Forecasting of Soft drinks production using LSTM

Zoheib Tufail Khan[1], Mohd Faiz Aslam[2]

[1]Aligarh Muslim University, Aligarh, India

[2]Aligarh College of Engineering and Technology, Aligarh, India

Abstract- In this paper the authors have proposed a methodology to predict the soft-drinks production using recurrent neural networks (LSTM). LSTM has a feedback connection and can be easily used in deep learning to map complex interdependencies. This method can be applied to any industrial case where forecasting is required. Data for this research has been assumed by the authors after analyzing variety of sources.

Keywords- Forecasting, Time series, LSTM, RNN

1. INTRODUCTION

Over the past few years, forecasting methods have been used to enhance the art of decision-making and predicting the future sales. Mostly statistical techniques are employed for this purpose, but nowadays artificial neural networks (ANN) are also in use (1). These methods are widely used in stock prediction (2), sales prediction (3), and construction cost prediction (4). Many researchers have used the historical demand information to generate autoregressive integrated moving average (ARIMA) models (5). Automobile sectors across the globe have adopted these methods to predict their sales (6). Many researchers have engaged themselves to predict the demand for electric vehicles using this approach (7). This method also proves to be beneficial for companies like Coca-Cola (8) and many fashion retailers (9).

2. RESEARCH BACKGROUND

In this study, the authors have used recurrent neural networks (LSTM) to prepare a forecasting model for soft drinks production. Long Short-Term Memory (LSTM) networks are upgraded version of recurrent neural networks, which makes it simple to recall previous data in memory. The vanishing gradient problem of RNN is resolved here. LSTM is convenient to categorize, process, and predict time series given time lags of unknown duration. It trains the model by using back-propagation. In an LSTM network, three gates are present as shown in figure 1. The input gate discovers which value from the input should be used to modify memory. The equation for the input gate is as follows:-

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{c}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c)$$

The forget gate discovers what details to be discarded from the block.

The equation for gate 2 is as follows:

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f)$$

Then finally the output gate decides the output of the block using the input and the memory of the block. The equation for the output gate is

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o$$

$$h_t = o_t * \tanh(C_t)$$

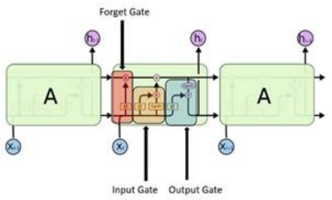

Figure 1: Three gates of LSTM

3. RESEARCH METHODS

Long Short-Term Memory (LSTM) is a type of recurrent neural network architecture that is developed to model sequences. These are machine learning models that take or give sequences of data and its vast range of dependencies more precisely than normal RNNs. To conduct this analysis the authors have taken data from 2007 to 2020 and then they have trained the data from 2007 to 2019 to predict the production for 2020. Later, the authors compared the actual value and the predicted value of 2020 for checking the performance of the trained model. This methodology can be extended in various industries for forecasting.

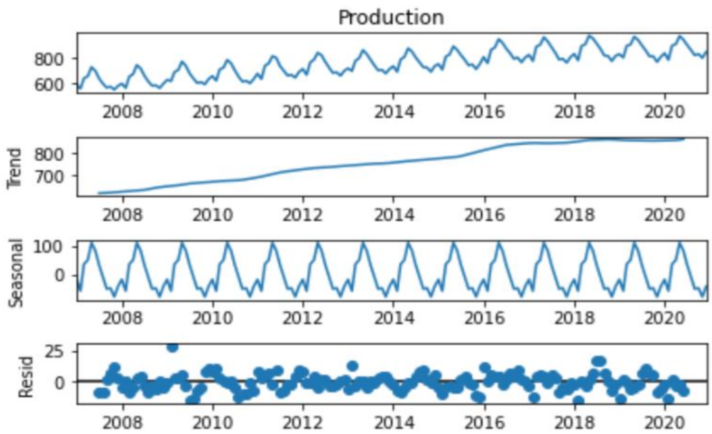

Figure 2: Seasonal decomposition

3.1 Dataset

Month	2007	2008	2009	2010	2011
January	589	600	628	658	677
February	561	566	618	622	635
March	640	653	688	709	736
April	656	673	705	722	755
May	727	742	770	782	811
June	697	716	736	756	798
July	640	660	678	702	735
August	599	617	639	653	697
September	568	583	604	615	661
October	577	587	611	621	667
November	553	565	594	602	645
December	582	598	634	635	688

***Data Table 1:** This table shows the data collected for year 2007 to 2011*

Month	2012	2013	2014	2015	2016
January	713	717	734	750	804
February	667	696	690	707	756
March	762	775	785	807	860
April	784	796	805	824	878
May	837	858	871	886	942
June	817	826	845	859	913
July	767	783	801	819	869
August	722	740	764	783	834
September	681	701	725	740	790
October	687	706	723	747	800
November	660	677	690	711	763
December	698	711	734	751	800

***Data Table 2:** This table shows the data collected for year 2012 to 2016*

Month	2017	2018	2019	2020
January	821	828	826	834
February	773	778	799	782
March	883	889	890	892
April	898	902	900	903
May	957	969	961	966
June	924	947	935	937
July	881	908	894	896
August	837	867	855	858
September	784	815	809	817
October	791	812	810	827
November	760	773	766	797
December	802	813	805	843

***Data Table 3:** This table shows the data collected for year 2017 to 2020*

4. **RESULT DISCUSSION**

In this research, the authors have forecasted the value of productions of soft drinks. Initially a model with dataset from 2007 to 2019 was trained and by using it the authors have predicted the production value of the year 2020 and later compared it with the actual value of production in 2020 presented in the dataset as shown in figure 3. The authors found that the predicted value is almost same as the actual value. Hence this method of forecasting is useful and gives better results.

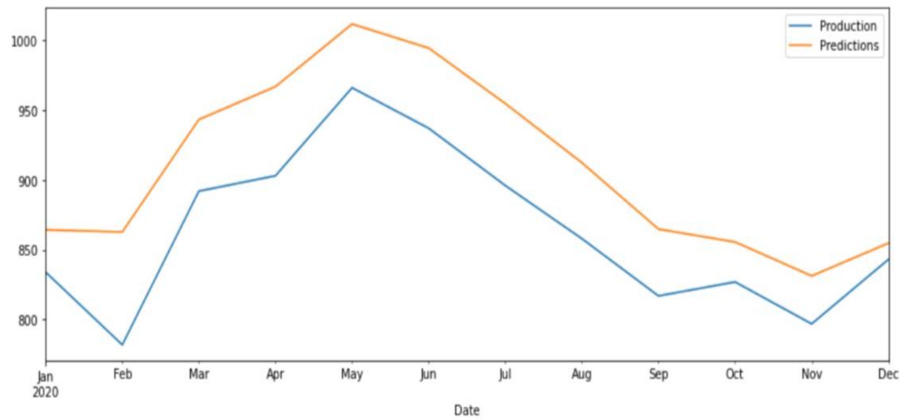

Figure 3: Comparison between Predicted value and the production value of year 2020

5. **CONCLUSION**

In this research, the authors have done the forecasting of soft drinks production using LSTM. The authors have analyzed the result obtained and the output. It can be seen that both the values are nearly the same. Hence it shows that this method of forecasting is highly accurate and should be employed in the future to solve similar forecasting and decision-making problems. It provides better results as compared to conventional statistical techniques.

References

1. Hill T., Marquez L., O'Connor M., Remus W. , 1994, "Artificial Neural Network models for forecasting and decision making", *International Journal of Forecasting*
2. Nayak A., Pai M., Pai R., 2016 , "Prediction Models for Indian Stock Market ", Procedia Computer Science
3. Chong A., Ngai E., Ch'ng E., Li B., 2015, "Predicting online product sales via online reviews, sentiments, and promotion strategies: A big data architecture and neural network approach", International Jour
4. Pandi V.,Nampoothiri N.,2014,"Analysis of construction cost prediction studies – global perspective", *International Review of Applied Sciences and Engineering*
5. Fattah J., Ezzine L., Aman Z., Moussami H., Lachhab A.,2018, "Forecasting of demand using ARIMA model", International Journal of Engineering Business Management
6. Sahabuddin S., 2009, "Forecasting automobile sales", Management Research News
7. Zhang Y., Zhong M., Geng N., Jiang Y., 2017 , "Forecasting electric vehicles sales with univariate and multivariate time series models: The case of China", PLoS ONE
8. Gardner J., Moeller S., McGowan Jr C., 2011, "Using Accounting Information For Financial Planning And Forecasting: An Application Of The Sustainable Growth Model Using Coca-Cola", Journal of Business Case Studies
9. Beheshti-Kashi S., Karimi H., Thoben K., Lütjen M., Teucke M., 2015, "A survey on retail sales forecasting and prediction in fashion markets", Systems Science & Control Engineering