Analysis and Survey on Grammar Checking and Correction Approaches for Punjabi and Other Languages

Jagdish Kaur[1], and Dr. C.P Kamboj[2]

[1]Research Scholar, Punjabi University Patiala, jagdishkaurmehrok@gmail.com

[2]Assistant Professor, Punjabi University Patiala, cpk@pbi.ac.in

ABSTRACT

Natural Language Processing (NLP) is a field under Artificial Intelligence (AI) which study human-computer interaction through computer processing of human language. A natural language consists of a large number of sentences which are linguistic units involving one or more words linked together in accordance with a set of predefined rules called grammar. Grammar checking is natural language processing task that analyzes the syntax of a corresponding language against the grammar rules of that particular language and correct those grammatical errors. NLP incorporates work in text generation by computers, text to speech and speech to text. Recently work in NLP field has been boomed up by the applications of machine learning, neural networks and deep learning. NLP is a broad topic for research which having applications for linguistics, teaching, business and other areas. One of the most known applications of NLP is using it to correct grammar. Grammatical error correction system is natural language processing task that attempts to correct are not only limited to grammatical errors, but also includes spelling error correction. Much research work has been done on grammar checker for English language. There has not been much research work done in other language like Punjabi, not due to lack of interest but due to lack of resources. In Punjabi language, there is need of more attention. Our purpose to draw the recent development of grammar checking and correction approaches used in past, present of various grammar checker for other language for developing new tool of automatic grammar checking and correction for Punjabi language.

Keywords: Artificial Intelligence, Deep learning, Transformer, Machine learning, NLP

1. INTRODUCTION

Punjabi language is the 10th most vocalized linguistic in the world with over 100 million speakers. People might make grammatical mistakes when writing some text. It becomes very inconvenience for them to check and correct grammatical errors on regular basis if the document they are writing or typing is large in size. Thus there is need of automatic grammar checker tool which automatically detect and correct such kinds of errors. The Most commonly existing grammatical errors in the Punjabi language are subject and verb agreement, modifier and noun agreement, noun and adjective, the order of verb in a verb phrase, an order of modifier of the noun in a noun phrase, punctuation errors etc. Some of errors are listed in Table 1.

A. Types of Sentences

The sentences of Punjabi language can be classified into three categories:

Simple Sentence: A Sentence that only contains one independent clause is known as simple sentence like as:

Transformer based with fine tuning and other additional features tried to get better results than existing approaches.

REFERENCES

[1] L.Jindal,V.Rana and S.K.Sharma," Punjabi Grammar Checker Components, Scope and Techniques",International Journal Of Engineering Sciences, December 2017, ISSN:2229-6913.

[2] N.Ehsan and H.Faili, "Statistical Machine Translation as a Grammar Checker For Persian Language", In the sixth International Multi-Conference on Computing in the Global Information Technology,2011.

[3] Alam, M.Jahangir, N.UzZaman and M. Khan," N-gram based Statistical Grammar Checker for Bangla and English", In Proc. Of ninth International Conference on Computer and Information Technology (ICCIT 2006).

[4] D.Naber,"A Rule-Based Style and Grammarchecker",Diplomarbeit.Technische Fakultat Bielefeld, 2003.

[5] N.H.Khan,G.C.Saha, B.Sarker and M.H.Rahman," Checking Correctness of Bangla Words using N-Gram", International Journal of Computer Applications, Vol. 89, No.11, March 2014.

[6] N.Ehsan and H.Faili," Towards grammar checker development for Persian language", IEEE International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE), 2010, pp.1-8

[7] S.Ahmadi, J.Le Roux and N.Tomeh," Attention based Encoder-Decoder Networks for Spelling and Grammatical Error Correction",Descarter University, 2017.

[8] Z.Yuan and T.Briscoe," Grammatical error correction using neural machine translation", In 2016 Conference of the North American chapter of the Association for Computational linguistics: Human Language technologies,2016,pp.380-386.

[9] S.Chollampatt and H.T.Ng," Connecting the dots: Towards Human Level Grammatical Error Correction", Proceeding of the 12th workshop on innovative use of NLP for building Educational Applications, pp.327-333, 2017.

[10] M.Gill and G.S.Lehal," A Grammar Checking System for Punjabi", Manchaster, August 2008, pp.149-152.

[11] A.Schmaltz, Y.Kim, A.M.Rush and S.Shieber," Sentence Level Grammatical error identification as sequence-to-sequence Correction", Proceeding of the 11th Workshop on Innovative Use of NLP for Building Educational Applications, 2016.

[12] Jaspreet Kaur and Kamaldeep Garg," Hybrid Approach for Spell Checker and Grammar Checker for Punjabi", International Journal of Advanced Research in Computer Science and Software Engineering Vol.4, Issue 6, June 2014.

[13] S.K.Sharma and G.S Lehal," Improving Existing Punjabi Grammar Checker", IEEE International Conference on Computational Techniques in Information and Communication Technologies(ICCTICT),PP.445-449,2016.

[14] L.Jindal and V.Rana," Grammar Checking Using Pattern Matching Approach", International Journal of Interdisciplinary Research and Innovations,Vol.6,Issue 3,pp.142-147,July-September 2018.

[15] A.Schmaltz, Y.Kim, A.Rush and S.Shieber ," Adapting Sequence Models For Sentence Correction", Proceeding of the 2017 Conference on Empirical Methods in Natural Language Processing", pp.2807-2813,2017.

[16] J.Schmidhuber, "Deep learning in neural networks: An overview", Neural Networks 61(2015), 85-117.

[17] A.Temesgen and Y.Assabie," Development of Amharic Grammar Checker Using Morphological Features of Words and N-gram Based Probabilistic Methods", Proceeding of the 13th International Conference on Parsing Technologies, pp.106-112, November 2013.

[18] N.Madi and H.S.Al-Khalifa," A Proposed Arabic Grammatical Error Detection Tool Based on Deep Learning", 4th International Conference on Arabic Computational Linguistics, November 2018,pp.352-355.

[19] J.Carlberger,R.Domeij,V.Kann and O.Knutsson," A Swedish Grammar Checker", 2002.

[20] N.Ehsan and H.Faili," Grammatical and context-sensitive error correction using a statistical machine translation framework",In wally Online Library,pp.187-206,January 2012.

[21] N.S.Bhirud, R.P.Bhavsar and B.V.Pawar," Grammar Checkers for Natural Languages: A Review", International Journal on Natural Language Computing (IJNLC) Vol.6, No.4, August 2017.

[22] Jose Silva," Grammatical Error Correction System with Deep Learning", 2018.

[23] J.Gudmundsson and F.Menkes," Swedish Natural Language Processing with Long Short-term Memory Neural Networks", 2018.

[24] T.Ge, F.Wei and M.Zhou ," Reaching Human-Level Performance In Automatic Grammatical Error Correction:An Empirical Study" ,2018.

[25] B.Manchanda,V.A.Athavale and S.K.Sharma," Various Techniques Used For Grammar Checking", International Journal of Computer Applications & Information Technology Vol.9,Issue.I, 2016.

[26] S.K.Sharma and G.S Lehal," Using Hidden Markov Model To Improve The accuracy Of Punjabi POS Tagger", IEEE International Conference on Computer Science and Automation Engineering 2,pp.697-701,2011.

[27] M.Kaur, M.Aggerwal and S.K.Sharma," Improving Punjabi Part of Speech Tagging By Using Reduced Tag Set", International Journal of Computer Applications & Information Technology,142,2014.

[28] J.Ji, Q.Wang, K.Toutanova, Y.Gong, S.Truong," Nested Attention Neural Hybrid Model For Grammatical Error Correction", Proceeding for Computational Linguistics, vol.1pp.753-762, 2017.

[29] R.Grundkiewicz and M.J.Dowmunt," Near human-level performance in grammatical error correction with hybrid machine translation", proceeding of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol.2, pp.284-290, 2018.

[30] J.Schmidhuber," Deep learning in neural networks: An overview", Neural Networks 61(2015), pp.85-117.

[31] S.Chollampatt and H.T.Ng," A Multilayer Convolutional Encoder-decoder Neural network for Grammatical Error Correction", In the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), pp.5755-5762, 2018.

[32] D.Dahlmeier, H.T.Ng and S.M.Wu," Building a large annotated corpus of learner English", The NUS Corpus of Learner English, Proceeding of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications.

[33] S.Lawrence, C.L.Giles and S.Fong," Natural Language Grammatical Inference with Recurrent Neural Networks", IEEE Transactions on knowledge and Data Engineering.

[34] S.N.Tran and A.S.A.Gareez," Deep logic networks: inserting and extracting knowledge from deep belief networks", IEEE Transactions on Neural Networks and Learning Systems, vol.29, no.2, pp.246-258, February 2018.

[35] Q.Ma," Natural language processing with Neural Networks", in Language Engineering Conference, 2002. Proceeding [online] Available: http://dx.doi org/10.1109/lec.2002.1182290.

[36] B.K.Bal and P.Shrestha,"Architectural and System Design of the Nepali Grammar Checker", www.panl10n.net/english/..../Nepal/Microsoft%20Word%20-%208_ok_N_400.pdf

[37] L.Bopche and G.Dhopavakar," Rule Based Grammar Checking System for Hindi", Journal of Information Systems and Communication, Vol. 3, Issue 1, pp.45-47, 2012.

[38] A.T.Gebru," Design and Development of Amharic Grammar Checker", 2013.

[39] S.Islam, M.F.Sarkar, T.Hussain, M.G.M.Hasan, D.M.Farid and S.Shatabda ," Bangla Sentence Correction Using Deep Neural Network Based Sequence to Sequence Learning", IEEE International Conference on Computer and Information Technology(ICCIT),pp.21-23,December 2018.

[40] Aiman Solyman[1], Zhenyu Wang[1], Qian Tao[1]," Proposed Model for Arabic Grammar Error Correction Based on Convolutional Neural Network", IEEE International Conference on Computer, Control, Electrical and Electronics Engineering (ICCCEEE19), 2019.

[41] Jiahao Wang1, Guimin Huang1,2*, and Yabing Wang1," An Automatic Grammar Error Correction Model Based on Encoder-Decoder Structure for English Texts", EAI Endorsed Transactions on Scalable Information Systems 10 2022 - 01 2023 | Volume 10 | Issue 1 | e10.

[42] N.Gupta and P.Mathu," Spell checking Techniques in NLP: A Survey", vol.2, Issue 12, December 2012, ISSN: 2277 128X, International Journal of Advanced Research in Computer Sciences and Software Engineering.

[43] B.Manchanda,V.A.Athavale and S.K.Sharma," Various Techniques Used For Grammar Checking", International Journal of Computer Applications & Information Technology Vol.9,Issue.I, 2016.

[44] A.T.Gebru," Design and Development of Amharic Grammar Checker", 2013.

[45] A.Musyafa, Y.Gao, A.Solyman , C. Wu and S.Kahan," Automatic Correction of Indonesian Grammatical Errors Based on Transformer", Applied Sciences 2022, 12, 10380. https://doi.org/10.3390/app122010380.

[46] M. Qin," A study on automatic correction of English grammar errors based on deep learning", Journal of Intelligent Systems 2022, 31: 672–680.