MEHMET AKİF ERSOY ÜNİVERSİTESİ
İKTİSADİ VE İDARİ BİLİMLER FAKÜLTESİ DERGİSİ
Mehmet Akif Ersoy University
Journal of Economics and Administrative Sciences Faculty

ISSN: 2149-1658
Cilt: 10 Sayı: 2 s.935-949
Volume: 10 Issue: 2 p.935-949
Temmuz 2023 July

PREDICTING USD/ TL EXCHANGE RATE IN TURKEY: THE LONG-SHORT TERM MEMORY APPROACH

TÜRKİYE'DE DOLAR/TL KURUNU TAHMİN ETMEK: UZUN-KISA BELLEK SİNİR AĞLARI YAKLAŞIMI

Ayten YAĞMUR[1], Zeynep KARAÇOR[2], Fatih MANGIR[3], Abdul-Razak B. YUSSIF[4]

1. Dr. Öğr. Üyesi, Akdeniz Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Çalışma Ekonomisi ve Endüstri İlişkiler, aytenyagmur@akdeniz.edu.tr, https://orcid.org/0000-0003-2138-240X
2. Prof. Dr., Selçuk Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat Bölümü, zkaracor@selcuk.edu.tr, https://orcid.org/0000-0003-2050-644X
3. Prof. Dr., Selçuk Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat Bölümü, fmangir@selcuk.edu.tr, https://orcid.org/0000-0003-1348-7818
4. Doktora Öğrenci, Manitoba Universitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat Bölümü, yussifa@myumanitoba.ca, https://orcid.org/0000-0002-1930-0287

Makale Türü	Article Type
Araştırma Makalesi	Research Article
Başvuru Tarihi	**Application Date**
02.04.2022	04.02.2022
Yayına Kabul Tarihi	**Admission Date**
01.12.2023	12.01.2023

DOI

https://doi.org/10.30798/makuiibf.1097568

Abstract

The prediction of the exchange rate time series has been quite challenging but is an essential process. This is as a result of the inherent noise and the volatile behavior in these series. Time series analysis models such as ARIMA have been used for this purpose. However, these models are limited due to the fact that they are not able to explain the non-linearity as well as the stochastic properties of foreign exchange rates. In order to perform a more accurate exchange rate prediction, deep-learning methods have been employed withremarkable rates of success. In this paper, we apply the Long-Short Term Memory Neural Network to predict the USD/TL exchange rate in Turkey. The result from this paper indicates that the Long-Short Term Memory Neural Network deep learning method gives higher prediction accuracy compared to the Auto Regressive Integrated Moving Average and the Multilayer Perception Neural Network models.

Keywords: Prediction, Exchange Rate, Time Series, ARIMA, LSTM, MLP.

Öz

Döviz kuru zaman serisinin tahmini oldukça zorlu, ancak önemli bir süreçtir. Bu, serilerdeki kalıtsal gürültü özelliğinin ve kırılgan davranışının sonucudur. Bu amaçla ARIMA gibi zaman serisi analiz modelleri kullanılmıştır. Ancak bu modeller döviz kurlarının stokastik özelliklerinin yanı sıra doğrusal olmama özelliklerini de açıklayamamaları nedeniyle sınırlıdırlar. Daha doğru bir döviz kuru tahmini gerçekleştirmek için, önemli başarı oranlarına sahip derin öğrenme yöntemleri uygulanmaktadır. Bu çalışma da, Türkiye'deki USD/TL kurunu tahmin etmek için Uzun-Kısa Vadeli Bellek Sinir Ağı yöntemi uygulanmaktadır. Bu makaleden elde edilen sonuç, Uzun-Kısa Süreli Bellek Sinir Ağı derin öğrenme yönteminin otoregresif hareketli ortalamalar yöntemi ile Çok katmanlı Yapay Sinir Ağı modellerine kıyasla daha yüksek tahmin yapmaktadır. .

Anahtar Kelimeler: Tahmin, Döviz Kuru, Zaman Serileri, ARIMA, LSTM, MLP.

GENİŞLETİLMİŞ ÖZET

Çalışmanın Amacı

Bu makale makroekonomik göstergeler üzerinde en temel belirleyicilerden biri olan döviz kuru tahminlerini kavramsal ve ampirik olarak incelemektir. Bu amaçla kur tahminini en doğru gerçekleştiren modeli tespit edebilmek adına klasik lineer ve lineer olmayan klasik yapay sinir ağları yöntemleriyle daha çok katman ve uzun hafıza ile çalışan Yığılmış Uzun kısa süreli bellek (LSTM) yönteminin karşılaştırmalı analizi yapılmış ve en doğru sonucu veren yöntem ortaya konmuştur. Elde edilen ampirik sonuçlar hem yatırımcılar hem konu üzerinde çalışanlar için döviz kuru tahmini konusunda literatüre katkı sağlanacaktır.

Araştırma Soruları

Bu çalışmanın cevaplamaya çalıştığı araştırma soruları; - Kur oynaklıkları yatırımcılar için ne anlama gelmektedir2- Kur tahmin yöntemlerinden zaman serilerinin kullanımı sonuçları açısından ne kadar etkindir?- Kur tahmin yöntemlerinden birbirine karşı olan üstün ve zayıf yönleri nelerdir?- ARIMA ve Yapay zeka modellerinin sonuçlarından hangisi daha doğru sonuç vermektedir?

Literatür Araştırması

Chandwani Deepika ve Manminder S. Saluja (2014), Okash K. Mahmoud (2014), Pallabi ve Kumari (2017), and Sujin vd. (2017) yapay zekâ modellerinin tahminlerde diğer modellere göre daha başarılı sonuçlar verdiğini ortaya koyan seçilmiş çalışmalardır. Literatürde döviz kurunun tahminini yapay sinir ağları ile inceleyen diğer çalışmalardan en önemlileri ise Greff et al (2016), Azzouni and Pujolle (2017), Kong Yun-Long vd. (2018), LiYi Fei, vd. (2018) tarafından gerçekleştirilmiş olup, Türkiye'de bu konuda çalışmalarda kurun yapay sinir ağları ve klasik zaman serileri karşılaştırmaları üzerinde analiz eden çalışmalarda genel ortak bir kanı oluşmamıştır. Literatürde döviz kurunun tahminini yapay sinir ağları ile inceleyen temel çalışmalardan en önemlileri Greff et al (2016), Azzouni and Pujolle (2017), Kong Yun-Long vd. (2018), LiYi Fei, vd. (2018) tarafından gerçekleştirilmiş olup, Türkiye'de bu konuda çalışmalarda kurun yapay sinir ağları ve klasik zaman serileri karşılaştırmaları üzerinde analiz eden çalışmalarda genel ortak bir kanı oluşmamıştır.

Yöntem

Literatürde, döviz kuruna ait zaman serilerinin tahmini için otoregresif hareketli ortalamalar (autoregressive integrated moving average-ARIMA) modeli ve yapay sinir ağları (NNP) modelleriyle tahminleri kullanan çalışmalara oldukça sık rastlanmaktadır. Ancak bu yöntemler hem kısa hafızalı olduğu hem de kur serisindeki doğrusal olmayan ilişki sorununu çözemediği için eleştirilmektedirler. Bu yüzden, ARIMA gibi doğrusal serilerle analiz yapan ve Yapay sinir ağları gibi kısa hafızalı modellerin yanında farklı yöntemlere gereksinim duyulmaktadır. Sepp Hochreiter ve Jurgen Schmidhuber'a göre, uzun kısa süreli bellek hafıza modelleri (LSTM) ilk olarak 1997 yılında tekrarlayan yapay sinir ağları (RNN)nı karakterize eden akış hata terimleri analizinden esinlenerek geliştirilmiştir. Bu hata terimleri, mevcut modellerin uzun süre gecikmeler nedeniyle sonuçlarında ortaya çıkan değerlerin doğruluğunu analiz edilme ihtiyacını ortaya koymuştur. LSTM'lerin klasik NN'lerin zaman serisindeki gecikmeli ve aralıklı olayları işleme ve tahmin etme yeteneği göz önüne alınarak finansal serilerin tahmini için uygun hale getirilmiştir. Bu çalışmada Türkiye Cumhuriyet Merkez Bankasında 2001 ve 2020 yılına ait nominal TL/ ABD dolar serisini tahmin etmek için ARIMA (Autoregressive Integrated Moving Average), çok katmanlı yapay sinir ağları (MLP) ve Uzun kısa vadeli hafıza ağları (LSTM) modeller seçilerek kur tahmin sonuçları karşılaştırılmıştır.

Sonuç ve Değerlendirme

Tüm ekonomik işlemlerin finansal karşılıkları, döviz kurlarının gelecekteki değerinden doğrudan etkilenmektedir; bu yüzden, döviz kuru hareketlerini yüksek oranda doğru tahmin etmek tüm piyasa oyuncuların yatırım kararları için çok kritiktir. Döviz kurlarının tahmin etmesi, ekonomik ve siyasi olanlar da dahil olmak üzere çeşitli faktörlere bağlı, zorlu bir görev olduğu için döviz kurlarının doğru tahmin eden ekonometrik modellerin tespiti ekonomik aktörlerin ekonomiye daha fazla yatırım yapmalarına neden olacaktır. Bu çalışma, döviz kuru tahmini için makine öğretisi yöntemleri ve klasik istatistiksel modellerin karşılaştırılmasına odaklanmıştır. Temel hedef, Dolar / Türk Lirası değerini tahmin etmek için en iyi tahmin yöntemini tespit etmektir. Bu amaçla, mevcut tarihsel verilere dayanarak Dolar/ TL serisini tahmin etmek için ARIMA, Çok Katmanlı Algılayıcı (MLP) ve Uzun Kısa Süreli Bellek (LSTM) olmak üzere üç tür istatistiksel yöntem kullanılmıştır. Bu verileri eğitim modelleriyle beslemeden önce, değerlerin modeller için daha uygun hale getirmek için veri kümesine ön işlem adımları uygulanmıştır. Bu ön işleme adımları, verilerin temizlenmesini, verilerin denetlenen bir öğrenme sorununa dönüştürülmesini ve farklılaşma uygulamalarını içermektedir. Elde ettiğimiz sonuca göre, ARIMA ve MLP sinir ağıyla karşılaştırıldığında, Yığılmış LSTM'nin döviz kurlarını tahmin etmek için gösterdiği doğruluğun daha başarılı olduğu ve bunları daha karmaşık hibrit derin öğrenme tekniklerine genişletmeye odaklanacak gelecekteki çalışmalara ilham verdiğidir. Ayrıca, gelecekteki çalışmalarda, modellerin en uygun zaman gecikmelerini otomatik olarak belirlemesine yardımcı olmak için değiştirilebilir zaman sırası giriş uzunluğuna sahip LSTM modelleri kullanılabilir.

1. INTRODUCTION

Economic actors such as investors involved in international business transactions are mostly faced with currency risk due to volatilities in the foreign exchange markets and the unexpected movements of exchange rates. Accurate forecasting of the exchange rate would therefore play a significant role in macroeconomic analysis for stability and economic growth. This goal has recently become popular among empirical analysis of economic studies.

In an integrated economic and financial markets, exchange rates play a significant role in macroeconomic policies towards ensuring economic goals focusing on the growth. Stable exchange rate would determine healthy imports and exports activities thereby achieving sustainable trade balance of the country.

Exchange rates are characterized with inherently noisy, non-stationary and deterministically chaotic tendencies which suggest that information about past behavior of such markets there is no completely capture the dependencies between future and past rates. Therefore, constructing rebuts exchange rate prediction is a lucrative factor for the future income of many businesses and fund managers.

Business environments rely on financial data to develop plans for future trade. A critical consequence of inaccurate forecast of financial indicators is that, it makes it harder for investors to make good decisions. Inaccurate forecast problems may grow catastrophic, steering to business collapse and become a serious problem even if they survive.

With the rapid integration of financial markets, the accurate prediction of systems is difficult mainly due to the chaotic conditions of these markets coupled with the fact that market features are quite volatile and non-linear. For several years, a variety of econometric analysis have been applied for analyzing and forecasting financial time series data.

These methods range from the classical time series forecasting methods such as the Auto Regressive Integrated Moving Average (ARIMA), Auto Regressive Moving Average (ARMA), Moving Average (MA), Auto Regressive (AR) as well as exponential smoothing. However, noisy, non-linear, complex, dynamic, nonparametric, and chaotic form of financial data are not correctly captured by these classical statistical methods. Most classical statistical models for time series forecasting have to assume that the time series is linear (Kleiner, 1977). However, the presence of non-linearity and noisy features in real world data yields some forecasting problems, which will result in inaccuracy of figures and unreliable results.

In Turkey, following the fast pace of development of the financial markets as well as the huge impact of both the domestic and international policies on key financial variables, the country has witnessed some exchange rates fluctuations. Given the essential and significant role that the financial industry plays and contributes to the economy, the exigence of making accurate exchange rate forecast cannot be over emphasized.

The accurate predictions of exchange rates are important to enable effective planning and policies settings including monetary policies and fiscal policies that will enhance currency performance in the international markets as well as support the economy. Unfortunately, however, the non-linear and stochastic features that characterize exchange rates do not allow for accurate predictions using the existing linear models.

In recent studies, deep learning neural networks have received great attention both in the industry and the academic world (LeCun, Bengio, and Hinton, 2015). The introduction of deep learning methods in recent decades has provided serious competition to the classical statistical models used for time series forecasting. Some of the commonly used models include artificial neural networks (ANN) and support vector machines (SVM).

These models are widely used due to their ability to successfully perform classification and prediction tasks on financial time series data by learning and modeling the nonlinearity in such data. In Chandwani Deepika and Manminder S. Saluja (2014), Okash K. Mahmoud (2014), Pallabi, and Kumari (2017), and Sujin et al. (2017) the high forecasting success of machine learning methods are discussed.

The fast pace at which financial forecasting is developing as well as the fact that financial data being a valuable information source for both domestic and international investors and traders calls for continuous studies. In modern financial forecasting studies, researchers largely focus on developing hybrid models. This is aimed at harnessing the combined strength of data mining models, statistical methods, and machine learning algorithms (Khairalla and Al-Jallad, 2017; Osório et al., 2019 and Das, Mishra, and Rout, 2020).

In the literature of exchange rate forecasting, the most commonly explored machine learning techniques include a series of recurrent neural networks like the Long-Short Time Memory (LSTM) used to handle classification and predictions problems in financial data including Greff et al (2016), Azzouni and Pujolle (2017), Kong Yun-Long et al. (2018) and LiYiFei, and Han Cao (2018). In Turkey, the neural network which is most suitable and offers the highest prediction accuracy for exchange rates remains unresolved.

The paper provides a compact view of a specific problem, related to forecasting exchange rate series for the case of Turkey. The originality of this article is that we applied the LSTM neural networks (NN) whıch outperforms both ARIMA and MLP models in Turkey. The Adam optimizer is used for the LSTM model with a learning rate of 0.001.

To establish a baseline for comparing our model, we tested the LSTM models with ARIMA and MLP models on current exchange rates of US Dollar (USD) to Turkish Lira (TRY). The results from this study shows that LSTM neural networks (NN) outperform both ARIMA and MLP models.

The rest chapters of this research paper is structured as follows; methodology, models and data description are included in Section II. The results of the models are shown in chapter III. Finally, chapter IV draws conclusions and suggests future works.

2. METHODOLOGY, MODELS AND DATA

2.1. Methodology

According to Sepp Hochreiter and Jurgen Schmidhuber, LSTM was first proposed in 1997 stemming from error analysis for flow that characterizes RNNs (Hochreiter et al., 2001). This error indicated that the existing architectures were out of reach due to long time lags. Given the ability of LSTM NNs to process and predict delayed and interval events in time series, this makes it suitable for predictions and for that matter prediction of the exchange rate.

2.1.1. Recurrent Neural Networks (The first version of LSTM)

A recurrent neural network (RNN) is an improvement of the multilayer perceptron network (MLP) containing the input, hidden and output layer. The functioning of RNNs occurs on sequences of inputs. The same task is performed on each sequence and the current output depends on the computations from the previous sequence. Unlike Feed-forward neural networks that have a one-to-one vector mapping of input to output, RNNs are capable of mapping from the squence of initial inputs to each output. Thus, RNNs do seq-2-seq mapping.

Theoretically, RNNs are capable of using information from long arbitrary sequences for prediction. In practice however, RNNs are restricted to recall for step-by-step.

(Fig. 1) shows a simple RNN containing one input layer, one recurrent hidden layer that is folded into a full network and one output layer. The input, output and hidden state at time step t are indicated by x_t, o_t, and s_t respectively. s_t is the memory of the network. U, V, W are parameters in different network layers. s_(t) is the memory structure consisting of U, V, W - parameters in different layers of the network.

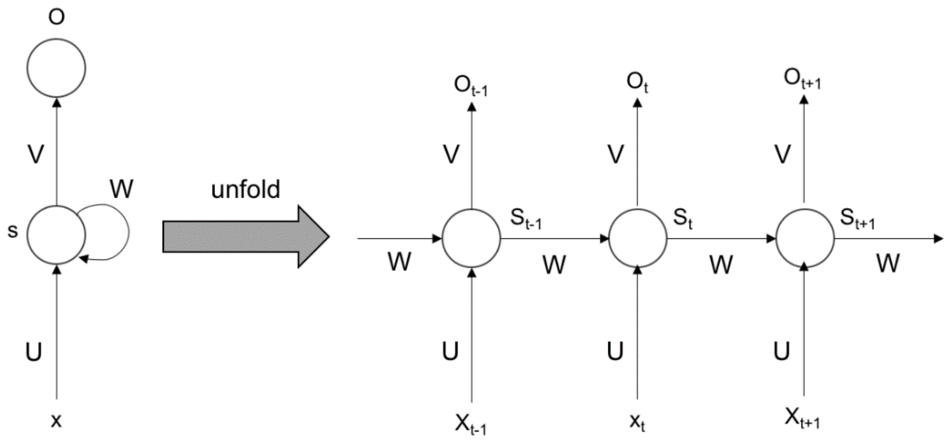
According to Faraway and Chatfield (1998), the traditional neural networks employ different parameters at each layer whereas RNNs parameters are used repeatedly across all steps as shown below. This reduces numbers of all variables that is required to train the model.

In the Forward propagation network, the current layer uses network connection and the weight to calculate the input data as outputs to the next layer. *Sj(t)* refers to the output of the hidden layer at each *t* interval step.

$S_j(t) = f(\sum_i^l x_i(t)v_{jt} + \sum_h^m s_h(t-1)u_{jh} + b_j)$

From the formula above, the input of the input layer at time step t is $\sum_i^l x_i(t)v_{jt}$. The input of the hidden layer at time step (t -1) is given by $\sum_h^m s_h(t-1)u_{jh} + b_j$. Where b_j represents the bias and *f(-)* is a nonlinear map function such as ReLU or tanh.

Figure 1. Unfolded Recurrent Neural Network (RNN)



Source: LeCun, Bengio, and Hinton, 2015

The RNN has a similar forward pass as a single hidden layer of the MLP with the only distinction being that RNN has activations to the hidden layer from two sources, the current external input and the hidden layer process from former intervals of time. Let's consider that an RNN has length *t* and *x* input sequence with I, H and K input units, hidden units and output units respectively.

Considering that the value of input *i* at time *t* be x^t and; a_h^t and b_h^tbe the network input to unit *h* and the processing of unit *h* at time *t* sequentially (Li et al., 2018). The hidden layer will be:

$$a_h^t = \sum_{i=1}^{l} w_{ih} x_i^t + \sum_{h'=1}^{H} w_{h'h} b_{h'}^{t-1} \quad (2)$$

Applying differentiable activation functions, we have:

$$b_h^t = \theta_h(a_h^t) \quad (3)$$

Then the network inputs to the output units is computed by:

$$a_k^t = \sum_{h=1}^{H} w_{hk} b_h^t \quad (4)$$

For RNN, a well-known algorithm, Backpropagation Through Time (BPTT) has been proposed to calculate weight derivatives (Jaeger Herbert, 2002, p.1). The standard BPTT entails repeatedly using the chain rule.

The activation of the hidden layer through its influence on the output layer, as well as its influence on the hidden layer at the following time step, determines the loss function in RNN (Bengio, Simard and Frasconi, 1994). This gives:

$$\delta_h^t = \theta'(a_h^t)\sum_{k=1}^{K}\delta_k^t w_{hk} + \sum_{h'=1}^{H}\delta_{h'}^{t+1} w_{hh'} \quad (5)$$

Where

$$\delta_j^t \stackrel{\text{def}}{=} \frac{\partial L}{\partial a_j^t} \quad (6)$$

The summation of all the sequences is used to obtain the derivatives for the entire network weights:

$$\frac{\partial L}{\partial w_{ij}} = \sum_{t=1}^{T}\frac{\partial L}{\partial a_j^t}\frac{\partial a_j^t}{\partial w_{ij}} = \sum_{t=1}^{T}\delta_j^t b_i^t \quad (7)$$

RNNs experience the problem of vanishing gradients. That is to say, in RNNs the perceptron of nodes in front of time nodes declines with time (Graves, Jaitly and Mohamed, 2013).

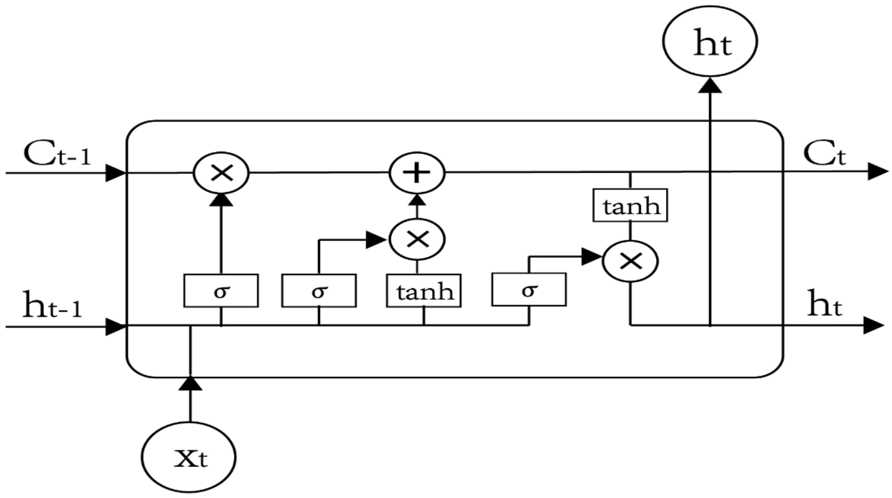
2.1.2. Long-Short Term Memory (LSTM)

Addressing the problem of disappearing gradients in RNNs, the LSTM creates memory cells that can store information across long distances by dividing into three inner-cell gates (Gers, Schmidhuber, and Cummins, 2000).

An LSTM neural network cell typically is made up of mainly four gates: Input, input modulation, forget and output. By accepting initial input data, the input gate performs these incoming data.

The Memory cell input gate process raw data from the output of the LSTM neural network structure in the last operation. The Forget gate choices the ideal moment lag for the data order by deciding when to clean output calculations. The Output gate generates data for the LSTM units by taking all results calculated into consideration.

Figure 2. A Simple Model of Long-Short Term Memory Neural Network Cells



Source: Qiu et al., 2020

In our exchange rate prediction model, the output layer of the LSTM cell consists of a linear regression layer. The hidden state cells is denoted as as H$(h_1, h_2, .., h_r)$ where as the input series and the output sequence are denoted by X = $(x_1, x_2, .., x_r)$ and Y = $y_1, y_2, .., y_r)$ respectively. Thus the LSTM neural networks do the following computations:

$$h_t = H(W_{hy}x_t + W_{hh}h_{t-1} + b_h); \quad \text{for t = 1,…,T} \quad (8)$$

$$y_t = W_{hy}h_t + b_y; \quad \text{for t = 1,…,T} \quad (9)$$

Then implemented as follows:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \text{ ; for t = 1,…,T} \quad (10)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f)\text{; for t = 1,…,T} \quad (11)$$

$$c_t = f_t * c_{t-1} + i_t *tan\ tanh\ (W_{xc}x_t + W_{hc}h_{t-1} + b_c)\text{;for t = 1,…,T} \quad (12)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o); \quad \text{for t = 1,…,T} \quad (13)$$

$$h_t = o_t\ tan\ tanh\ (c_t); \quad \text{for t = 1,…,T} \quad (14)$$

$$\sigma(x) = \frac{1}{1+ e^x} \quad (15)$$

The specific element wise activation functions of the LSTM are σ and tan h (Gers, Schraudolph, and Schmidhuber, 2002). Equation 9 defines the classic sigmoid function defined representing as σ. The sigmoid function describes how much of each component that should be let through and take a range of values between zero and one [0, 1].

When the value is zero, it means ‘let nothing through’ whereas a value of one implies ‘let everything through.’ We represent the input gate by *i*, the forget gate by *f* and lastly the output gate by *o*. The hidden vector *h* needs to be equal to *c*. The weight matrix is denoted by W. The input gate determines how the incoming vectors x_t alter the memory cell state. The forget gate allows the memory cell to either remember or forget its previous state. And finally, the effects of the memory cell on the outputs is carried out through the output gate.

2.2. Data and Model

In this study, the exchange rate time series were generated from the Republic of Turkey Central Bank data bank. The effective selling exchange rate of the Turkish lira to the United States dollar was used. The daily exchange rate values are for the periods January 1, 2001 to August 10, 2020 excluding weekends and public holidays. This consists of 4,930 data points used to train our models.

The data was normalized to the range between -1 and 1 using the Minimax Scalar. In this experiment, we divided the data into training dataset and test dataset. We used 80% of the dataset for training while the remaining 20% is used to test the prediction accuracy of our models. Before feeding this data into the training models, preprocessing steps needed to be applied on the dataset to make the values more appropriate for the models. Preprocessing steps included cleaning the data, converting the data to a supervised learning problem andapplying differencing.

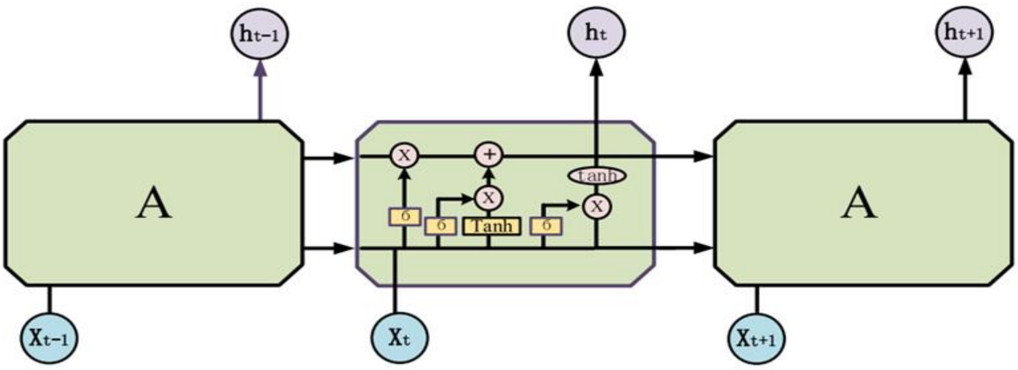
2.2.1. Model Design

The time series forecasting is performed using deep learning methods. In our study, we developed two neural network models for exchange rate series predictive modeling. These are the MLP neural network and Stacked LSTM neural network.

The models are developed by first creating a differenced series of the entire dataset with an interval of one. Thus subtracting the observation at period (t-1) from current period's (t) observation. This is done to make the data set stationary.

Establishing stationarity in univariate time series forecasting impacts model predictability and reliability. In a univariate time series forecast, the data is then transformed into a supervised learning problem. Thus the observation from the previous 5 time steps is used as the input for the current observation.

Figure 3. MLP Neural Network



Source: Hu et all., 2020: 1487.

The MLP model consists of an Input layer and three hidden layers. The models require a single neuron in the output layer with a linear activation function to predict the exchange rate at the next time step. The first hidden layer has 100 neurons while the second and third hidden layers consist of 64 and 32 neurons respectively.

We use the Adam optimizer with a learning rate of 0.0001 and the Mean Square Error (MSE) as the loss function in the model. The activation functions for the three layers including the Input layer is the ReLU function. A batch size of one is used and the model is trained for 1000 epochs. In addition, a dropout layer with a probability of 0.1 was added in order to reduce overfitting.

The LSTM model is made up of an input layer and two stacked LSTM layers, a linear layer and one output layer with a single neuron. The stacked LSTM layers both have 128 neurons and the last hidden layer is an MLP layer with 10 neurons. The learning rate used for the Adam optimizer is 0.001.

The mean square error (MSE) is used as the loss function. The ReLU activation function is used in the LSTM layers. The LSTM model is trained for 1000 epochs with a batch size of one in order to make one-step forecasts on the test.

2.2.2. Experiment

The experiment was developed on an open-source machine learning library called TensorFlow 1.15.0. Keras which runs on Tensorflow is also used. Keras enables quick and easy prototyping, provides the base for convolutional, recurrent neural networks and combinations thereof, and runs seamlessly on central processing units (CPUs) and Graphical Processing Unit (GPUs).

2.2.3. Performance Evaluation

In order to evaluate the forecasting performance of the proposed methods, the following error evaluation criteria are used: the MAE (mean absolute error), and RMSE (root mean square error).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y'_i - y_i| \tag{16}$$

$$RMSE = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}(|y_, - y'_I|)^2\right]} \tag{17}$$

Where $y_i$is the actual output (exchange rate) at time *i* and y'_i is the predicted output at time *i.*

3. RESULTS

(Fig. 3) and (Fig. 4) show the graph of the actual and predicted series for the test dataset with MLP neural network and stacked LSTM respectively. The results as well as the performance of the MLP neural network and the Stacked LSTM model are discussed below. The actual test data is in blue and the predicted values are in orange.

Figure 3. Exchange Rate USD/TL prediction with MLP neural network

Figure 4. Exchange Rate USD/TL prediction with stacked LST

From the results above, we observe that both the MLP neural network and the Stacked LSTM models performed really well in fitting the test dataset. However, the Stacked LSTM provided high accuracy for short term prediction with error less than 0.04 for daily exchange rate predictions.

In order to further evaluate the performance of our models, we compare them with a linear conventional time series forecasting model like the ARIMA. The following Tables 1 shows overall results of methods ARIMA, MLP, and Stacked LSTM respectively in terms of MAE, MSE, Stacked LSTM, evaluation metrics**.**

Table 1. The Average Error Rates Of ARIMA, MLP And Stacked LSTM Models Across All Lags

	ARIMA	MLP neural network	Stacked LSTM
MAE	1.86	**0.394**	**0.018**
RMSE	2.056	**0.395**	**0.037**

From (table 1), the two deep learning methods: MLP neural network and Stacked LSTM for predicting exchange rates achieve the best performance in all cases with Stacked LSTM having an edge over MLP neural network and thus it is beneficial for modeling exchange rates of Turkey.

4. CONCLUSION AND FUTURE WORKS

Financial flows of all economic payments are directly affected by the future value of the exchange rates; therefore, forecasting exchange rate movements with high accuracy is very important for all active market participants. Even though forecasting of exchange rates isa challenging task since it depends on several factors including economic and political ones. The accurate predictions of exchange rates have led economic actors to invest more in economy.

This study focused on the comparison of neural networks and other statistical models for exchange rate forecasting. Our main aim is to determine which deep learning prediction methods are best to forecast the exchange rate between USD and TL. We applied the MLP neural network and the Stacked LSTM models to daily USD/TL exchange rates from January 2001 to August 2020. Before feeding this data into the training models, preprocessing steps were applied on the dataset to make the values more appropriate for the models. These preprocessing steps included cleaning the data, converting the data to a supervised learning problem and applying differencing.

In order to evaluate the prediction performance of our models, the MAE and the RSME performance metrics were employed in this study.

Our conclusion is that the accuracy shown by MLP neural network and Stacked LSTM for predicting exchange rates are encouraging and inspires future studies that would focus on extending these to more complex hybrid deep learning techniques. Additionally, in future studies LSTMs models with changeable length of time sequence inputs can be used to help the models automatically determine the optimal time lags.

REFERENCES

Azzouni, A. and Pujolle, G. (2017). A Long Short-Term Memory Recurrent Neural Network Framework For Network Traffic Matrix Prediction. Accessed Adress *arXiv preprint arXiv:1705.05690*

Bengio, Y., Simard, P. and Frasconi, P. (1994). Learning Long-Term Dependencies With Gradient Descent İs Difficult, *Ieee Transactions On Neural Networks, Vol.* 5, No. 2: 157-166.

Chandwani, D., and Manminder S.S. (2014). Stock Direction Forecasting Techniques: An Empirical Study Combining Machine Learning System With Market İndicators İn The Indian Context, *International Journal of Computer Applications, Vol. 92*, No. 11: 8-17.

Das, S.R., Mishra, D. and Rout, M. (2020). A Hybridized ELM-Jaya Forecasting Model For Currency Exchange Prediction, *Journal of King Saud University-Computer and Information Sciences, Vol. 32*, No:3: 345-366.

Faraway, J., and Chatfield, C. (1998). Time Series Forecasting With Neural Networks: A Comparative Study Using The Airline Data, *Journal of the Royal Statistical Society: Series C (Applied Statistics), Vol. 47*, No. 2: 231-250.

Gers, F.A., Schmidhuber, J. and Cummins, F. (2000). Learning To Forget: Continual Prediction With LSTM, Neural Computation, Vol.12, No. 10: 2451-2471.

Gers, F. A., Schraudolph, N.N. and Schmidhuber. J. (2002). Learning Precise Timing With LSTM Recurrent Networks, *Journal Of Machine Learning Research, Vol. 3*, No.1: 115-143.

Graves, A., Jaitly, N. and Mohamed, A. (2013). Hybrid Speech Recognition With Deep Bidirectional LSTM, IEEE Workshop On Automatic Speech Recognition and Understanding. Accessed Adress *https://ieeexplore.ieee.org/document/6707742*

Greff, K., Srivastava, R.K., Koutník,J., Steunebrink, B.R. and Schmidhuber. J. (2016). LSTM: A search space odyssey, *IEEE transactions on neural networks and learning systems*, Vol.28, No.10: 2222-2232.

Hochreiter, S., Bengio, Y., Frasconi, P. and Schmidhuber. J. (2001). Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. Accessed Adress http://www.bioinf.jku.at/publications/older/ch7.pdf

Hu, Jiaojiao, Wang, Xiaofeng, Zhang, Ying, Zhang, Depeng, Zhang, Meng and Xue, Jianru (2020). Time Series Prediction Method Based on Variant LSTM Recurrent Neural Network, *Neural Processing Letters*, Vol. 5: 1–2

Jaeger, H. (2002). Tutorial On Training Recurrent Neural Networks, Covering BPPT, RTRL, EKF And The And The"Echo State Network Approach. Accessed Adress https://citeseerx. ist.psu. edu/viewdoc/download? doi=10.1.1.378.4095&rep=rep1&type=pdf

Khairalla, M., and AL-Jallad. N.T. (2017). Hybrid Forecasting Scheme For Financial Time-Series Data Using Neural Network And Statistical Methods, *International Journal of Advanced Computer Science and Applications*, Vol.8, No.9: 319-327.

Kleiner, B. (1977). Time series analysis: Forecasting and control, *Technometrics, Vol. 19*, No.3:343-344.

Kong, Y., Huang, Q.. Wang, C., Chen, J., Chen, J. and He, D. (2018). Long short-term memory neural networks for online disturbance detection in satellite image time series, *Remote Sensing Vol.10*, No.3: 452.

LeCun,Y., Bengio, Y. and Hinton, G. (2015). Deep Learning", *Nature*, Vol.521, No.7553: 436-444.

Li, Y. and Cao, H. (2018). Prediction For Tourism Flow Based On LSTM Neural Network, *Procedia Computer Science*, Vol.129: 277-283.

Okasha, M.K. (2014). Using Support Vector Machines İn Financial Time Series Forecasting, *International Journal of Statistics and Applications*, Vol.4, No.1: 28-39.

Osório, G. J., Lotfi, Shafie-khah, M., Campos, V. and Catalão, J. (2019). Hybrid Forecasting Model For Short-Term Electricity Market Prices With Renewable Integration, *Sustainability* Vol.11, No.1: 57.

Pallabi, P. and Kumari, B. (2017). Stock Market Prediction Using ANN SVM ELM: A Review, International Journal of Emerging Trends & Technology in Computer Science, Vol.6, No.3: 88-94.

Qiu, J., Wang, B., & Zhou, C. (2020). Forecasting stock prices with long-short term memory neural network based on attention mechanism, *PloS one*, Vol. 15, No.1:1-15

Sujin, P., Lee, J., Cha, M. and Jang, H. (2017). Predictability of machine learning techniques to forecast the trends of market index prices: Hypothesis testing for the Korean stock markets, *PloS one*, Vol.12, No.11: e0188107.