

Gated Recurrent units (GRU) for Time Series Forecasting in Higher Education

Hassan Bousnguar
IRF-SIC, Faculty of Science
Ibn Zohr University
Agadir, Morocco

Amal Battou
IRF-SIC, Faculty of Science
Ibn Zohr University
Agadir, Morocco

Lotfi Najdi
LIMA Laboratory, ENSA
Ibn Zohr University
Agadir, Morocco

Abstract— Accurate forecasting is essential for decision-making in higher education. Traditional time series forecasting models have limitations in handling complex relationships and nonlinear trends in data. This paper explores the use of a gated neural network, a type of deep learning technique, for time series forecasting in the context of higher education. The gated neural network selectively filters information based on the relevance of past inputs and has shown promise in various domains. We present a case study that demonstrates the effectiveness of this approach in predicting student enrollment numbers, a critical factor for resource planning in universities. The results show that the gated neural network outperforms traditional models and can help universities better plan their resources, allocate budgets, and adapt to changes in demand. The findings suggest that the use of a gated neural network can improve time series forecasting accuracy in higher education.

Keywords—RNN; GRU; Forecasting; Education data mining;

I. INTRODUCTION

Forecasting plays a crucial role in decision-making processes[1], particularly in the field of higher education[2]. Accurate predictions of future trends and patterns can help universities and other educational institutions to better plan their resources, allocate budgets, and adapt to changes in demand. Time series forecasting is a common approach used to make predictions based on historical data. However, traditional time series forecasting models often have limitations, such as their inability to handle complex relationships and nonlinear trends in data.

Various techniques and approaches are utilized in the pursuit of accurate forecasting, and multiple disciplines are engaged in developing solutions to this end. Statistical scientists were among the pioneers in this field, and they introduced forecasting models as early as the 1950s. One such model, the exponential smoothing model, was presented by Brown et al. in that decade. This basic model has since undergone numerous modifications and extensions, including seasonal exponential smoothing (i.e., the Holt-Winters method) and state space models with exponential smoothing (ETS). Moreover, researchers have endeavored to incorporate additional features such as trend and seasonality into the basic model. Recent studies have examined the use of exponential

smoothing for demand forecasting across various industries, including healthcare and retail.

The rest of this paper is organized as follows: In the second section, we discuss the Statistical method for time series forecasting, providing an overview of the key techniques used. The third section outlines the methodology that have been employed for time series forecasting. The fourth section presents the preliminary results of this work. Lastly, we conclude with a summary and an outlook for future work in the final section.

II. LITERATURE REVIEW

ARIMA models represent one of the most commonly used time series forecasting techniques in both academic research and practical applications. The core concept behind ARIMA involves modeling a time series as a combination of autoregressive (AR) and moving average (MA) components, optionally preceded by a differencing step to eliminate trends or seasonality. The ARIMA methodology was initially introduced by Box and Jenkins, who also proposed a systematic approach for model identification, parameter estimation, and diagnostic checking. Since then, a plethora of extensions and refinements of ARIMA have been put forward, including seasonal ARIMA (SARIMA) models, vector ARIMA (VARIMA) models, and transfer function models with ARIMA errors.

Machine learning algorithms can be broadly categorized into two families: those designed for sequential data and those developed for traditional machine learning tasks. The former family includes techniques such as recurrent neural networks (RNNs)[3] and long short-term memory (LSTM)[4] networks, which are capable of capturing dependencies between sequential data points. In contrast, traditional machine learning algorithms, such as decision trees, random forests[5] and KNN[6], are designed for working with static data.

In recent years, deep learning techniques have shown great promise in improving time series forecasting accuracy[4]. One such technique is the gated neural network, which is a type of recurrent neural network that can selectively filter information based on the relevance of past inputs. The gated neural

network has been applied successfully in various domains, including finance[7], Cyber-attacks [8], and energy forecasting[9]. In this paper, we explore the use of a gated neural network for time series forecasting in the context of higher education. We present a case study that demonstrates the effectiveness of this approach in predicting student enrollment numbers, which is a critical factor for resource planning in universities.

III. METHODOLOGY

A. RNN

Over the past few years, recurrent neural networks (RNNs) have been utilized in a range of areas such as automated translation, speech recognition, and voice dialogue generation. These models excel at handling historical data spanning extended periods, illustrated in Figure 14, making them particularly effective in predicting data from time series. Unlike conventional neural networks, RNNs take into account the impact of past data on future predictions (Lipton et al., 2015), achieved recursively for each sequence element. To achieve this, previous observations are stored in a "memory" layer that is continuously updated and contains the current state of the series (referred to as the hidden layer or hidden state).

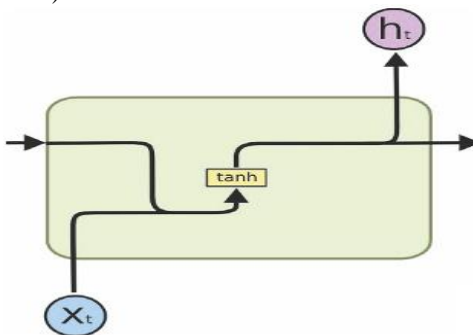


Fig. 1. RNN cell

B. LSTM

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) that are designed to overcome some of the limitations of traditional RNNs, such as the vanishing gradient problem.

The vanishing gradient problem occurs when the gradients used to train the neural network become very small during backpropagation, making it difficult to update the weights of the network. This problem is particularly prevalent in traditional RNNs because they rely on the repeated use of the same weight matrix, which can cause the gradients to become very small or even zero over time.

LSTM networks address this issue by introducing a memory cell that can selectively remember or forget information over time, allowing the network to retain important information while discarding irrelevant or redundant information. This enables LSTMs to handle long-term dependencies in the data, making them more effective than traditional RNNs for tasks such as language modeling, speech recognition, and image captioning.

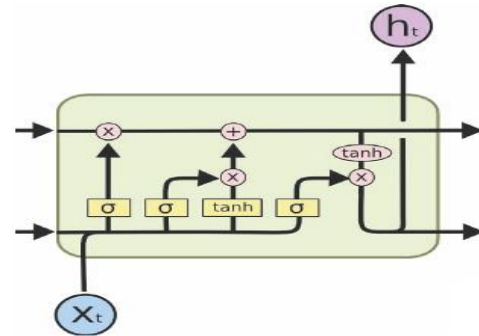


Fig. 2. LSTM Cell

C. GRU

GRU (Gated Recurrent Unit) and LSTM (Long Short-Term Memory) are both types of recurrent neural networks (RNNs) that are commonly used in natural language processing (NLP) and other sequence modeling tasks.

While both GRU and LSTM are designed to address the vanishing gradient problem in RNNs, GRU has fewer parameters than LSTM and is therefore faster to train and less likely to overfit the data. Additionally, GRU has a simpler architecture than LSTM, with only two gates (reset and update) compared to LSTM's three gates (input, output, and forget).

In some cases, using GRU instead of LSTM may result in comparable or even better performance, especially when the dataset is small and the sequence length is short.

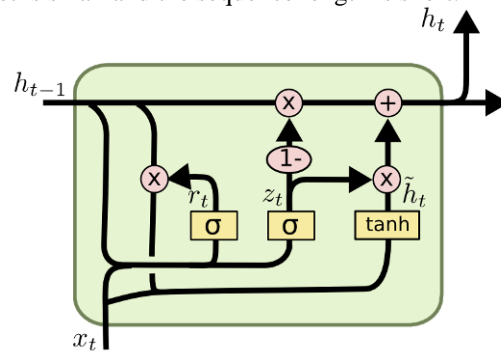


Fig. 3. GRU Cell

IV. RESULTS AND DATASET

A. DataSet

The Dataset that we use in this paper is extracted from enrollments in Alabama University between 1831 and 2019, we use an extract from this dataset to build our time series. In the Fig. 4 bellow we illustrate the time series.

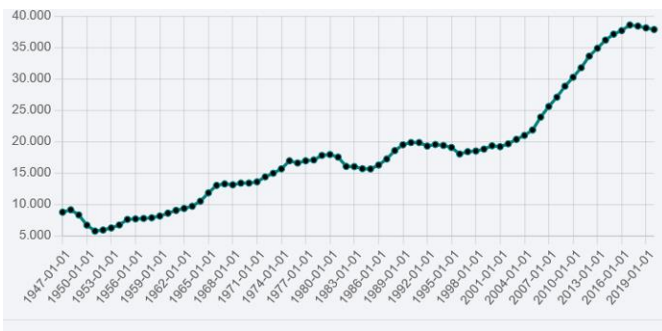


Fig. 4. Alabama university enrollments Time serie

B. Résultats

For the evaluation of models we use two metrics :

- MAE (Mean Absolute Error)
- RMSE (root-mean-square error)

Those metrics allow us to evaluate the accuracy of models. We train two models, LSTM and GRU, and we try to compare the results. In Fig. 5 we illustrate the GRU Models trained with Alabama Time series, and in Fig. 6, we illustrate LSTM With the same tome series.

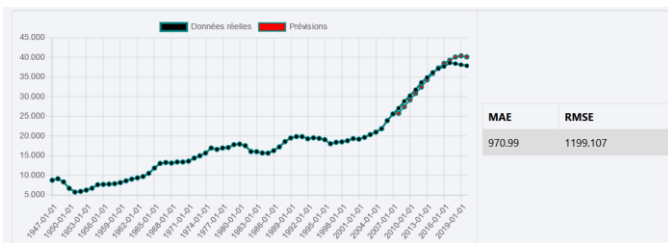


Fig. 5. GRU Model

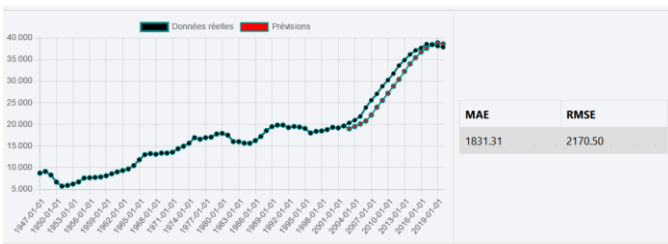


Fig. 6. LSTM Model

In the TABLE1, we presenting all resultants that we obtained with the two algorithms, and we can see that GRU give more accurate results than LSTM.

TABLE I. GRU and LSTM accuracy

Models	Metrics	
	RMSE	MAE
LSTM	2170.50	1881,31
GRU	970.99	1199.107

V. DISCUSSION

We analyzed the Alabama Enrollment time series dataset to evaluate the performance of two popular forecasting algorithms, GRU (Gated Recurrent Unit) and LSTM (Long

Short-Term Memory). Our objective was to compare the MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error) metrics of these algorithms and determine which one produced more accurate forecasts.

After training both models on the Alabama Enrollment dataset, we calculated the MAE and RMSE for each algorithm's predictions. The results showed that GRU outperformed LSTM in terms of both metrics, indicating that GRU produced more accurate forecasts.

The Alabama Enrollment dataset is a popular time series dataset that contains historical data on the number of students enrolled in public schools in Alabama. Despite its relatively it's not very wide, this dataset is commonly used to evaluate the performance of forecasting algorithms due to its complexity and non-stationary nature.

Our results suggest that GRU may be a better choice than LSTM for time series forecasting tasks, particularly when dealing with relatively small datasets like the Alabama Enrollment dataset. However, the choice of algorithm ultimately depends on the specific requirements of the task at hand and the characteristics of the data being analyzed.

REFERENCES

- [1] K. Krishnaiyer y F. F. Chen, «Web-based Visual Decision Support System (WVDSS) for letter shop», *Robotics and Computer-Integrated Manufacturing*, vol. 43, pp. 148-154, feb. 2017, doi: 10.1016/j.rcim.2015.09.016.
- [2] H. Bousnguar, L. Najdi, y B. Amal, «Forecasting approaches in a higher education setting», *Education and Information Technologies*, ago. 2021, doi: 10.1007/s10639-021-10684-z.
- [3] S. Smyl, «A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting», *International Journal of Forecasting*, vol. 36, n.º 1, pp. 75-85, ene. 2020, doi: 10.1016/j.ijforecast.2019.03.017.
- [4] A. H. Elsheikh, V. P. Katekar, O. L. Muskens, S. S. Deshmukh, M. A. Elaziz, y S. M. Dabour, «Utilization of LSTM neural network for water production forecasting of a stepped solar still with a corrugated absorber plate», *Process Safety and Environmental Protection*, vol. 148, pp. 273-282, abr. 2021, doi: 10.1016/j.psep.2020.09.068.
- [5] A. Bell y K. Jones, «Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data*», *Political Science Research and Methods*, vol. 3, n.º 1, pp. 133-153, ene. 2015, doi: 10.1017/psrm.2014.7.
- [6] S. K. Sharma y V. Sharma, «TIME SERIES PREDICTION USING KNN ALGORITHMS VIA EUCLIDIAN DISTANCE FUNCTION: A CASE OF FOREIGN EXCHANGE RATE PREDICTION», *undefined*, 2013. /paper/TIME-SERIES-PREDICTION-USING-KNN-ALGORITHMS-VIA-A-Sharma-Sharma/a1156a35b9190209ecb0bdc022df14431c0b164b (accedido 11 de enero de 2021).
- [7] L. Munkhdalai, M. Li, N. Theera-Umpon, S. Auephanwiriyaikul, y K. Ryu, «VAR-GRU: A Hybrid Model for Multivariate Financial Time Series Prediction», 2020, pp. 322-332. doi: 10.1007/978-3-030-42058-1_27.
- [8] D. Lavrova, D. Zegzhda, y A. Yarmak, «Using GRU neural network for cyber-attack detection in automated process control systems», en *2019 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, jun. 2019, pp. 1-3. doi: 10.1109/BlackSeaCom.2019.8812818.
- [9] B. Liu, C. Fu, A. Bielefeld, y Y. Q. Liu, «Forecasting of Chinese Primary Energy Consumption in 2021 with GRU Artificial Neural Network», *Energies*, vol. 10, n.º 10, Art. n.º 10, oct. 2017, doi: 10.3390/en10101453.