

Using Artificial Neural Networks to Forecast Egyptian economic indicator

Ahmed Nabil

Business Information Systems Dept.
Helwan University,
Cairo, Egypt
Ahmed.Nabil21@commerce.helwan.edu.eg
Khaled M. Metwally
Comp. Eng. & A.I. Dept.
Military Technical Collage
Cairo, Egypt
k.metwally@mtc.edu.eg

Omar Saker

Business Information Systems Dept.
Helwan University,
Cairo, Egypt
omersaker@yahoo.com
I.M. El-Henawy
Faculty of Computers and informatics
Zagazig University,
Zagazig, Egypt
henawy2000@yahoo.com

Abstract—Gross Domestic Product (GDP) refers to the market value of all final goods and services produced within a country in a given period. GDP can be thought as the size of an economy, this allowed many decision makers relied on GDP in their economic polices and decisions. Thus, GDP considered the foremost important macroeconomic performance measure of a country. Regarding the GDP prediction economic impact, GDP prediction has gained a great interest last years, this could help decision makers in their economic decisions and applied policies. Future GDP expectations becomes the primary determinant in many other sectors, e.g., investment sector, employment sector, economic sector, and also in stock market activities. This research work presents a GDP prediction case study on Egypt economy. The presented study implemented alternative prediction models that relied on Artificial Intelligence (AI) techniques, e.g., Long Short Term Memory (LSTM). An enhanced LSTM models were proposed for GDP prediction. The proposed enhanced LSTM models have implemented uni-variate (GDP) time-series and multivariate (GDP, Unemployment, Inflation rate) time-series schemes integrated with an ensemble approach for selecting the best prediction results. The obtained results showed that multivariate with two macroeconomic indicators (GDP and Inflation rate) outperformed uni-variate and multivariate with three indicators in Root Mean Square Error (RMSE) and Coefficient-of-Determination (R^2).

Keywords—Forecasting Gross Domestic Product, Economic Indicators, Artificial Neural Networks.

I. INTRODUCTION

Recently, macroeconomic prediction gained a potential interest world wide in decision makers policies. Decision makers aimed to achieve sustainable development and profitable growth for their countries economy. Moreover, market participants heavily rely on different indicators to assess their business reputation. Last decade has witnessed a great improvement in the developing countries economic policies. This achievement has been carried out by the fast emerging of statistical and prediction techniques in the economic field. Moreover, these techniques allowed decision makers testing their polices prior applying.

Economic enhancement in the developing countries is

mainly based on a set of metrics and indicators that are frequently monitored from financial analysts and specialists to discover valuable insights that help decision makers in drawing an accurate economic policy. Inflation rate, exchange rate, Gross Domestic Product (GDP), unemployment rate are examples of macroeconomic indicators that indicated the economic status of the country.

Indicators contain information that can help understand and forecast GDP indicator. Elliott et. al. in [1] argued that knowing the economy's possible direction and events in advance will improve the process for decision makers. Government policy makers, economists, businessmen, investors, employees and consumers all rely on forecasts for future judgment and base their strategic decisions on this information [2]. Therefore, it is important that economic indicators are reliable and provide accurate information in order for different players to interpret them correctly.

Other indicators like Consumer Price Indices (CPIs), also known as Inflation rate, measure the changes over time in the prices of consumer goods and services acquired, used or paid for by households. CPIs aim to cover all the goods and services consumed by a country's population within its territory. The price changes are measured against a representative set of consumer goods and services. Yet to date, their impact in the field of economic policy has been largely muted or exploratory in nature [3].

The unemployment rate is an important indicator with both a social and an economic dimension. From economic perspective, unemployment indicates unused available labour (in persons). Rising unemployment may result in loss of income for individuals and increased pressure on government spending on social benefit schemes.

Using AI techniques facilitate these ideas in the financial sector although historical data is available. ANN as one of the popular techniques in AI is a system resembling biological neural systems and uses working principles of human brain as a base. ANN can be applied in various fields for the purposes of forecasting, classification, optimization, data binding and

so on. ANN has been frequently used in financial applications in recent years [4]. In this research, alternative ANN models have been used in predicting GDP of the Egyptian economy. With the advent of the digital era, Artificial Intelligence (AI) has become a potentially fundamental technology impacting significant real-world applications [5].

In recent decades, the Artificial Neural Network (ANN) model has been considered one of the popular AI models used in macroeconomic prediction including GDP, exchange rate, and other indicators.

Various attempts have been made in the context of macroeconomic prediction to achieve more accurate and reliable forecasting results, using various popular AI techniques, e.g., ANN, and its alternatives. It is worth mentioning that, the indicators impacts varies from country to other and from quarter to another, as well as the decision makers policy changes periodically. In Egypt, after years of economic instability and financial crisis, an increased attention towards studying the economic state from past, current and predict the future of the Egyptian economy. So, GDP is a summary measure for economic production. It is generally considered to be an overall indicator of the development of the economy.

The impressive performance of ANNs in other domains raises the question of their performance in predicting macroeconomic indicators. ANNs have been applied to macroeconomic prediction [6]. However, because of the time-series nature of many economic prediction applications, the Long Short-Term Memory (LSTM) architecture is better suited to the problem than the traditional feed forward architecture explored in [6]. LSTMs considered an extension of Recurrent Neural Network (RNN) architecture, that introduces a temporal component to ANNs. LSTMs have been used to forecast meteorological events [7] as well as GDP [8].

However, use of LSTMs in forecasting economic variables remains in its early stages, perhaps partly because of high barriers to their implementation as well as the availability of the data for learning. Many common deep learning frameworks, including Keras and Python, include provisions for LSTMs. However, the implementations are general and require knowledge of the frameworks to successfully implement.

The remainder of this paper is organized as follows: section II will further explain the related work and its challenges; section III will explore back ground on ANN, RNN, and LSTMs; section IV explains the conducted research methodology; section V illustrates the adopted models in the implementation, section VI presents the conducted experiments to evaluate the proposed models and their performance evaluation. Section VII performs a discussion on the obtained results Finally, section VIII conclude and examine areas of future research.

II. RELATED WORK

The growing interest in applying machine learning techniques in forecasting popular economic indicators, motivated this research progress in that context. A growing number of studies have applied recent machine learning models in macroeconomic indicators forecasting. Some of these macroeconomic indicators prediction related work will be presented, e.g., exchange rate, inflation rate, GDP, and unemployment rate using different techniques machine learning techniques.

In [9], the author demonstrated the adoption of deep learning models in predicting the top-traded currencies exchange rates. The author applied ANN, Support Vector Regressor (SVR), Long Short-Term Memory (LSTM), and Neural Network (NN) with one and two hidden layer neurons to predict the multi-currency exchange rates. The adopted models predicted the exchange rate of the world's top-traded currencies; USD/GBP, USD/EUR, USD/AUD, USD/JPY, USD/CAD, USD/CHF, USD/CNY, USD/SEK, USD/NZD, USD/MXN and USD/INR from the daily currency exchange rate data collected in the period from 1980s to December 2018 (30–39 years). The performance evaluation results showed that the prediction model average accuracy exceeded 99%. The authors in this research proved the applicability of using deep learning in multi-currency exchange rates prediction.

In [10], the authors proposed a deep learning-based model in forecasting foreign currency exchange rate. The authors incorporated event sentiments in the proposed model to achieving accurate prediction, since the exchange market is a volatile market that always affected by the ongoing social and political events. Moreover, the authors considered other volatile factors that had highly impacts e.g., gold and crude oil prices. The proposed model was tested over 3 currency exchange rates; British pound sterling to US dollar (GBP/USD), Hong Kong Dollar to US dollar (HKD/USD), and Pak Rupee to US dollar (PKR/USD).

The proposed model validity was tested against linear regression and support vector regression models. For performance evaluation, MAE and RMSE were the most commonly used metrics in evaluating the foreign exchange rate prediction accuracy. The obtained results with/without sentiment showed that the proposed deep learning-based model outperformed the commonly used statistical techniques as well as support regression model in the prediction of foreign currency exchange rate. As a conclusion, incorporating the social media sentiment in the prediction model contributed significantly on the currency exchange rate prediction accuracy of Hong Kong, Pakistan, and the UK.

In [11], the authors aimed to achieve two objectives; the first objective was to explore the deep neural networks prediction accuracy in currency exchange rate compared with the well-known models of neural network and time-series analysis. The second objective was adapting deep neural network models for predicting currency exchange rates in emerging markets, i.e., Eastern Partnership (EaP) countries. In this research, the authors adopted raw exchange rate data of US dollar against developed countries currency e.g., Euro (EUR/USD), British Pound (GBP/USD), and Japanese Yen (USD/JPY) as input features to the model. The authors adopted data from 2000 to 2015 in training and testing the proposed models. Preliminary results suggested that indeed deep convolution networks outperformed other existing methods. The authors also adopted the exchange rates of Eastern Partnership (EaP) countries e.g., Armenia, Azerbaijan, Belarus, Georgia, Moldova, and Ukraine to US Dollar: AMD/USD, AZN/USD, BYR/USD, UAH/USD, GEL/USD, and MDL/USD. The authors claimed that in EaP countries' economies, the stability if the currency exchange market considered as one of the most important indicators in achieving sustainable development and growth on those countries.

Preliminary results confirmed that in the developed currency market, the proposed deep neural network model achieved significantly higher prediction accuracy than the baseline models: random walk model, two time-series models (ARIMA and ETS). For emerging currency market, the authors also proposed a novel set of input features that may help improve the prediction accuracy of such models.

In the work presented in [12], the authors discussed GDP as one of the most important indicators of economic growth, welfare, and health. The authors aimed in this study to investigate the possibility of adopting ANN with feed-forward back-propagation learning techniques in GDP prediction based on non-economic data. In achieving this objective, the authors tested different ANN models with more than one hidden layer and more neurons. The following performance metrics were used for ANN training evaluation including Value of correlation coefficient (R), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), Logarithmic Transformation Variable (e). Many experiments have been performed on the developed models, and according to the obtained results, it was observed that among the developed ANN models, the feed-forward back-propagation technique provided the best optimal ANN model. Therefore, ANN can be used as GDP estimation model based on non-economic parameters and the authors proved this observation with quite good and satisfactory results.

In [13] the authors aimed to enhance the prediction accuracy Anhui province GDP using an improved Back Propagation (BP) neural network. The improved BP model was proposed to overcome the shortcoming of the traditional BP neural network e.g., the local minimum and slow convergence speed, etc. The proposed model was compared against the traditional BP model and other models, such as ARMA and SVR model. The obtained prediction results showed the proposed model superiority over the alternatives on both RMSE and MAPE. It also showed that the proposed model has the good predictive ability in forecasting Anhui province GDP.

In [14] the author demonstrated the prediction of GDP development in Euro zone countries until the year 2025. The authors applied neural networks looking for the GDP growth time-series in the period 1960 to 2015, then based on the obtained results, they estimated the GDP growth of Eurozone countries until the year 2025. Radial Basic Function neural network (RBF) and Multiple Perceptron Neural network (MLP) with different structures were selected in the prediction process. Based on the conducted analysis, the RBF-1 10 1 network was determined to be the best since it exhibited satisfactory numerical characteristics, i.e., minimal residues. From the obtained results, it can be stated that the RBF models appeared to be the most useful tool for predicting GDP.

The authors in [15] aimed to forecast Albania's GDP. They designed an ANN model equipped with the Genetic Algorithm (GA) as a learning algorithm instead of the traditional gradient descent, called the "neuro-genetic" model. The authors adopted GA to train the weights of different ANN architectures, then compared the output of these models and find the best ANN architecture that achieved GDP forecasting with high accuracy. The neuro-genetic forecasting GDP growth accuracy was evaluated against real GDP growth using Mean Forecasting Error (MFE), Mean Absolute Error (MAE), Track-

ing Signal (TS), and Mean Square Error (MSE). The obtained results showed that the forecasting model tended to slightly over forecast, with an average absolute error of 0 : 195.

The Consumer Price Index (CPI), i.e., the inflation rate reflects the increased cost of living. The CPI is calculated by measuring the costs of essential services and goods, including vehicles, professional services, medical care, clothing, shelter, transportation, and electronics. Then, inflation is determined by the average increased cost of the total basket of goods over a period of time.

In [16], the authors presented a new approach that forecasts the inflation rate precisely considering the financial markets non-linear nature and complex behavior. The authors introduced a hybrid model that attempted to forecast the inflation rate in the presence of subtractive clustering technique and a fuzzy inference neural network. The authors started to select macroeconomic factors to predict the CPI historical data from the US markets. This selection has been carried out in three stages; the first stage involved performing a literature survey on macroeconomic variables, as a result a list of the candidate macroeconomic and financial variables have been generated. After that, the financial expert started to choose the most widely used and significant variables. At the second stage, rules identification was performed by applying the subtractive clustering algorithm. In the final stage, the generated rules in the second stage were fed to the Fuzzy Inference Neural Network. The authors aimed at these three stages to forecast the CPI prices change in coming periods.

In [17], the author demonstrated the adoption of ARIMA model in forecasting the Kingdom of Bahrain inflation rates since it is one of the most widely used methods in forecasting time-series data. The authors explained their steps in defining the ARIMA model components AR and MA, and its usage. After identifying the appropriate model, they considered only the AIC as the criteria for choosing the best forecasting model for Bahrain's inflation rate, and so, the ARIMA (0; 1; 1) model was accurately selected.

The used data for evaluating the prediction model was the annual inflation rates in Bahrain in the period from 1966 to 2017 and was collected from the World Bank for training the model. The authors in this research decided to forecast inflation rate in the Kingdom of Bahrain for the upcoming period from 2018 to 2027 and the defined model with accurate fitting results was selected. The authors confirmed that the chosen ARIMA (0; 1; 1) model was stable and suitable for predicting inflation in Bahrain over the period under study. The forecasting results obtained in the period from 2018 to 2027 showed that Kingdom of Bahrain inflation rate was expected to be around 1 : 5% in the next 10 years (approximately 1 : 5% by 2020).

In [18], the author demonstrated that macroeconomic modeling and forecasting is a widely researched area in the applied economic literature. Predicting the unemployment rate is one of the most important applications for economists and policymakers. In this paper, the authors reported the forecasting competition between Autoregressive (AR), Moving Average (MA), GARCH, EGARCH and TGARCH models of the UK monthly unemployment rate series. They compared the forecasting techniques based on the following symmetric error

statistics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percent Error (MAPE). The results from the comparisons of static and dynamic forecasts by the time-series models showed that the simplest models were the most appropriate for forecasting.

In [19], the authors main aim was to predict how individuals who are unemployed in previous term pass to employment or stay at current state with his/her characteristics and qualifications. In order to do so, Turkish labor market was selected as the study market where the individuals previous year state was known, however, present state was taken as unknown, which can be unemployment (0) or employment (1). In this case, the defined problem turned out to be a binary choice model. Binary choice models can be solved with logistic regression in econometrics and classification methods in machine learning literature. The family of the adopted classifiers was Nave Bayesian, Support Vector Machines, Decision Trees, Discriminant, Stochastic Gradient Classifier and Ensemble techniques. By comparing accuracies, specifically the XGBoost classifier which is in Decision Trees, and Random Forest that in Ensemble have nearly the same, 67% accuracy rate to predict present state of individuals in the Turkish labor market. However, logistic regression with the same variables has slightly lower accuracy, 63%; this exceeds half of ML algorithms in general. Since the results obtained belong to the test data, the variables and models can be used to estimate new data produced in the future to predict an individual's transition probability.

III. BACKGROUND

It is obviously clear that ANNs have been the catalyst to numerous advances in a variety of fields and disciplines in recent years including the economic field. One type of ANN alternatives, the Long Short-Term Memory network (LSTM), is particularly well-suited to deal with economic time-series. Further advantages include their ability to handle large numbers of input features in a variety of time frequencies. A brief background on ANN and its alternatives RNN and LSTM will be presented in this section.

A. ANN and RNN

ANNs are made up of various inter-connected layers composed of groups of nodes or neurons. The structure's conceptual similarity to the biological sort is the source of their name. Each of these nodes receives inputs either from the external data source, the "input layer", or from previous layers, "hidden" and "output" layers, the latter if the final output of the model. The output of a node is found by taking the weighted sum of all its inputs, the connections between individual nodes being the weights, and then running it through a non-linear activation function. In training, these weights are initially randomized, and when the data has passed through all layers of the network, an output is obtained, which is then run through a predefined cost function to assess performance. Then, using calculated gradients, or derivatives of the cost function, the network adjusts its weights to obtain an output with a smaller error, and the process is repeated. This is of course an oversimplification of the process, however, there exists a vast literature outlining and explaining the methodology of ANNs for those desiring a deeper examination of

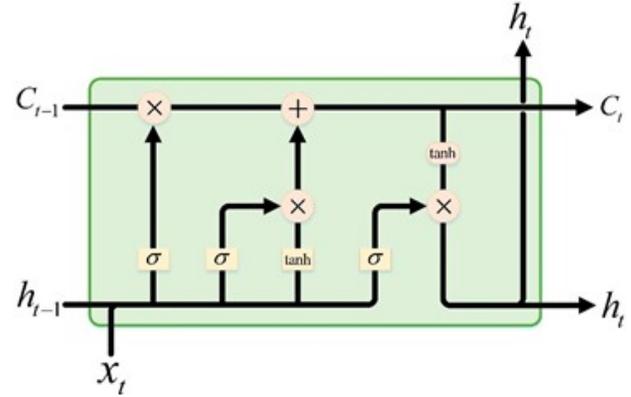


Fig. 1. Long-Short Term Memory architecture

their mathematics. Those interested can see [20], [21], or even [6] for an explanation in the nowcasting context.

Traditional feed forward ANNs have a long history of use in time-series forecasting [22]. These models, however, lack an explicit temporal aspect. This can be introduced to their architecture, resulting in Recurrent Neural Networks (RNN) [23]. As opposed to the unidirectional relationship between inputs and outputs in feed forward networks, RNNs introduce a feedback loop, where layer outputs can be fed back into the network [22]. This architecture makes RNNs well-suited to applications with a temporal aspect or flow, such as natural language processing or speech processing. However, due to vanishing gradients, RNNs tend to have a very "short" memory, limiting their usefulness in the nowcasting application [24].

B. LSTM

Long short-term memory (LSTM) networks are a special kind of RNN that can overcome the main issue of RNN, i.e., vanishing gradients using the gates to retain relevant information and discard unrelated details selectively. The structure of an LSTM neural network is shown in Figure 1, LSTMs introduce a memory cell and three gates: an *input*, *output* and *forget* gate [25]. Crucially, this architecture then allows gradients to flow unchanged through the network, mitigating the vanishing gradient problem of RNNs and rendering them more suitable for application to the forecasting problem.

IV. RESEARCH METHODOLOGY

The presented article research methodology followed the steps depicted in figure 2. Four stages summarized the conducted research work; Raw data retrieval, Preprocessing process, Building Prediction model, and finally Model Evaluation.

A. Raw Data

The adopted data was collected in the period from 1990s to December 2019 (30 years) for the Egyptian economy. The quarter inflation rate of the country was obtained from the *Central Agency for Public Mobilization and Statistics*, considering that 1990 is the initial base. The GDP data were obtained from the *Ministry of Planning*. Unemployment data were obtained from the *Central Agency for Public Mobilization and Statistics*.

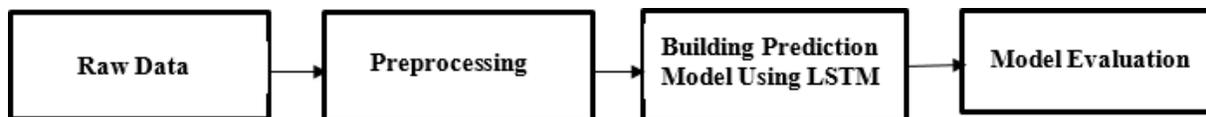


Fig. 2. Research Methodology

B. Preprocessing

Unfortunately, the obtained raw data is highly susceptible to noise, missing values, and inconsistency. The quality of data impacts the learning process and so affects the prediction accuracy. In order to help improve the quality of the data and, consequently, the prediction results, raw data has to be preprocessed so as to improve the efficiency of the prediction process. Data preprocessing is one of the most critical steps in machine learning and prediction process which deals with the preparation and transformation of the initial dataset. Data preprocessing stage is divided into following processes:

- **Data cleaning** The first process in data preprocessing is Data cleaning that recognizes partial, incorrect, imprecise or inappropriate parts of the data from datasets [26]. Data cleaning process involves eliminating typographical errors, ignoring tuple contains missing values, and other tasks. The data then becomes complete, free of errors, and consistent with other datasets available in the system.
- **Data integration** Data Integration is the method of merging data derived from different sources of data into a consistent dataset. Data collected from the web is usually expanding in size and complexity, and is either unstructured or semistructured. Integration of data is an extremely hard and iterative process. The considerations during the integration process are mostly related to standards of heterogeneous data sources. Moreover, the process of integrating new data sources to the existing dataset is timeconsuming, ultimately results in inappropriate consumption of valuable information.
- **Data transformation** Raw data is usually transformed into a format suitable for analysis. Data can be normalized for instance transformation of the numerical variable to a common range. Data normalization can be achieved using range normalization technique or *zscore* method. Categorical data can also be transformed using aggregation which merges two or more attributes into a single attribute. Generalization can be applied on lowlevel attributes which are transformed to a higher level.
- **Data reduction** Multifaceted exploration of huge data sources may consume considerable time or even be infeasible. When the number of predictor variables or the number of instances becomes large, machine learning algorithms suffer from dimensionality handling problems [27]. The last stage of data preprocessing is data reduction. Data reduction makes input data more effective in representation without loosening its integrity, it may or may not be lossless.

Many techniques can be used for data reduction, e.g., Encoding techniques and hierarchy distribution data cube aggregation. Instance selection [28] and Instance generation are two major approaches used by machine learning algorithm to reduce data size.

C. Building Prediction Model Using LSTM

In this research we aimed to predict the GDP as a case study in Egyptian economy. This proposal was achieved in two parts:

- **Part 1:** Uni-variate enhanced LSTM model that adopted an ensemble approach over a set of implemented LSTMs models that predicted uni-variate macroeconomic indicator in the Egyptian economy, i.e., GDP.
- **Part 2:** Multivariate enhanced LSTM model that adopted the highest LSTMs models prediction accuracy obtained in *Part-1* in GDP prediction. The adopted multivariate indicators was unemployment rate and inflation rate.

D. Model Evaluation

The main objective in this phase is to assess the trained model prediction accuracy. So, in this case, there is a need to work with at least training and test or validation datasets. Training datasets are used to create the model while test or validation datasets must be used to prove that the model is reliable and ready for prediction. Validating a predictive model is necessary to ensure that a model is effectively able to predict accurately the values of a variable of interest, in our case is GDP. Many performance metrics have been used in such prediction evaluation models, e.g., bias, Coefficient of Determination R^2 , Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and others. None of these aforementioned metrics is definitely better than the others. On the contrary, all the metrics must be used in conjunction to provide a better understanding of the prediction accuracy.

V. PROPOSED SYSTEM IMPLEMENTATION

Aforementioned, building the LSTM prediction model has been achieved in two parts, i.e., uni-variate and multivariate time-series approach. Different uni-variate LSTM models have been tested in GDP prediction, each implemented model has different layers and hyper-parameters.

- **Stacked-LSTM:** This LSTM model structured as multiple hidden LSTM layers that are stacked one on top of another. Each LSTM layer required a three-dimensional input and produces by default a two-dimensional output as an interpretation from the end of the stacked LSTM layers sequence.

```
Shape of training set: (74, 5)
Shape of test set: (36, 5)
Model: "sequential_88"
```

Layer (type)	Output Shape	Param #
conv_lstm2d_30 (ConvLSTM2D)	(None, 1, 1, 64)	17920
flatten_30 (Flatten)	(None, 64)	0
dense_135 (Dense)	(None, 16)	1040
dropout_88 (Dropout)	(None, 16)	0
dense_136 (Dense)	(None, 1)	17

```
Total params: 18,977
Trainable params: 18,977
Non-trainable params: 0
```

Fig. 3. Conv-LSTM model architecture

```
Total number of samples in the original training data = 79
Total number of samples in the generated data = 69
Model: "sequential_87"
```

Layer (type)	Output Shape	Param #
bidirectional_36 (Bidirectional LSTM)	(None, 64)	8704
dropout_87 (Dropout)	(None, 64)	0
dense_134 (Dense)	(None, 1)	65

```
Total params: 8,769
Trainable params: 8,769
Non-trainable params: 0
```

Fig. 4. LSTM-Bidirectional model architecture

- **Bidirectional LSTM:** In this LSTM type, the established LSTM model learn the input sequence both forward and backwards and concatenate both interpretations. This can be beneficial on some time series or sequence prediction cases. In our case study, a bidirectional LSTM for uni-variate time series forecasting has been implemented by wrapping the first hidden layer in a wrapper layer called Bidirectional.
- **CNN-LSTM:** A Convolutional Neural Network (CNN), is a type of neural network developed for working with two-dimensional image data. CNN models can automatically extracting and learning features from one-dimensional sequence data such as uni-variate time series data. Moreover, the integration of CNN with LSTM backend formed the hybrid CNN-LSTM model where the CNN is used to interpret subsequences of input that together are provided as a sequence to an LSTM model to interpret.
- **Conv-LSTM:** Although this LSTM type adopted convolutional operation, it differs from CNN-LSTM in such a way that the convolutional reading of input is built directly into each LSTM unit not as a separate CNN layer as in CNN-LSTM.

In this research work, three types of LSTM mentioned above have been implemented and evaluated, the obtained results showed that Bidirectional LSTM and Conv-LSTM outperformed other models. Figures 3 4 represented the convolutional and bidirectional LSTM models architecture for uni-variate and multivariate series.

The main objective in *Part-2* was improving the prediction accuracy as well as testing the impacts of other economic indicators on GDP prediction. Thus, the multivariate LSTM model was proposed. The achieved experiments have been carried out using the highest performance prediction models achieved in *Part-1* in a multivariate time-series prediction model. Unemployment rate and the inflation rate were the selected macroeconomic indicators in this phase.

VI. PERFORMANCE EVALUATION

A. Evaluation Metrics

Proposed economic prediction models need to be evaluated from different perspectives like: how much the predicted values deviated from the actual values; the adopted model used for prediction is useful or not; strength of linear relationship between dependent and independent variables,...etc. In this research, the prediction performance have been evaluated using the Root Mean Squared Error (RMSE) and Coefficient of Determination (R^2) evaluation metrics.

- **Root Mean Square Error (RMSE):** performance metric that measures the differences between predicted values by a model and the observed values.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{(d_i - f_i)^2}{\sigma_i}} \quad (1)$$

Where n the number of data, d_i is the actual value and f_i is the forecast value.

- **Coefficient of Determination (R^2):** The R^2 is a statistical measurement that examines how differences in one variable can be explained by the difference in a second variable, when predicting the outcome of a given event. In other words, this coefficient assesses how strong the linear relationship is between two variables. This metric is very important in trend analysis, the closer R^2 is to 1, the better the prediction.

R^2 and RMSE can be used together to study the variation in the errors in evaluating prediction models.

B. working environment

Data preparation and handling is entirely conducted in Python 3.8, relying on the packages numpy and pandas. The adopted deep learning LSTM models have been developed with keras over tensorflow on Google Colab platform. Regards the building models hyperparameters, The initial learning rate was 0.005, and the maximum number of epochs was fixed at 250 with batch size= 10. The gradient threshold is set to 1 to prevent the gradients from exploding. For each part in the solution we have run the experiments many times and averaged the obtained results, for uni-variate, two-multivariate, and three-multivariate.

C. Experimental results

The following results represented the achieved experimental results in uni-variate and multivariate experiments. In uni-variate experiments, four prediction models have been tested and evaluated; LSTM, Conv-LSTM, LSTM-Stacked, and LSTM-Bidirectional.

TABLE I. UNI-VARIATE PERFORMANCE EVALUATION

Exp	Conv-LSTM	LSTM-Sack	LSTM-Bidirection	LSTM
RMSE-Train	8480.64	9298.48	11780.81	15736.98
RMSE-Test	89488.09	257336.99	82358.65	236379.92
R^2 -Train	0.9821	0.9864	0.96627	0.9398
R^2 -Test	0.9221	0.36233	0.9401	0.5064

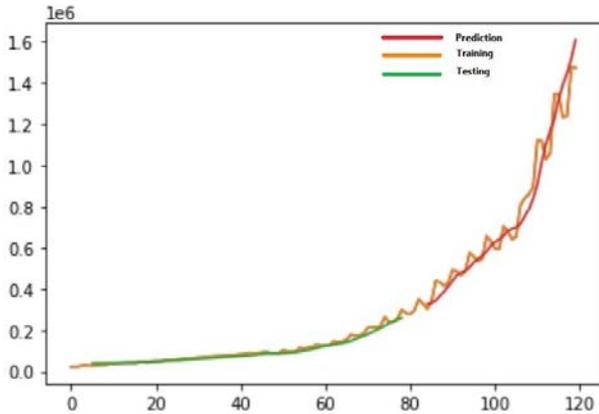


Fig. 5. Conv-LSTM uni-variate prediction model results

Table I represented *Part1* experiments results, the obtained preliminary results indicated that Conv-LSTM and LSTM-Bidirection achieved the highest scores in the evaluation metrics, thus they have been selected for *Part2* experiments. Moreover, figures 5,6 showed the prediction accuracy in univariate prediction model.

The following results illustrated the multivariate prediction results, the adopted variables in these experiments are GDP, Inflation rate, and Unemployment rate. The conducted experiments have been grouped in to two groups; two-multivariate experiments, i.e., GDP and Inflation rate, three-multivariate experiments, i.e., GDP, Inflation rate, and Unemployment rate.

TABLE II. TWO-MULTIVARIATE PERFORMANCE EVALUATION (GDP & INFLATION RATE)

Exp	Conv-LSTM	LSTM-Bidirection
RMSE-Train	11228.98	10697.43
RMSE-Test	98017.27	97946.14
R^2 -Train	0.9804	0.98221
R^2 -Test	0.9151	0.9152

Table II illustrated the two-multivariate performance evaluation, these experiments adopted the GDP and Inflation rate in predicting GDP using Conv-LSTM and LSTM-Bidirection models.

Table III represented the performance evaluation in predicting GDP using three-multivariate prediction model (GDP, Unemployment rate & inflation rate). The adopted prediction models were Conv-LSTM and LSTM-Bidirection.

TABLE III. THREE-MULTIVARIATE PERFORMANCE EVALUATION (GDP, UNEMPLOYMENT RATE & INFLATION RATE)

Exp	Conv-LSTM	LSTM-Bidirection
RMSE-Train	13562.43	11761.23
RMSE-Test	120335.39	142982.01
R^2 -Train	0.9714	0.9785
R^2 -Test	0.8720	0.8194

Fig. 6. Stacked-LSTM uni-variate prediction model results

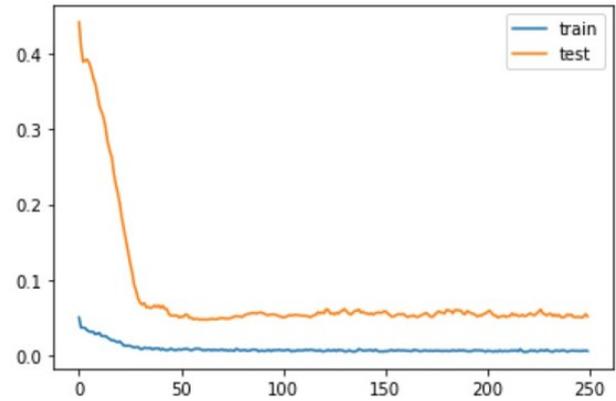


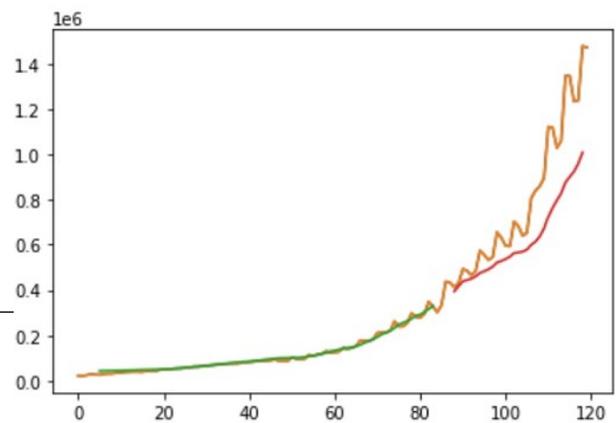
Fig. 7. LSTM-Bidirectional multivariate model (2 variables) learning curve

The following figures illustrated the adopted models learning curve for two-multivariate time-series, i.e., GDP and inflation rate. Also, learning curve three-multivariate time series, i.e., GDP, unemployment, and inflation rates. The conducted experiments aimed to predict Egypt GDP.

VII. DISCUSSION

In this section, we discuss and compare the results of our proposed model, other approaches, and the most related works. It was clear from the aforementioned related work that most of the achieved work mainly relied on ANN, ARIMA, and other techniques, e.g., SVM and linear regression models although they mentioned deep learning in their works. Most of the achieved work evaluated their models using RMSE, MAE, and Mean Absolute Percent Error (MAPE) performance metrics.

This research achieved the Egyptian GDP prediction using LSTM prediction models, this selection was motivated by the advantages on LSTM over other ANN alternatives. The conducted experiments covered three categories with multiple model architectures; uni-variate, two-multivariate, and three-multivariate. From the architectural point of view, figures 3 and 4 represented both LSTM-Bidirectional and Conv-LSTM model architectures. The number of the trainable parameters in *Conv LSTM* approximately doubled the number of the trainable parameters in *LSTM-Bidirectional* with more hidden layers. These parameters enhanced the learning process



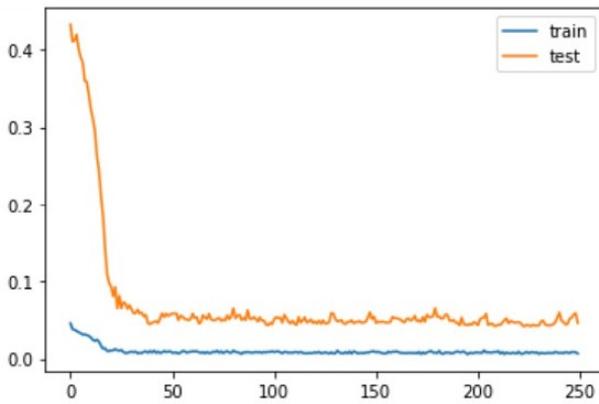


Fig. 8. Conv-LSTM multivariate model (2 variables) learning curve

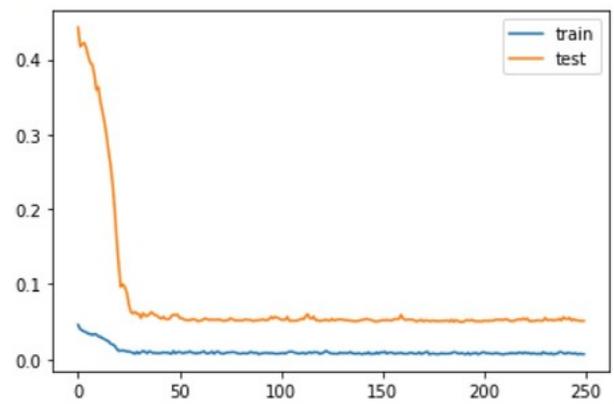


Fig. 10. Conv-LSTM multivariate model (3 variables) learning curve

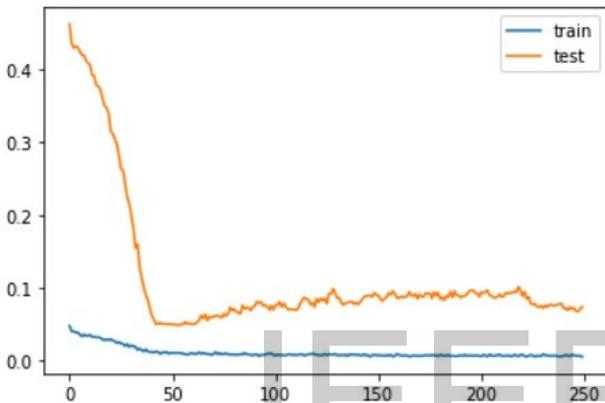


Fig. 9. LSTM-Bidirectional multivariate model (3 variables) learning curve

in terms of *RMSE* accuracy of *Conv-LSTM* as shown in tables I, II, and III, *Conv LSTM* achieved the least *RMSE*, and the highest R^2 compared to the other models. Also in multivariate case, *Conv LSTM* achieved better R^2 than *LSTM Bidirectional* model and near by *RMSE* accuracy. However, *Conv-LSTM* consumed longer time to converge to attain accuracy near to that of a *LSTM Bidirectional*. A lower value of *RMSE* and a higher value of R^2 indicate a good model fit for the prediction.

In uni-variate prediction models, figure 5 presented *Conv-LSTM* uni-variate GDP model prediction results, as it shown, it covered all the adopted dataset period, i.e., 120 record. The figure showed that an acceptable prediction results in the last period around the original values with minimum errors, i.e. oscillations. On the contrary, the prediction results obtained by *Stacked LSTM* model in figure 6 showed that the prediction values are far from the real dataset values on the same period. In multivariate prediction models, the convergence between the training and testing in the learning curves in figures 7,9, 8, 10, indicated that multivariate prediction model with two variables, i.e., GDP and Inflation rate converged faster than multivariate with three variables. Moreover, it has also been observed that the loss function converges smoothly in the case of *Conv-LSTM* (Figures 8, 10). As shown in the obtained results and figures the best result was in *Con-LSTM* two-multivariate experiments that outperformed uni-variate and three-multivariate time-series approaches.

VIII. CONCLUSION

This article tackled the problem of GDP prediction in the Egyptian economy. We first presented an efficient alternatives ANN learning models, i.e., LSTM with different architectures and hyper-parameters settings. These models are; Stacked-LSTM, Bidirectional LSTM, CNN-LSTM, and Conv-LSTM. We performed the prediction experiments using uni-variate GDP time-series from 1990 to 2019 in quarterly bases. Moreover, another set of prediction experiments have been achieved using multivariate time-series using Unemployment and inflation rates economic indicators in order to improve the obtained prediction accuracy in uni-variate models. As presented in the results the multivariate analysis with inflation rate and GDP indicators outperformed the other prediction models either uni-variate or multivariate with unemployment and inflation rates.

Some interesting directions for our future work become clear. First, we plan to investigate the applicability of the proposed models using wide range of economic indicators that are highly correlated with each others and directly impact the Egyptian economy. Moreover, try to increase the available data even by using machine learning data augmentation techniques like Generative Adversarial networks (GAN).

REFERENCES

- [1] C. W. J. Granger, G. Elliott, and A. Timmermann, *Handbook of Economic Forecasting*. Elsevier, 2006, vol. 1.
- [2] W. H. Schmidt, K. Heier, L. Bittner, and R. Bulirsch, *Variational Calculus, Optimal Control and Applications*. Springer, 1998.
- [3] L. Falat and L. Pancikova, "Quantitative modelling in economics with advanced artificial neural networks," *Procedia economics and finance*, vol. 34, pp. 194–201, 2015.
- [4] J. Logemann, "Measuring and managing expectations: Consumer confidence as an economic indicator, 1920s–1970s," *Futures Past. Economic Forecasting in the 20th and 21st Century*, 2020.
- [5] M. Haseeb, L. W. Mihardjo, A. R. Gill, K. Jermsittiparsert *et al.*, "Economic impact of artificial intelligence: new look for the macroeconomic assessment in asia-pacific region," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 1295–1310, 2019.

- [6] J. Loermann and B. Maas, "Nowcasting us gdp with artificial neural networks," 2019.
- [7] L. Galluzzi, J. Bravo-San Pedro, I. Vitale, S. Aaronson, J. Abrams, D. Adam, E. Alnemri, L. Altucci, D. Andrews, M. Annicchiarico-Petruzzelli *et al.*, "Essential versus accessory aspects of cell death: recommendations of the nccd 2015," *Cell Death & Differentiation*, vol. 22, no. 1, pp. 58–73, 2015.
- [8] Y. Igarashi, T. Kogure, Y. Kurihara, H. Miura, T. Okumura, Y. Satou, Y. Takahashi, and N. Yamaguchi, "A review of cs-bearing microparticles in the environment emitted by the fukushima dai-ichi nuclear power plant accident," *Journal of environmental radioactivity*, vol. 205, pp. 101–118, 2019.
- [9] A. R. Nagpure, "Prediction of multi-currency exchange rates using deep learning," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, no. 6, pp. 316–322, 2019.
- [10] M. Yasir, M. Y. Durrani, S. Afzal, M. Maqsood, F. Aadil, I. Mehmood, and S. Rho, "An intelligent event-sentiment-based daily foreign exchange rate forecasting system," *Applied Sciences*, vol. 9, no. 15, p. 2980, 2019.
- [11] S. Galeshchuk and S. Mukherjee, "Deep learning for predictions in emerging currency markets," 02 2017, pp. 681–686.
- [12] A. E. Tmer and A. Akku, "Forecasting Gross Domestic Product per Capita Using Artificial Neural Networks with Non-Economical Parameters," *Physica A: Statistical Mechanics and its Applications*, vol. 512, no. C, pp. 468–473, 2018.
- [13] Z. Xu, X. Wang, and Y. Jin, "Regional gdp prediction based on improved bp neural network model," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 9, no. 4, pp. 51–62, 2014.
- [14] J. Vrbka, "Predicting future gdp development by means of artificial intelligence," *Littera Scripta [online]*, vol. 9, no. 3, pp. 154–167, 2016.
- [15] D. Gjylapi and V. Kasemi, "A neuro-genetic model in gdp forecasting: Case study albania," 2013.
- [16] D. Enke and N. Mehdiyev, "A hybrid neuro-fuzzy model to forecast inflation," *Procedia Computer Science*, vol. 36, pp. 254 – 260, 2014, complex Adaptive Systems Philadelphia, PA November 3-5, 2014.
- [17] T. Nyoni, "Predicting inflation in the kingdom of bahrain using arima models," 2019.
- [18] C. Floros, "Forecasting the uk unemployment rate: model comparisons," *International Journal of Applied Econometrics and Quantitative Studies*, vol. 2, no. 4, pp. 57–72, 2005.
- [19] Y. KÜTÜK and B. GÜLOĞLU, "Prediction of transition probabilities from unemployment to employment for turkey via machine learning and econometrics: a comparative study," *Journal of Research in Economics*, vol. 3, no. 1, pp. 58–75, 2019.
- [20] M. H. Sazli, "A brief review of feed-forward neural networks," *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, vol. 50, no. 01, 2006.
- [21] R. K. Singh *et al.*, "Artificial neural network methodology for modelling and forecasting maize crop yield," *Agricultural Economics Research Review*, vol. 21, no. 347-2016-16813, pp. 5–10, 2008.
- [22] G. Dematos, M. S. Boyd, B. Kermanshahi, N. Kohzadi, and I. Kaastra, "Feedforward versus recurrent neural networks for forecasting monthly japanese yen exchange rates," *Financial Engineering and the Japanese Markets*, vol. 3, no. 1, pp. 59–75, 1996.
- [23] A. Amidi and S. Amidi, "Vip cheatsheet: Recurrent neural networks," 2018.
- [24] R. Grosse, "Lecture 15: Exploding and vanishing gradients," *University of Toronto Computer Science*, 2017.
- [25] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.
- [26] T. Dasu and T. Johnson, *Exploratory data mining and data cleaning*. John Wiley & Sons, 2003, vol. 479.
- [27] W. Jia, D. Dai, X. Xiao, and H. Wu, "Arnor: Attention regularization based noise reduction for distant supervision relation classification," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 1399–1408.
- [28] S. Vijayarani, M. J. Ilamathi, M. Nithya *et al.*, "Preprocessing techniques for text mining-an overview," *International Journal of Computer Science & Communication Networks*, vol. 5, no. 1, pp. 7–16, 2015.