Investment Strategy – AI Trader (Foreign Exchange)

Hanlun Wang [*] **Zhaomeng Wang** [*]

Abstract

This report will cover the following six sections: Task Goal, Current Progress, Model, Performance Evaluation, Graphic User Interface and Future Work which are related to the Investment Strategy – AI Trader (Foreign Exchange) project.

1. Task Goal

In this project, we want to implement a forex AI trading system by using neural networks and machine learning techniques to learn to predict forex quotes from big data and to automate forex trading in a sophisticated way. This system should have the following features: low latency, which means it is a real-time system for volatile forex markets; user-selected investment strategies including aggressive, conservative and hedging strategies, which means the system can operate in different modes depending on the user; and finally, the system should have a built-in risk control mechanism that eliminates the risk of losing large amounts of money.

This AI trader system could be very helpful and significant as it overcomes the weaknesses of human traders such as greed, fear, and fluke. It prevents the effects of investor mood swings and ensures maximum returns while controlling risk.

This topic attempts to combine multidisciplinary knowledge that encompasses financial and economic literacy, especially quantitative investment using mathematics, statistics, information technology and other methods to manage investment portfolios. After collecting and analyzing a large amount of data, advanced mathematical models are used to replace subjective judgments with the help of computer systems.

2. Current Progress

We finish building the model. We used multidimensional inputs, including trends in foreign exchange prices, country GDP ratios, interest rate ratios, and several other relevant factors. We also completed the training of the model. We use LSTM model for training and adjust the parameters according to the training results to achieve the best accuracy. Finally, we completed the forecasting with the trained model. We input one month of data to predict the results for the coming week. We tested 100 results and finally got relatively stable error values. Then we completed the development of a visualization application using the trained model. A GUI was added to make the application more user-friendly.

3. Literature Survey

Gunho Jung and Sun-Yong Choi(Jung & Choi, 2021) have proposed a hybrid model that combines the LSTM and autoencoder models to accurately predict the foreign exchange volatility. They employed the Foreign Exchange Volatility Index (FXVIX) as a measure of foreign exchange volatility. In particular, the three major FXVIX indices (EUVIX, BPVIX, and JYVIX) from 2010 to 2019 are considered, and they predict future prices using the proposed hybrid model.

In this paper, we found that the Long Short-Term Memory (LSTM) is known to perform well in time-series prediction for forecasting FX volatility. LSTM is a type of recurrent neural network capable of learning order dependence in sequence prediction problems, and it has the mechanism of “gate” in order to eliminate the vanishing gradient and exploding gradient which makes it more advanced compared to the traditional RCNN.

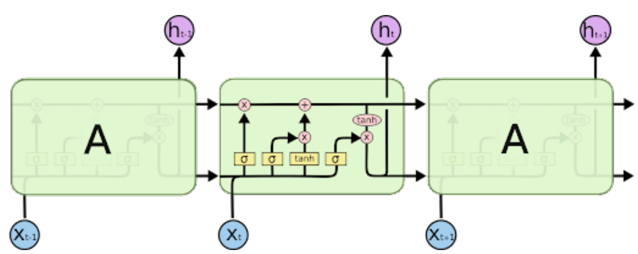

Figure 1. Schematic diagram of lstm algorithm.

In RNN models, usually we use the relation f: h’, y=f(h,x) where h’ and h are the input and output of the hidden layer, and y and x are the input and output of the nodes. In general, we connect nodes in series and pass information between nodes by h and h’. This model is effective for short sequences, but the problem of gradient disappearance and gradient explosion usually occurs in long sequence examples. Therefore the LSTM model will be relatively more suitable.

Compared with the single transmission state h in RNN, LSTM has two transmission states, c and h. c in LSTM is h in RNN. First there is a sigmoid activation function to convert the input x and h into a gate signal. This gate is used to control whether the input from the previous node is forgotten or not. The input of the previous node is added to the signal of this node after choosing whether to forget it or not. And the signal of this node is also selectively remembered by a sigmoid activation function. The c obtained in the previous stage is also scaled by a tanh activation function. Finally, the output to the next node is also controlled by a sigmoid function.

The above is the internal structure of LSTM. It controls the transmission state by gating the state, remembers the information that needs to be remembered for a long time, and forgets the unimportant information. Unlike the ordinary RNN, which can only have one kind of memory overlay in a single way. This is especially useful for many tasks that require "long-term memory".

Haoyi Zhou (Zhou et al., 2021) have designed an efficient transformer-based model for long sequence time-series forecasting, named Informer. This model addresses the issues such as quadratic time complexity, high memory usage, and inherent limitation of the encoder-decoder architecture with three distinctive char-acteristics: (i) a ProbSparse Self-attention mechanism, which achieves O(Llog L) in time complexity and memory usage and has comparable performance on sequences' dependency alignment. (ii) the self-attention distilling highlights dominating attention by halving cascading layer input, and efficiently handles extreme long input sequences. (iii) the generative style decoder, while conceptually simple, predicts the long time-series sequences at one forward operation rather than a step-by-step way, which drastically improves the inference speed of long-sequence predictions. Extensive experiments on four large-scale datasets demonstrate that Informer significantly outperforms existing methods and provides a new solution to the LSTF problem.

In the paper "Stock Market and Foreign Exchange Market in India: Are they Related?"(Mishra, 2004), the researchers attempt to examine whether stock market and foreign exchange markets are related to each other or not. On monthly stock returns, currency rates, interest rates, and money demand for the period April 1992 to March 2002, the study applies Granger's Causality test and Vector Auto Regression approach. The major findings of the study is that there exists a unidirectional causality between the exchange rate and interest rate and between the exchange rate return and demand for money.

4. Datasets

Our main data are the exchange rates of national currencies to the US dollar. We have collected exchange rate data for each day of the last ten years from https://fred.stlouisfed.org/. This is from the Federal Reserve Economic Data, which we believe is a more reliable source of data. Machine learning does not accomplish accurate learning predictions with exchange rates alone. This is because the trend of exchange rates is related to many factors which are explained in the Theoretic Foundation section. After reading and analyzing some aforementioned relevant economics papers, we decided to pick the following data as auxiliary data for training.

Gross Domestic Product: A country's economic growth is measured by the growth of GDP, which is the fundamental determinant of long-term exchange rate changes(Jayachandran, 2013). There is a relationship in the long run, and this is the Balassa–Samuelson effect, where per capita income is high, the real exchange rate is also high, in other words, prices are more expensive in developed countries. When the production efficiency of the traded goods sector (manufacturing) increases rapidly, the growth rate of wages in that sector also increases. Wage levels tend to average out regardless of the domestic industry, so even though the increase in productivity in the non-traded sector (services) is not large, wages in other industries will rise in roughly the same proportion. This causes an increase in the relative price of non-traded products to traded products. Assuming that the price level of traded products (in foreign exchange terms) is a certain level, this change in relative prices will cause an increase in the price of non-traded products under a fixed exchange rate, which in turn will cause an increase in the overall price level (weighted average of traded and non-traded products). If a floating exchange rate is adopted in order to stabilize domestic prices, it will cause an increase in the exchange rate. In either case, the real exchange rate will rise.

Interest Rate: Interest rate changes have a significant impact on the exchange rate. First, interest rate policy has an impact on the exchange rate by affecting the current account(Hamrita & Trifi, 2011). When interest rates rise, credit tightens, lending decreases, investment and consumption decrease, prices fall, imports are suppressed to some extent, exports are promoted, foreign exchange demand is reduced, foreign exchange supply increases, prompting the foreign exchange rate to fall and the local currency exchange rate to rise. Contrary to the rise in interest rates, when interest rates fall, credit expands, money supply (M2) increases, stimulating investment and consumption, prompting prices to rise, discouraging exports and favoring imports. In this case it will increase the demand for foreign exchange, prompting the foreign exchange rate to rise and the local currency exchange rate to fall. Second, indirectly, by affecting international capital flows have an impact on the exchange rate. When a country's interest rate rises, it will attract international capital inflows, thus increasing the demand for local currency and the supply of foreign exchange, so that the local currency exchange rate rises foreign exchange rate falls. Moreover, an increase in a country's interest rate promotes an increase in international capital inflows and a decrease in capital outflows, resulting in a decrease in the balance of payments deficit and an increase in the local currency exchange rate. Contrary to the rise in interest rates, when interest rates fall, it may lead to international capital outflows, increasing the demand for foreign exchange and reducing the balance of payments surplus, prompting the foreign exchange rate to rise and the local currency exchange rate to fall.

Infaltion: The relationship between inflation and exchange rates is very close, but very complex. However, we still have a simple way to understand the relationship between inflation and exchange rates(Odusola & Akinlo, 2001). Since the short-term debt cycle is so important for inflation, and the sharp expansion of debt triggers a large amount of money creation, bringing about a large increase in the money stock, and exchange rate fluctuations, and either a large purchase of foreign currency by domestic currency or a large purchase of domestic currency by foreign currency, this implies that the credit short cycle should have an extremely important nexus role in understanding the inflation-exchange rate relationship. Based on the above micro case, we can speculate that, assuming other things being equal, a domestic credit expansion will trigger domestic inflation and lead to a depreciation of the domestic exchange rate; a foreign credit expansion will trigger foreign inflation and lead to an appreciation of the domestic RMB. Because of the liquidity of tradables, inflation is then globally consistent in the tradable goods sector.

Export and Import: The rise or fall of the exchange rate has a significant impact on the import and export of goods(Nguyen & Do, 2020). The depreciation of the local currency refers to the lower exchange rate of the local currency, i.e., the depreciation of the ratio of the local currency to the foreign currency, which can play the role of promoting exports and discouraging imports. After the devaluation of the local currency, if you export the same goods and receive the same amount of dollars, you can exchange more local currency from the bank and make more profit, so it is beneficial to the export; while the import is just the opposite, in order to import goods, you must first exchange the local currency to the bank for dollars and then buy the goods from abroad. After the devaluation of the local currency, the same amount of local currency will be exchanged for less dollars, and the amount of goods purchased will be smaller, so the cost of importing will increase, so it is not good for importing. In the case of appreciation of the local currency, if the exchange rate of the local currency rises, i.e., the ratio of the local currency to the foreign currency rises, the result is exactly the opposite of the above situation, which is favorable to imports and unfavorable to exports.

Because these data have small daily variations, we have chosen to train on longer-term data. For GDP, we chose annual data, and for interest rates, inflation, and imports and exports, we chose monthly data.

5. Environment

Our model was built based on Pytorch, and we used the Google colab platform which was equipped with both CPU and GPU for the model training.

6. Model

In this work, we determined to use the LSTM model for exchange rate prediction. We developed a prototype model at this time, which combined the previous exchange rate and interest rate as input (which means in the prototype model the input size is 5), and then we defined the number of recurrent layers as 2 and the number of features in the hidden state is 128. We applied the MSE (Mean Squared Error) as the loss function and Adam optimizer to optimize to obtain a convergence model. The structure of the model is shown in Figure 2

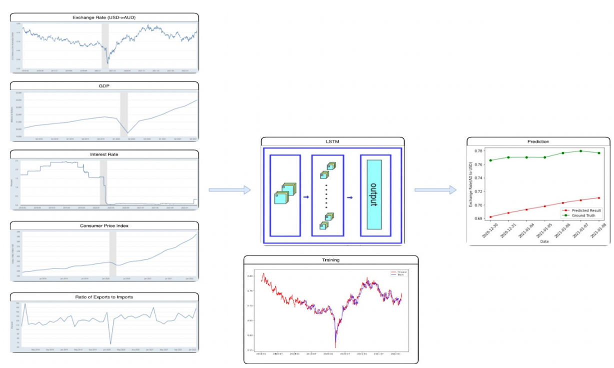

Figure 2. Schematic diagram: shows the input data, training process, and output.

7. Performance Evaluation

After 2500 epochs of training, our model achieved convergence. The mean square error (MSE) of our model fluctuated at 0.02. The result in the training process over the validation set is shown in Figure 3.

Nextly, we applied our model to do the prediction by using

Figure 3. The schematic diagram shows the training process. The red curve is the input data and the blue curve is the training result data.

real-world data in the testing set, in which we tried to use the past 30 days data to predict the next 7 days data. The result is shown in Figure 4. The MSE of the testing process is 0.020. And the comparison between the model produced prediction and the ground truth is shown in Table 1.

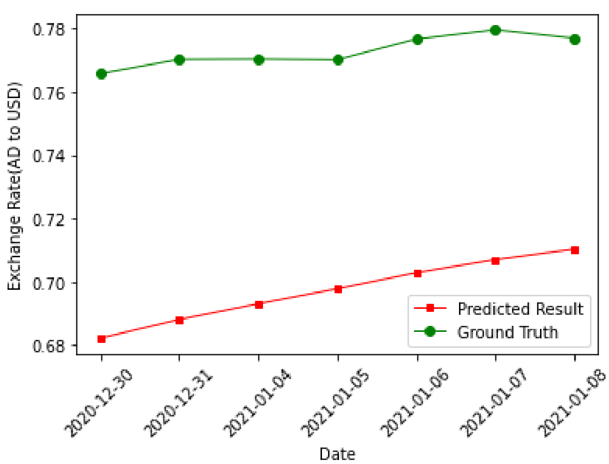

Figure 4. The schematic diagram shows the prediction result. The green curve is the real data and the red curve is the prediction result data.

In the end, we conducted a test and calculated the MSE of the next 100 days (since the last day of the input data) prediction. The results are shown in Figure 5 and Figure 6, from which we can clearly see that the MSE trend fluctuates within the range of 0.01 0.04. We believe that our model achieved a relatively high accuracy.

Table 1. Comparison Between Model Predicted Result and Ground Truth.

MODEL PREDICTED RESULT	GROUND TRUTH
0.6822	0.7657
0.6881	0.7702
0.6931	0.7703
0.6978	0.7701
0.7029	0.7766
0.7070	0.7795
0.7103	0.7769

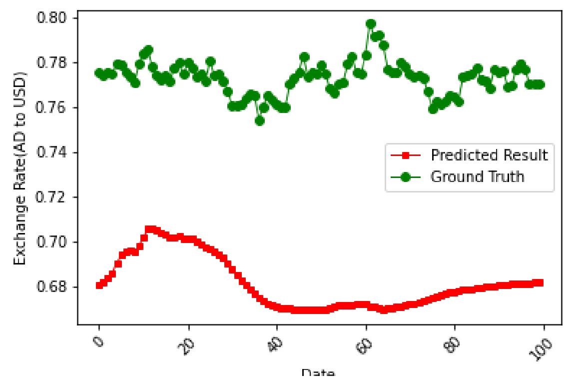

Figure 5. The schematic diagram shows the prediction result of 100 days. The green curve is the real data and the red curve is the prediction result data.

8. Optimization

After the aforementioned training process of 2500 epochs, we still want optimize our prototyped model. Through studying the literatures, we decided to modify the "num_layers" and "hidden_size" parameters in our model.

8.1. Increase the Number of the Layers

The "num_layers" indicated the number of the layers in the model, which also refer to the 'depth' of the neuron network model. And by adding the layer, we actually developed a stacked LSTM. An LSTM model with numerous LSTM layers is known as a stacked LSTM architecture. An LSTM layer above sends a series of values to the LSTM layer below, rather than a single value. One output per input time step, as opposed to one output time step for all input time steps. The architecture of stacked LSTM model is shown in Figure 7.

The reason why we want to increase the number of the layers is because that in paper 'Speech Recognition with Deep Re-

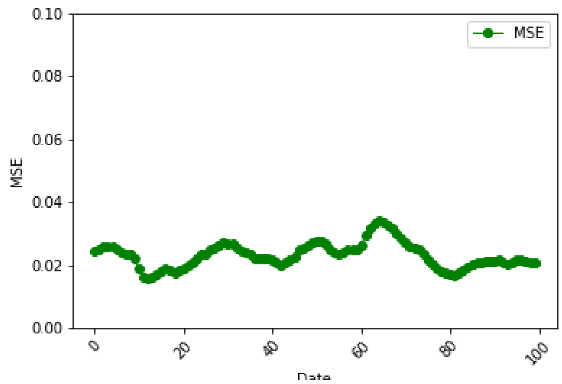

Figure 6. The schematic diagram shows MSE Trend of Prediction.

curent Neural Network'(Graves et al., 2013) the researchers pointed out that stacked LSTM layers is more effective than adding memory cells inside LSTM model. And generally speaking, the addition of layers increases the depth of the model which means increases the level of abstraction of input observations over time as well. Besides, we also noticed that stacked LSTM improves training efficiency and accuracy by adding network depth.

8.2. Increase the Hidden Size in each layers

he "hidden_size" indicated the number of the hidden states inside each layers, which is illustrated in Figure 8 at the position where the red arrow point to.

In most of the cases, the larger hidden size is, the model complexity may increase and over-fitting phenomenon may occur, but the improvement to a certain extent can improve the model accuracy. Meanwhile, more neurons means we can project our data into a higher dimensional space and learn more parameters for more 'connections'. However, there is no exact rule for how to set a specific value of this parameter, which means the problem of how to tuning this parameter still remains unsolved. Therefore, this kind of optimization is somehow purely depends on experience and experiments. But commonly, if we have more training examples we are supposed to add more multiple hidden units. After several preliminary experiment, we finally decided to set the 'hidden_size' in our model to 256.

8.3. Optimized Model Evaluation

In order to evaluate the performance of our optimized model to see if there is ant improvement. we conducted a testing just as same as that mentioned in Section 7. We applied our model to do the prediction by using real-world data in

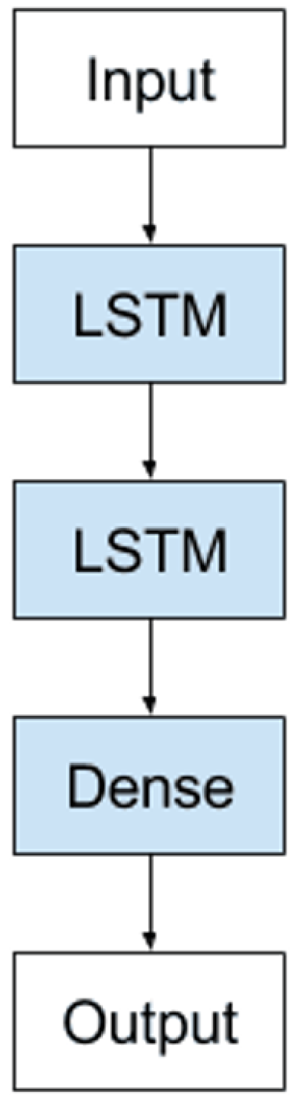

Figure 7. The Architecture of Stacked LSTM Model

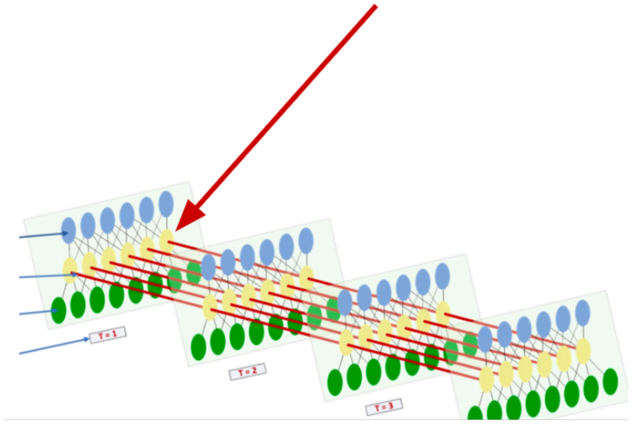

Figure 8. The hidden state in LSTM Model

the testing set, in which we tried to use the past 30 days data to predict the next 7 days data and drew the figure and calculate the MSE.

The comparison between the predicted result and ground truth are shown in Table 2. And the MSE of the testing process is 0.006 which is three times better than before. From the Figure 9, we can easily draw the conclusion that the performance of the optimized model is greatly improved in a qualitative way and in the red circle we can clearly see that the improved model can better reflect the fluctuation trend of the exchange rate.

Table 2. Comparison Between Optimized Model Predicted Result and Ground Truth.

OPTIMIZED MODEL PREDICTED RESULT	GROUND TRUTH
0.7143	0.7657
0.7181	0.7702
0.7235	0.7703
0.7306	0.7701
0.7373	0.7766
0.7399	0.7795
0.7386	0.7769

9. Graphic User Interface

We have realized that our foreign exchange prediction system may aim at non-professional users such as traders in financial markets without technical background. Therefore we developed a graphic user interface (GUI) for our system in order to make those users accessible to our application. Figure 10 shows the main interface of our system, in which the user can choose the functions which means they can choose the type of exchange rate. And after choosing, the

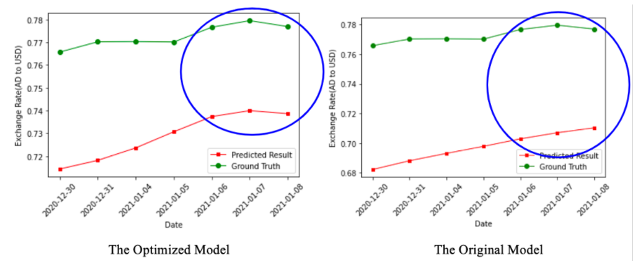

Figure 9. The Comparison between Optimized Model and Original Model

user should click the button to upload the past days' data (currently CSV file only) for the system to do the prediction, which is shown in Figure 11.

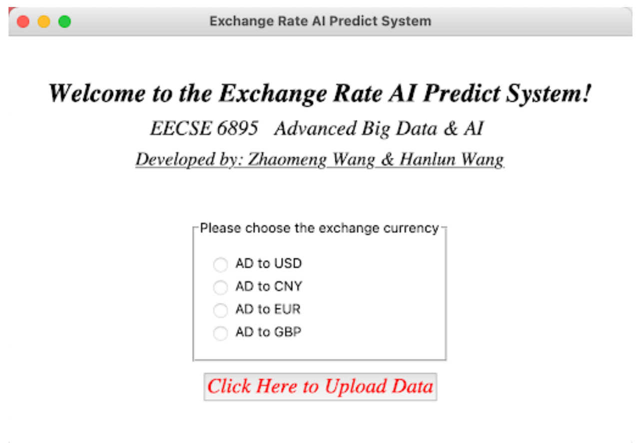

Figure 10. Main Interface

When the system receives the file and finishes processing, the prediction result will be displayed. In the new window, the figure of the prediction data will be shown which helps the user to form a direct and clear view of the trend of exchange rate change. And in the below area, the specific data will be shown in a table which gives the users a quantitative result. This window is shown in Figure 12.

References

Graves, A., Mohamed, A.-r., and Hinton, G. Speech recognition with deep recurrent neural networks. In *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 6645–6649. Ieee, 2013.

Hamrita, M. E. and Trifi, A. The relationship between interest rate, exchange rate and stock price: A wavelet analysis. *International Journal of Economics and Financial Issues*, 1(4):220–228, 2011.

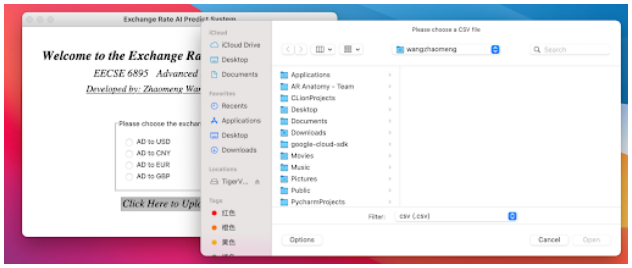

Figure 11. Upload File

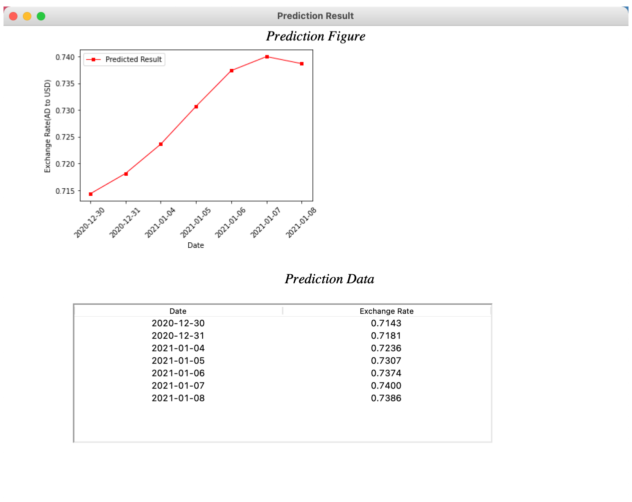

Figure 12. Prediction Result Window.

Jayachandran, G. Impact of exchange rate on trade and gdp for india: A study of last four decade. *International journal of marketing, financial services & management research*, 2(9):154–170, 2013.

Jung, G. and Choi, S.-Y. Forecasting foreign exchange volatility using deep learning autoencoder-lstm techniques. *Complexity*, 2021, 2021.

Mishra, A. K. Stock market and foreign exchange market in india: are they related? *South Asia Economic Journal*, 5(2):209–232, 2004.

Nguyen, V. C. and Do, T. T. Impact of exchange rate shocks, inward fdi and import on export performance: a cointegration analysis. *The Journal of Asian Finance, Economics and Business*, 7(4):163–171, 2020.

Odusola, A. and Akinlo, A. Output, inflation, and exchange rate in developing countries: An application to nigeria. *The Developing Economies*, 2001.

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., and Zhang, W. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of AAAI*, 2021.