Deep learning with encoder-decoder architecture for exchange currency rates model predictions

M F Andrijasa*
Department of Information Technology State Polytechnic of Samarinda, Samarinda, Indonesia

*andrijasa@gmail.com

Abstract. Exchange currency rate becomes one of the most important things on country economic growth. In some countries, the rate may affect seriously to economic growth and political stability. Governments need to have some actions in order to stabilize the rate. Knowing the pattern of exchange currency rate is a mandatory by governments for determining their future policies or as consideration in future decision-making. Therefore, government needs an analysis tool for modelling the rate prediction. Our previous study emerged that using deep learning method with encoder-decoder architecture has succeeded to model the rate prediction with teacher forcing and give only a one-step prediction. In this study, the authors aim to predict the pattern of exchange rate using the previous method with zero input to decoder to obtain multiple time steps prediction. After training and testing over 4,344 daily data series of IDR to USD exchange currency rate over 17 years taken from Bank Indonesia official website, the result found that the method still showed an interesting model to predict the next 40 sequence data series of exchange rate. In conclusion, the proposed method can be utilized to tackle the others prediction problem and the resulted model can be applied as an analysis tool by government for predicting their rate for the coming months or years.

1. Introduction
Exchange currency rate may affect to economic stability of most countries. In developing countries, the rate has an effect to their economic growth. This circumstance can avoid some investors to invest their capital. Most of financial transaction with foreigners use USD as a payment instrument and it makes USD rates outperform over the others currency rates [1]. Finally, this can seriously affect to inflation level, politic and security instability. Government need to analyse this issue accurately and make some policies in order to keep monetary and economic stabilities safe. Therefore, government need some analysis tools such as fundamental analysis and technical analysis [2] that can be utilized for predicting the rates using data in the past to forecast the future rates [3] and exploit time series data in the past as the basis for decision making [4].

In machine learning era, there are a lot of methods that can be applied for predicting time series data. AutoRegressive Integrated Moving Average (ARIMA) models has become popular for predicting [5] as well as its variant like ARIMAX [6]. Deep learning as the newer method is used for prediction time series data recently which has more complex architecture such as Multi-Layer Perceptron (MLPs) or Recurrent Neural Network (RNNs) [7]. Sequence-to-sequence (seq2seq) model which is as RNNs variant outperform ARIMA model for predicting Bitcoin USD [8].
Based on the above-background and their findings, the authors proposed to use encoder-decoder RNN model for predicting exchange currency rates for the next 40 days using past time series data daily. This article will explain as follows:

2. Method
RNNs can feed their network model with sequential data and process the sequences with memory buffer capability. This capability is served by internal state $h_t$ to hold the previous data stay on the network and can be reused by another neuron on the same network. An encoder-decoder (seq2seq) RNN uses two specialized cells. One for handling past data series and memorizing the important event (encoder) and one for predicting the future value using encoder internal output and decoder input (decoder). In press used $\{x_1, \ldots, x_n\}$ as input sequence on encoder and $\{y_1, \ldots, y_n\}$ as target of decoder output with decoder input fed by sequence of shifted by 1 from decoder output [9]. In this paper, we employ similar model with all decoder input set to zero as shown in figure 1. It can force the network model working harder and memorizing longer.

3. Experimental evaluation

3.1. Data
The data was extracted from Bank Indonesia official websites. The data that consists of three features, date, buy and sell cover the time period between January 24th, 2001 and October 17th, 2018. In pre-processing data, we employ normalization.

3.2. Network hyper-parameter optimization
We use an encoder-decoder RNN with utilizing gated recurrent unit (GRU) cells. That model was created on Keras with Tensorflow backend and based on the previous study. Manual selection of network hyper-parameter optimization was still applied in order to obtain best model based on literature suggestion and intuition. Input and output dimensions were set to 40 whereas target dimension was set to 80; the batch size was set to 512; the number of GRU cell that is used by RNN was 128 neurons with 2 stacked recurrent cells. Furthermore, learning rate was set to its default value of 0.007 with 1000 training iteration and decay and momentum to 0.9 and 0.5 in turn. At last, we employed RMSProp optimizer to optimize the network [10].

![Figure 1. Encoder-decoder model architecture.](image-url)
3.3. Empirical results
As can be seen in Figure 2, the model result RMSE value for training and validation for 14 epochs. Training RMSE decrease significantly to epoch 2 and start decrease slowly from epoch 4 to the end epoch period. The training graph illustrate that the model improves during training. On the other hand, validation RMSE had a fluctuation by just over 2.0 during the epoch time period. The validation graph show that the model does not improve using validation data. Literature speaking, to decrease validation loss we can do by tweaking or manual selection of hyper-parameter especially by increasing the number of neurons and stacked. This will make training very expensive and time consuming.

Interestingly, as shown in Figure 3, the model can predict the next 40 days of data time series using the trained model. The prediction graph show that there is a similar pattern of rates between day 40 and day 80 and followed by the graph pattern for the next 40 days as prediction graph.

4. Discussion
At this point, we can assume that the model will perform better or as the best model with additional related data input in decoder input such as social data source, news, and sentiment analysis. Moreover, proposed model architecture and its hyper-parameter can be fined tune in order to obtain the best model, such as cell selection (GRU or LSTM).
5. Conclusions
Although decoder input was set to zero, the encoder-decoder model still results the prediction pattern for targeted days. The proposed method can be utilized to tackle the others prediction problem and the resulted model can be applied as an analysis tool by government for predicting their rate for the coming months or years.

References
[1] Ernita D, Amar S and Syofyan E 2013 Analisis Pertumbuhan Ekonomi, Investasi, dan Konsumsi di Indonesia Jurnal Kajian Ekonomi 12
[2] Luca C 1995 Tranding in the Global Currency Markets (New Jersey: Prentice Hall)
[3] Gaspersz 2002 Ekonomi Manajerial, Pembuatan Keputusan Bisnis (edisi revisi dan perluasan) (Jakarta: Gramedia Pustaka Utama)
[4] Hanke J. and Wichern D 2015 Business forecasting Pearson
[5] Ariyo A A, Adewumi A O and Ayo C K 2014 Stock price prediction using the ARIMA model In Computer Modelling and Simulation (UKSim), 2014 UKSim-AMSS 16th International Conference on 106-112 IEEE
[6] Kongcharoen C and Kruangpradit T 2013 Autoregressive integrated moving average with explanatory variable (ARIMAX) model for Thailand export In 33rd International Symposium on Forecasting, South Korea 1-8
[7] Sutskever I, Vinyals O and Le Q V 2014 Sequence to sequence learning with neural networks In Advances in neural information processing systems 3104-3112
[8] Rebane J, Karlsson I, Denic S and Papapetrou P 2018 Seq2Seq RNNs and ARIMA models for Cryptocurrency Prediction: A Comparative Study
[9] Andrijasa M F, Hidayat H and Sari W E 2018 Prediksi Nilai Tukar Mata Uang IDR Terhadap USD Dengan Teknik Deep Learning Menggunakan Model Recurrent Neural Network
[10] Riedmiller M and Braun H 1992 RPROP-A fast adaptive learning algorithm In Proc. of ISCIS VII Universitat II