(IJAER) 2022, Vol. No. 24, Issue No. I, July **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

DEVELOPING THE DEEP LEARNING NEURAL NETWORK MODEL IN THE SUITABILITY/ FEASIBILITY ANALYSIS OF THE NATURAL LANGUAGE PROCESSING (NLP) IN INDUSTRY APPLICATIONS

Ahmed Abbas Naqvi

New Delhi, India

ABSTRACT

Due to its ability to computationally represent and analyze human language, natural language processing, or NLP, has recently received a lot of attention. Its expanded applications include machine translation, email spam detection, information extraction, summarization, medical diagnosis, and question answering. Deep learning and neural networks are used to analyze natural language syntax in this study. The first topic of this research is a transfer-based dependent syntax analyzer-based feedforward neural network-based classifier. A long-term memory neural network-based dependent syntactic analysis paradigm is the focus of this study. This model, which will be used as a feature extractor, is based on the feedforward neural network model. A syntactic analyzer's characteristics are used as input to train a recursive neural network classifier that is sentence-optimized and trained using a long short-term memory neural network. After learning how to use the feature extractor, this is done. Syntactic analysis, which models the analysis of the entire sentence, takes the place of independent modelling analysis. The experiment's outcomes demonstrate that the model has outperformed the benchmark techniques.

INTRODUCTION

A language is a set of rules or invariant symbols for expressing ideas or information. Natural Language Processing (NLP) helps druggies who need additional time to learn new languages or come complete in the bones they formerly know because only some druggies have a strong background in the machine-specific language. The thing of NLP, a subfield of linguistics and artificial intelligence, is to enable computers to comprehend mortal language statements and words(1). It was made to do work lightly and to satisfy the stoner's desire to communicate with a machine in standard English. It can be divided into two orders generation of natural language and appreciation of natural language," which advances comprehension and text production [2].

Phonology is a branch of language study that deals with sound; morphology, which is concerned with the formation of words and sentence structure; semantics, which is concerned with syntax; and pragmatics, which is concerned with understanding. Noah Chomsky, one of the earliest linguists of the 12th century to develop syntactic ideas, held a special place in the field of theoretical linguistics because the author redefined the study of syntax [3]. Natural language generation, or NLG, uses an internal representation to create meaningful words, sentences, and paragraphs.

(IJAER) 2022, Vol. No. 24, Issue No. I, July **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

Studying particular linguistic structures and norms, such as determining the rules for sentence word order and categorizing words, is referred to as "grammar" in computer linguistics [4]. Part-of-speech tagging and language models were two methods by which the linear laws of those languages could be expressed. Syntactic parsing is a prominent area of natural language processing research that has significant application and research relevance [5]. It is an important technique in many activities that use natural language.

In the 1940s, researchers first used the term "neural network" to describe systems for processing biological information [6]. Due to ongoing advancements in computer performance, deep neural networks can be trained extensively. Consequently, research into a wide range of machine learning domains has benefited greatly from the Deep Learning approach. To learn intricate structural representations, deep learning extensively uses massive amounts of information. Backpropagation and error-driven optimization techniques are used to alter the network parameters between multiple layers of artificial neural networks to achieve this kind of learning [7].

METHODOLOGY

In problems involving sequence identification or forecasting, recursive neural networks transform input sequences into output sequences. However, numerous real-world issues demonstrate how difficult it is to train recursive neural networks. The sequences in these problems typically span a longer period. Because their gradient will eventually disappear, it is harder for recursive neural networks to learn a long-distance memory. Long Short-Term Memory (LSTM), as suggested by the author [19], could be the answer to this problem. A "door" in this model lets the network choose when to "Forget" and when to add new "memory."

The long-term memory neural network, a variation of the recursive neural network, addresses the gradient disappearance of conventional recursive neural networks. While reading an input vector xt from a vector sequence (X1, X2,..., Xn), the typical recursive neural network calculates a new hidden layer state h1. Still, it cannot be used to describe long-distance dependence due to the problem of gradient disappearance. Long-short-term memory neural networks created a "Memory Cell" and three "Control Gates" to control when to select "memory" and when to select "forget."

Table 1: Statistical Data

Data Set	Total sentences	Projectable sentences
Training set	29834	29786
Development set	1710	1707
Test set	1710	1707

RESULTS

This issue is contrasted with the Malt Parser and MST Parser, two more well-known dependency parsers, in contrast to the baseline approach. The current work uses Malt Parser's stackproj and overeager training options. These choices represent both the arc-eager analysis algorithm and the arc-standard analysis method. Also, provide the MST Parser results in [125]. The test results are shown in Table 2. The table demonstrates that the long- and short-term memory neural network-based dependency syntax analyser has produced specific results in modelling the analysis sequence of phrases.

Table 2: WSJ Results

Analyzer	Development Set		Test Set	
	UAS	LAS	UAS	LAS
Present Model	90.5	90	91.2	90.4
Greedy feature extractor	89.6	89.2	90.3	90.2
Malt: eager	89.9	89.6	90.2	90.1
MST parser	90.1	89.8	90.4	90.1
Malt: standard	89.5	88.9	89.9	89.8
Baseline method	89.3	88.9	89.8	89.6

This model achieved a UAS accuracy of 90.50 per cent and LAS accuracy of 90.00 per cent on the Penn Tree Bank development set, a 0.60 per cent improvement over the baseline method's greedy neural network dependency parser. With UAS accuracy rates of 91.20 per cent and LAS accuracy rates of 90.40 per cent, the model presented in this study outperformed the greedy neural network-dependent syntax analyzer of the baseline method by approximately 0.55 per cent on the test set.

The experimental results show that the greedy feedforward neural network performs worse than the dependency syntax analysis model, which is based on long and short-term memory neural networks. In contrast to the greedy model, this model employs both long-term and short-term memory neural networks to model the entire sentence. Using pattern and historical analysis data, it can classify analysis activities. The dependent syntax analyzer's performance is enhanced as a result. The results of the test on the Pennsylvania Tree Bank. This article employs a column search during testing, and the appropriate beam size is 12.

The statistics in the table show that the dual attention method can significantly reduce the number of output result errors. In the final output, the model's Fl value reached 0.795.

CONCLUSION

This study investigates a neural network model of dependency syntactic analysis based on transfer learning. The current paper shows how a feedforward neural network can be used as a classifier in the dependent syntax analyzer. After evaluating the model, the neural network changes its parameters to get better results. This study suggests a long-term memory neural network-based dependent syntactic analysis paradigm.

REFERENCES

[1] Chi EC, Lyman MS, Sager N, Friedman C, Macleod C (1985) A database of computer-structured narrative: methods of computing complex relations. *In proceedings of the annual symposium on computer application in medical care* (p. 221). Am Med Inform Assoc

[2] Chomsky N (1965) *Aspects of the theory of syntax. MIT Press*, Cambridge, Massachusetts

[3] Choudhary N (2021) LDC-IL: the Indian repository of resources for language technology. *Lang Resources & Evaluation* 55:855-867. https://doi.org/10. l 007/sl 0579-020-09523-3

[4] Chouikhi H, Chniter H, Jarray F (2021) Arabic sentiment analysis using BERT model. In *international conference on computational collective intelligence* (pp. 621-632). Springer, Cham

[5] Chung J, Gulcehre C, Cho K, Bengio Y, (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. *Presented in NIPS 2014 Deep Learning and Representation Learning Workshop*, arXiv preprint arXiv:1412.3555

[6] Cohen WW (1996) Learning rules that classify e-mail. *In AAAI spring symposium on machine learning in information access* (Vol. 18, p. 25)

[7] Collobert R, Weston J (2008) A unified architecture for natural language processing. *In proceedings of the 25th international conference on machine learning* (pp. 160-167)

[8] Dai Z, Yang Z, Yang Y, Carbonell J, Le QV, Salakhutdinov R, (2019) Transformer-xi: attentive language models beyond a fixed-length context. *ACL 2019 long paper,* arXiv preprint arXiv: 1901.02860

[9] Davis E, Marcus G (2015) Commonsense reasoning and commonsense knowledge in artificial intelligence. *Commun ACM* 58(9):92-103

[10] Desai NP, Dabhi VK (2022) Resources and components for Gujarati NLP systems: a survey. *Artif Intell Rev:*1-19

[11] Devlin J, Chang MW, Lee K, Toutanova K, (2018) Bert: pre-training of deep bidirectional transformers for language understanding. *Computation and Language*, arXiv preprint arXiv:1810.04805

[12] Diab M, Hacioglu K, Jurafsky D (2004) Automatic tagging of Arabic text: From raw text to base phrase chunks. In *Proceedings of HLT-NAACL 2004: Short papers* (pp. 149-152). Assoc Computat Linguist

[13] Doddington G (2002) Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. *In proceedings of the second international conference on human language technology research* (pp. 138-145). Morgan Kaufmann publishers Inc

[14] Drucker H, Wu D, Vapnik VN (1999) Support vector machines for spam categorization. *IEEE Trans Neural Netw* 10(5):1048-1054

[15] Dunlavy DM, O'Leary DP, Conroy JM, Schlesinger JD (2007) QCS: A system for querying, clustering and summarizing documents. *Inf Process Manag* 43(6):1588-1605

[16] Müller, M.; Ewert, S. (2011), Chroma Toolbox: MATLAB implementations for extracting variants of chroma-based audio features. *In Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR)*, Miami, FL, USA, 24–28 October 2011.

[17] Fuentes, B.; Liutkus, A.; Badeau, R.; Richard, G. (2012), Probabilistic model for main melody extraction using constant-Q transform. *In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, 25–30 March 2012; pp. 5357–5360.

[18] Durand, S.; Bello, J.P.; David, B.; Richard, G. (2016), Robust downbeat tracking using an ensemble of convolutional networks. *IEEE/ACM Trans. Audio Speech Lang. Process.*, 25, 76–89.

[19] Di Giorgi, B.; Mauch, M.; Levy, M. (2020), Downbeat tracking with tempo-invariant convolutional neural networks. *Proceedings of the 21st International Society for Music Information Retrieval Conference, ISMIR 2020,* arXiv 2021, arXiv:2102.02282.

[20] Hung, Y.N.; Wang, J.C.; Song, X.; Lu, W.T.; Won, M. (2022), Modeling beats and downbeats with a time-frequency Transformer. *In Proceedings of the ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* Singapore, 23–27 May 2022; pp. 401–405.

[21] Desblancs, D.; Hennequin, R.; Lostanlen, V. (2021), *Zero-Note Samba: Self-Supervised Beat Tracking*; hal-03669865. 2022.

[22] Zonoozi, A.; Kim, J.j.; Li, X.L.; Cong, G.(2018), Periodic-CRN: A convolutional recurrent model for crowd density prediction with recurring periodic patterns. *In Proceedings of the IJCAI, Stockholm*, Sweden, 13–19 July 2018; pp. 3732–3738.

[23] Chen, C.; Li, K.; Teo, S.G.; Zou, X.; Wang, K.; Wang, J.; Zeng, Z. (2019), Gated residual recurrent graph neural networks for traffic prediction. *In Proceedings of the AAAI conference on artificial intelligence*, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 485–492.

[24] He, Z.; Chow, C.Y.; Zhang, J.D. (2019), STCNN: A spatio-temporal convolutional neural network for long-term traffic prediction. *In Proceedings of the 2019 20th IEEE International Conference on Mobile Data Management (MDM)*, Hong Kong, 10–13 June 2019; pp. 226–233.

[25] Karim, M.E.; Maswood, M.M.S.; Das, S.; Alharbi, A.G. (2021), BHyPreC: A novel Bi-LSTM based hybrid recurrent neural network model to predict the CPU workload of cloud virtual machine. *IEEE Access 2021*, 9, 131476–131495.

[26] Wu, H.; Ma, Y.; Xiang, Z.; Yang, C.; He, K. (2022), A spatial–temporal graph neural network framework for automated software bug triaging. *Knowl.-Based Syst*. 2022, 241, 108308.