International Current Journal of Engineering and Science (ICJES)
Volume 4, Issue 1, pp 10-14, January-2025
Article DOI: https://doi.org/10.47001/ICJES/2025.401002
ISSN (Online): 2583-7257

Neural Network-Based Linguistic Framework for Successive Word Prediction to Enhance Language Processing

[1]Md. Sohal Hasan, [2]Saifullah Ilam, [3]Md. Sejuti Mondol, [4]Anika Sara, [5]Mahmud Khan Rana

[1,2,3,4,5]Department of Computer Science and Telecommunication Engineering, Noakhali Science and Technology University, Bangladesh

***Abstract* - The next word prediction process is crucial for improving natural language processing applications and aiding in the creation of coherent and contextually appropriate texts. This project focuses on the use of Long Short-Term Memory (LSTM) networks for next word prediction tasks, which are a specific type of recurrent neural network (RNN). Utilizing the strengths of LSTM, the research aims to develop a system that can accurately predict the following word in a text sequence. The study demonstrates the effectiveness and accuracy of LSTM by examining its methodological framework, architectural structure, training techniques, and performance evaluation metrics.**

Keywords: LSTM, Efficiency, Lingusitic, Evaluation Metrics.

I. INTRODUCTION

We propose a language modeling based framework, which enhances the rate of communication for those who type slowly. This framework uses word prediction methods that infer the next word which is meant to be typed based on the already typed words. Consequently, the texting or emailing is done a lot faster, while the system suggests the next word that is most likely to be part of the text.

Our framework is aimed at improving the speed of instant communication by providing suggestions that are relevant for the words being typed. Using sophisticated algorithms, our system is able to make accurate predictions about the next word and hence assists the user spend minimal time and effort typing. This functionality is not only useful for improving the efficiency of communication, but also helps users who struggle with traditional means of typing because of physical or other limitations.

Furthermore, our framework is designed to be adaptable and configurable, allowing users to change different parameters to suit their unique writing tastes and styles. This adaptability guarantees that the word prediction tool can successfully serve a wide range of customers and take into account various language patterns and communication situations. To sum up, our approach represents a significant step forward in facilitating rapid electronic communication. Our goal is to improve writing's accessibility and efficiency by utilizing predictive technology, which will result in more smooth and successful communication interactions.

II. LITERATURE SURVEY

In this study, the author uses a neural network-based linguistic framework to present a novel approach for predicting the next word in Ukrainian. Different model types and data preprocessing techniques are investigated in order to improve the model's performance. Additionally, the model's efficacy is evaluated and a modified version that takes contextual information into account is proposed. High levels of accuracy and efficiency are demonstrated by the suggested method in a wide range of natural language processing applications. The field of natural language processing for the resource-poor Ukrainian language is greatly advanced by this work. However, the amount of data used for assessment and the scope of the suggested method limit the research.

The application of recurrent neural networks (RNNs) to word prediction tasks is investigated in this paper. The authors stress the importance of next word prediction in a number of fields, such as machine translation, speech recognition, and text completion. After providing a brief introduction to RNNs and their design, they provide a next word prediction model that makes use of an LSTM network. A wide range of text data is used to evaluate the model, and it is contrasted with other models that already exist. Results show that the LSTM-based model outperforms other models in terms of accuracy and effectiveness.

The application of pre-training strategies for federated text models in the field of next word prediction is the main emphasis of this study. The authors suggest a framework that combines federated learning with pre-training techniques like BERT and GPT-2 in order to overcome the difficulties

Article DOI: https://doi.org/10.47001/ICJES/2025.401002
ISSN (Online): 2583-7257

involved in next word prediction. A sizable dataset is used to assess the suggested system, and its results are compared to those of other top models. According to the authors, the accuracy and scalability of their system surpass those of the most advanced models.

Recent research has shown a great deal of interest in the task of using recurrent neural networks (RNNs) to predict the next word in Telugu. Data communication and text sharing through typing have become the main forms of engagement in modern culture. The amount of text a user must enter can be significantly reduced by having the capacity to predict the next word in a series. A useful application of natural language processing (NLP), also known as language modeling, is this predictive capability. These prediction systems are used in many applications, such as autocorrect capabilities that are mostly used in email and messaging platforms. Additionally, these approaches are used by search engines like Google and software like Microsoft Word to predict the next word based on users' previous interactions. This research investigates the applications of NLP, Long Short-Term Memory (LSTM) networks, and deep learning techniques to enhance the process of next-word prediction.

III. METHODOLOGY

Conventional approaches in the current framework find it difficult to keep up with the way language usage changes over time and the dynamic character of linguistic patterns. Long short-term memory (LSTM) networks are able to model complex linguistic structures, adapt to changing linguistic trends, and capture long-range dependencies in order to address this complex problem. This improves the efficiency of language translation tasks like next-word prediction.

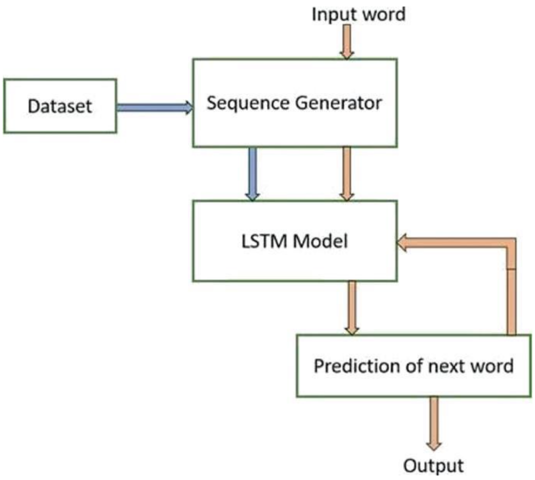


Figure 1: System Architecture

Our approach uses an LSTM-based architecture to build the model, which has a four-layer sequential design with two LSTM layers and two dense layers in between. The sequence length is set at ten, and the model is configured with three input sizes and one output size. Training takes place across 500 epochs with a batch size of 64. The LSTM model uses a gating method that combines input, output, and forget gates to efficiently control and update the state of its memory cells. Three sigmoid activation functions and one hyperbolic tangent activation function are used to describe each memory cell. Figure 3 depicts the arrangement of the LSTM memory cells.

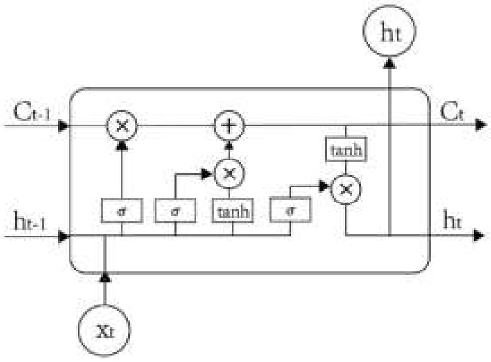


Figure 2: Basic Architecture of LSTM

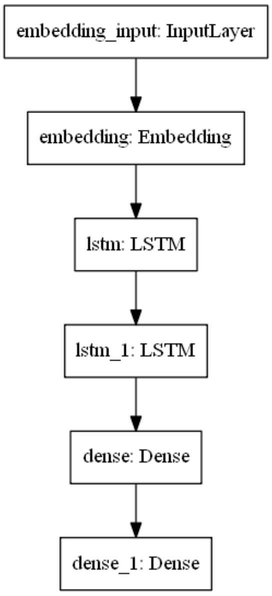


Figure 3: Data Flow Diagram

The analysis of the next word prediction system utilizing LSTM encompasses the following stages:

1. Collecting and preparing data: The first step involves gathering a sizable corpus of text data from various sources, such as books, papers, and websites. Preprocessing, which may include tokenization, lowercase conversion, punctuation removal, and segmentation into fixed-length sequences, is applied to the text data to guarantee cleanliness and consistency.

2. Encoding and feature extraction: Digital representations of every word in the corpus are created using techniques like word embedding and one-hot encoding. After that, word sequences are produced, each of which consists of the following word (target) and a predetermined amount of input words (context).

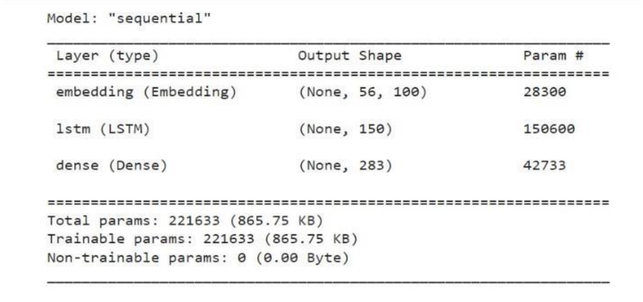
3. Architecture of the Model: The architecture of the LSTM model is fixed and usually consists of one or more LSTM layers, which are followed by one or more dense layers. While the dense layers are in charge of transferring the outputs from the LSTM layers to the vocabulary space for word prediction, the LSTM layers are essential for identifying long-term dependencies in input sequences. In order to reduce overfitting and improve the model's generalization skills, additional layers like batch normalization or pruning may be used.

4. Training: A prepared dataset is used in the training phase, in which input sequences serve as features and the target label is the word that best matches them. By using backpropagation of gradients throughout the network, the model iteratively modifies its parameters (weights and biases) during training in order to minimize a selected loss function, such as category cross-entropy.

5. Validation and Evaluation: To gauge the model's capacity for generalization and spot any possible overfitting, its performance is evaluated using a validation dataset after training. To measure the model's predicted ability statistically, a number of metrics can be used, such as BLEU score, precision, or perplexity.

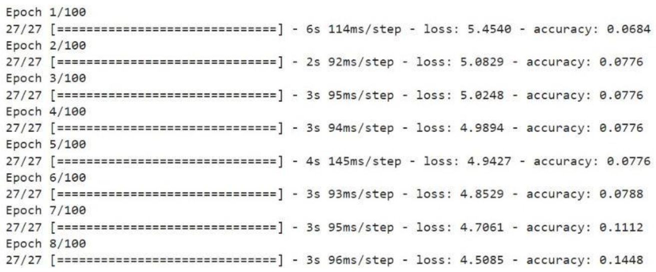
6. Implementation and Completion: The model is ready for deployment in real-time inference applications when training and validation are successfully finished. By processing the input through the trained LSTM network and choosing words with the highest probability based on the output distribution, the model predicts subsequent words given a sequence of input words. The anticipated words can be used to generate text, offer recommendations in text editing software, or help users with a variety of activities related to natural language processing.

7. Monitoring and Maintenance: Constant monitoring is essential to guaranteeing the deployed model's continued dependability and performance. Updating dependencies to fix security flaws or compatibility issues, modifying hyperparameters to improve speed, and retraining the model with fresh data to keep up with evolving language trends are some examples of maintenance tasks.



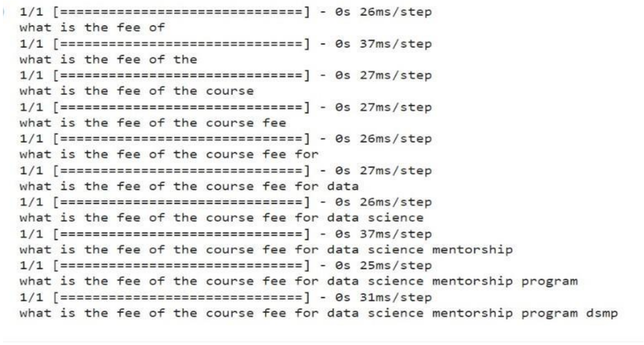
```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 56, 100)           28300

 lstm (LSTM)                 (None, 150)               150600

 dense (Dense)               (None, 283)               42733

=================================================================
Total params: 221633 (865.75 KB)
Trainable params: 221633 (865.75 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Figure 4: Model Prediction



```
Epoch 1/100
27/27 [==============================] - 6s 114ms/step - loss: 5.4540 - accuracy: 0.0684
Epoch 2/100
27/27 [==============================] - 2s 92ms/step - loss: 5.0829 - accuracy: 0.0776
Epoch 3/100
27/27 [==============================] - 3s 95ms/step - loss: 5.0248 - accuracy: 0.0776
Epoch 4/100
27/27 [==============================] - 3s 94ms/step - loss: 4.9894 - accuracy: 0.0776
Epoch 5/100
27/27 [==============================] - 4s 145ms/step - loss: 4.9427 - accuracy: 0.0776
Epoch 6/100
27/27 [==============================] - 3s 93ms/step - loss: 4.8529 - accuracy: 0.0788
Epoch 7/100
27/27 [==============================] - 3s 95ms/step - loss: 4.7061 - accuracy: 0.1112
Epoch 8/100
27/27 [==============================] - 3s 96ms/step - loss: 4.5085 - accuracy: 0.1448
```

Figure 5: Training Data



```
1/1 [==============================] - 0s 26ms/step
what is the fee of
1/1 [==============================] - 0s 37ms/step
what is the fee of the
1/1 [==============================] - 0s 27ms/step
what is the fee of the course
1/1 [==============================] - 0s 27ms/step
what is the fee of the course fee
1/1 [==============================] - 0s 26ms/step
what is the fee of the course fee for
1/1 [==============================] - 0s 27ms/step
what is the fee of the course fee for data
1/1 [==============================] - 0s 26ms/step
what is the fee of the course fee for data science
1/1 [==============================] - 0s 37ms/step
what is the fee of the course fee for data science mentorship
1/1 [==============================] - 0s 25ms/step
what is the fee of the course fee for data science mentorship program
1/1 [==============================] - 0s 31ms/step
what is the fee of the course fee for data science mentorship program dsmp
```

Figure 6: Model Prediction

The sequence length is set at ten, and the model is configured with three input sizes and one output size. Training takes place across 500 epochs with a batch size of 64. The LSTM model uses a gating method that combines input, output, and forget gates to efficiently control and update the state of its memory cells. Three sigmoid activation functions and one hyperbolic tangent activation function are used to describe each memory cell.

IV. CONCLUSION

A thorough set of procedures, including data collection, preprocessing, model architecture design, training, assessment, implementation, and continuing support, are involved in the use of Long Short-Term Memory (LSTM) networks for post-word prediction. These systems are excellent at producing contextually relevant word predictions by utilizing the sequential nature of text and the memory capacities of LSTM architectures. This improves a variety of text-related

Article DOI: https://doi.org/10.47001/ICJES/2025.401002
ISSN (Online): 2583-7257

applications and user experiences. Our project's results demonstrate how well deep learning techniques work to get around the drawbacks of conventional post-word prediction models.

In addition to demonstrating the potential for text prediction tasks, the use of LSTM networks in language modeling suggests prospective developments in the broader field of natural language understanding. Our experiment provides important insights into the future of increasingly complex and context-aware interactions with digital content and serves as an example of the continuous improvements in language processing systems as technology advances.

REFERENCES

[1] Shakhovska, K., Dumyn, I., Kryvinska, N., & Kagita, M. K. (2021). An Approach for a Next-Word Prediction for Ukrainian Language. Wireless Communications and Mobile Computing, 2021, 1-9.

[2] Ambulgekar, S., Malewadikar, S., Garande, R., & Joshi, B. (2021). Next Words Prediction Using Recurrent Neural Networks. In ITM Web of Conferences (Vol. 40, p. 03034). EDP Sciences.

[3] Stremmel, J., & Singh, A. (2021). Pretraining federated text models for next word prediction. In Advances in Information and Communication: Proceedings of the 2021 Future of Information and Communication Conference (FICC), Volume 2 (pp. 477-488). Springer International Publishing.

[4] Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. Procedia Computer Science, 170, 1168-1173.

[5] Afika, R., Suprih, W., Atikah, D. A., & Fadlan, B. H. (2022). Next word prediction using LSTM. Journal of Information Technology and Its Utilization, 5(1), 10-13.

[6] Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals, 135, 109864.

[7] Violos, J., Tsanakas, S., Androutsopoulou, M., Palaiokrassas, G., & Varvarigou, T. (2020, September). Next position prediction using LSTM neural networks. In 11th Hellenic Conference on Artificial Intelligence (pp. 232-240).

[8] R. Sharma, N. Goel, N. Aggarwal, P. Kaur and C. Prakash, "Next Word Prediction in Hindi Using Deep Learning Techniques", 2019 International Conference on Data Science and Engineering (ICDSE), pp. 55-60, 2019.

[9] O. F. Rakib, S. Akter, M. A. Khan, A. K. Das and K. M. Habibullah, "Bangla Word Prediction and Sentence Completion Using GRU: An Extended Version of RNN on N-gram Language Model", 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 2019.

[10] Wessel Stoop and Antal van den Bosch, "Using idiolects and sociolects to improve word prediction", Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, April 2014.

[11] Hozan K. Hamarashid, Soran A. Saeed and Tarik A. Rashid, "Next word prediction based on the N-gram model for Kurdish Sorani and Kurmanji", Neural Computing and Applications NCAA-D-19–02773R1, July 2020.

[12] Aurnhammer Christopha and Stefan L. Frankb, "Evaluating information-theoretic measures of word prediction in naturalistic sentence reading", Neuropsychologia, vol. 134, pp. 107198, 2019, ISSN 0028-3932.

[13] Partha Pratim Barman and Abhijit Boruah, "An RNN-based Approach for next word prediction in Assamese Phonetic Transcription", 8th International Conference on Advances in Computing & Communications (ICACC-2018), 2019.

[14] Eisape Tiwalayo, Zaslavsky Noga and Levy Roger, "Cloze Distillation: Improving Neural Language Models with Human Next-Word Prediction", Association for Computational Linguistics, 2020.

[15] K Smagulova and AP James, "A survey on LSTM memristive neural network architectures and applications", The European Physical Journal, 2019.

Citation of this Article:

Md. Sohal Hasan, Saifullah Ilam, Md. Sejuti Mondol, Anika Sara, & Mahmud Khan Rana. (2025). Neural Network-Based Linguistic Framework for Successive Word Prediction to Enhance Language Processing. *International Current Journal of Engineering and Science - ICJES*, 4(1), 10-14. Article DOI: https://doi.org/10.47001/ICJES/2025.401002

****