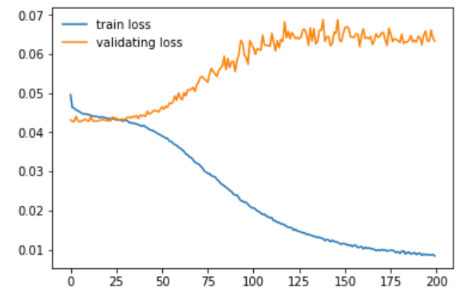
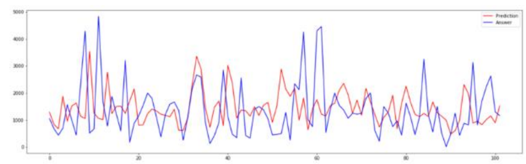


Fig. 21. The Loss Function training state with 200 epochs in the Adam optimizer.


FastDTW: 4191190.85
Fig. 22. Fit results for 200 epochs in the Adam optimizer.

From above FastDTW values, the Adam optimizer with 100 epochs has the best parameters selection among these experiments. The SGD has the worst prediction results, in which the prediction is hard to find significant match.

Figure 23 shows the loss function and fitting results of a conventional method [1]. Figure 24 depicts the fitting results. Our study shows that the proposed model performs better than the conventional method [1].

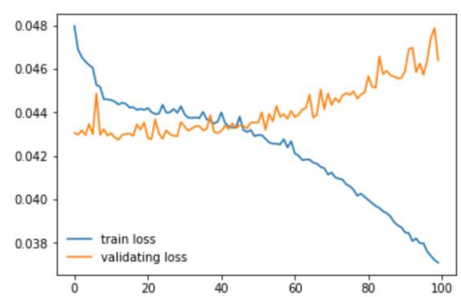


Fig. 23. Loss Function Training Status by the Adam optimizer [1].

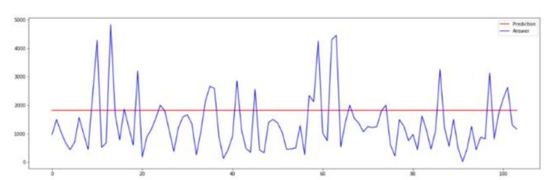


Fig. 24. Fitting results by the Adam optimizer [1].

C. *Experiment 3: House price prediction by multi-layered LSTM*

The house price is predicted by using a two-layered LSTM neural network with 256 neurons, 1 layer of MLP, and 1 output layer. Moreover, the drop-out rate is 0.2, batch size is 20, and epochs times is set to 100. Figure 25 shows the proposed multi-layered LSTM structure.

```
Layer (type)                 Output Shape              Param #
=================================================================
lstm_1 (LSTM)                (None, 30, 256)           268288
_________________________________________________________________
dropout_1 (Dropout)          (None, 30, 256)           0
_________________________________________________________________
lstm_2 (LSTM)                (None, 2)                 2072
_________________________________________________________________
dropout_2 (Dropout)          (None, 2)                 0
_________________________________________________________________
dense_1 (Dense)              (None, 4)                 12
_________________________________________________________________
dense_2 (Dense)              (None, 1)                 5
=================================================================
Total params: 270,377
Trainable params: 270,377
Non-trainable params: 0
_________________________________________________________________
```

Fig. 25. The proposed multi-layered LSTM structure.

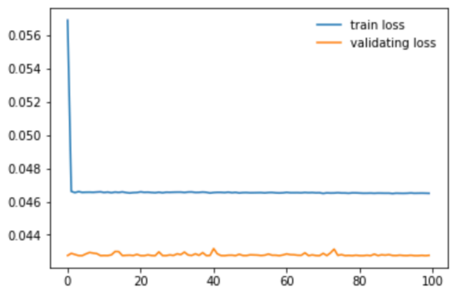


Fig. 26. Loss function training results of the multi-layered model.

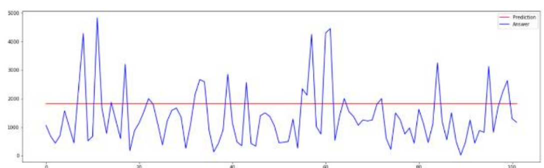


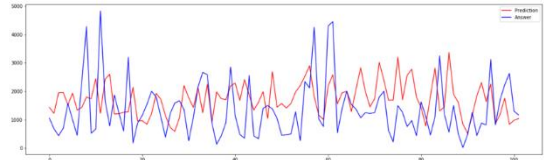
Fig. 27. Fitting results of the multi-layered LSTM model.

Figures 26 and 27 show the loss function and fit results of the multi-layered LSTM model. These two results show that using multi-layered LSTM model cannot acquire better results.

D. *Experiment 4: House price prediction LSTM with Consumer Price Index (CPI) feature*

In addition to the Actual Price Registration dataset, this paper to use the consumer price index (CPI) as an additional feature, Wan Teng Lim at el [5] mentioned the use of CPI can be used of house price prediction feature.

At last, we add feature of CPI to the previous one-layered LSTM model for finding the effect of CPI feature. Similar to previous structure, the experiment selects Adam optimizer, 20 batch size, and 100 epochs. The fitting results is illustrated in Figure 28. Comparing with Figure 20, the CPI feature cannot improve the prediction result. Further analyzing the CPI feature shows that the CPI features are very close in recent three years.



FastDTW Distance: 3626743.53

Fig. 28. Fitting results with adding consumer price index (CPI).

V. CONCLUSIONS

This paper predicts house price of Taipei city by different machine learning techniques, including linear regression, MLP, and LSTM models. The curves similarity function FastDTW is adopted for measuring the predict performance of each model. Among these models, the one-layered LSTM model with the Adam optimized, 20 batches, and 100 epochs is the best model in our limited experiments. Moreover, the CPI feature cannot improve the prediction performance. Since our experiments only analyze recent three-year data, more data with longer period merits our study for checking the importance of the CPI feature.

REFERENCES

[1] Jing-Quan Chen, Apply Deep Learning to Forecasting Average Price of Regional Residential Houses- with Application to Actual Price Registration in Taipei City, Master Thesis, Ming Chuan University, 2019.

[2] S. Hochreiter & J. Schmidhuber, LONG-SHORT-TERM MEMORY, Neural Computation, Vol.9, No.8, p.1735-1780, 1997.

[3] Diederik P. Kingma & Jimmy Lei Ba, ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION, 3rd International Conference for Learning Representations, San Diego, 2015.

[4] Sifei Lu & Zengxiang Li & Zheng Qin & Xulei Yang & Rick Siow Mong Goh, A Hybrid Regression Technique for House Prices Prediction, Institute of High Performance Computing, IEEE International Conference on Industrial Engineering and Engineering Management 2017.

[5] Wan Teng Lim & Lipo Wang & Yaoli Wang & Qing Chang, Housing Price Prediction Using Neural Networks, International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, ICNC-FSKD 2016.

[6] Li Yu & Chenlu Jiao & Hongrun Xin & Yan Wang & Kaiyang Wang, Prediction on Housing Price Based on Deep Learning, International Journal of Computer and Information Engineering, Vol.12, No.2. p.90-99, 2018.

[7] Stan Salvador & Philip Chan, FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space, Journal Intelligent Data Analysis, Vol.11 Issue 5, p.561-580, 2007.

[8] Boston Housing, https://www.kaggle.com/c/boston-housing