

Automated Essay Assessment: Hybrid Approach Using BERT and Deep Learning LSTM Models

*Agastya Seth**, *Dr Rajeev Singh ***, *Anurag Seth****

**Cadence Design Systems, Noida, India*

***Shiv Nadar University, Delhi NCR, India*

****LiqVid eLearning, India & Tembusu ICT Fund, Singapore*

<http://www.arvixe...>

Abstract- Mention the abstract for the article. An abstract is a brief summary of a research article, thesis, review, conference proceeding or any in-depth analysis of a particular subject or discipline, and is often used to help the reader quickly ascertain the paper's purpose. When used, an abstract always appears at the beginning of a manuscript, acting as the point-of-entry for any given scientific paper or patent application.

Index Terms- About four key words or phrases in alphabetical order, separated by commas. Keywords are used to retrieve documents in an information system such as an online journal or a search engine. (Mention 4-5 keywords)

I. INTRODUCTION

Automated computer based assessments have been used for quite some time now and have become the norm at much of the competitive/testing exams like SAT, GRE, JEE, IELTS, TOEFL etc. However, while this has worked well for objective/MCQ type of questions it has not become mainstream for subjective assessments like testing for language writing skills. In the post COVID19 scenario, it has become even more important to be able to apply advances in NLP and Deep Learning to build good Automated Essay Scoring engines.

AES research started in 1966 with the "Project Essay Grader (PEG)" by Ajay et al. [7]. PEG evaluated writing characteristics such as grammar, diction, construction etc to grade the essay. Shermis et al released a modified version of PEG in 2001 which focused on grammar checking with a correlation between human evaluators and the system. Folz et al. [8] introduced an Intelligent Essay Assessor (IEA) by evaluating content using latent semantic analysis to produce an overall score. Powers et al. [11] proposed E-rater and Intellimetric by Rudner et al. [10] and Bayesian Essay Test Scoring System (BESTY) by Rudner and Liang [9]. These tools employed traditional approaches like pattern matching and statistical approaches including classification methods like Bayesian Classifiers and

k-nearest-neighbors [5] on rudimentary features like number of characters in a document, number of sentences, number of words etc. Since the last decade, the essay grading systems started using regression-based and NLP techniques. AES systems like Dong et al. [12] and others developed from 2014 onwards used Deep Learning techniques inducing syntactic, semantic features resulting in better results than the previous ones.

In 2019, Jiawei Liu et al. [16] proposed an interesting 2 stage approach which combines Deep Learning (LSTM) models for semantic, coherence and prompt relevance scoring using BERT [15] embeddings in stage 1 and combining that with multiple hand-crafted features (spelling errors, grammar errors, essay length, vocabulary etc) in stage 2 with statistical machine learning methods for computing the final essay scores. We believe that the proposed approach resulted in very low explainability & intuition of esp. the prompt relevance and coherence models. We propose an alternate approach for more explainable prompt relevance and coherence models and believe that this would result in more robust and explainable essay grading system.

II. IDENTIFY, RESEARCH AND COLLECT IDEA

It's the foremost preliminary step for proceeding with any research work writing. While doing this go through a complete thought process of your Journal subject and research for it's viability by following means:

- 1) Read already published work in the same field.
- 2) Goggling on the topic of your research work.
- 3) Attend conferences, workshops and symposiums on the same fields or on related counterparts.
- 4) Understand the scientific terms and jargon related to your research work.

Identify the constructs of a Journal – Essentially a journal consists of five major sections. The number of pages may vary

depending upon the topic of research work but generally comprises up to 5 to 7 pages. These are:

- 1) Abstract
- 2) Introduct

Now it is the time to articulate the research work with ideas gathered in above steps by adopting any of below suitable approaches:

A. Bits and Pieces together

In this approach combine all your researched information in form of a journal or research paper. In this researcher can take the reference of already accomplished work as a starting building block of its paper.

Jump Start

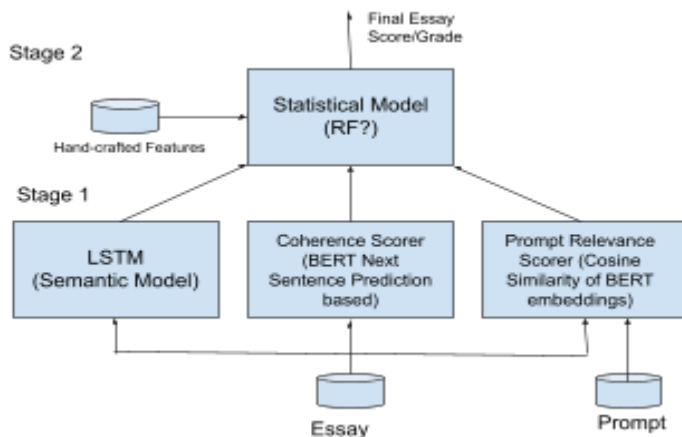
This approach works the best in guidance of fellow researchers. In this the authors continuously receives or asks inputs from their fellows. It enriches the information pool of your paper with expert comments or up gradations. And the researcher feels confident about their work and takes a jump to start the paper writing.

B. Use of Simulation software

There are numbers of software available which can mimic the process involved in your research work and can produce the possible result. One of such type of software is Matlab. You can readily find Mfiles related to your research work on internet or in some cases these can require few modifications. Once these Mfiles are uploaded in software, you can get the simulated results of your paper and it eases the process of paper writing. As by adopting the above practices all major constructs of a research paper can be written and together compiled to form a complete research ready for Peer review.

III. OUR APPROACH

Here comes the most crucial step for your research publication. Ensure the drafted journal is critically reviewed by your peers or any subject matter experts. Always try to get maximum review comments even if you are well confident about your paper.



IV. IMPROVEMENT AS PER REVIEWER COMMENTS

Analyze and understand all the provided review comments thoroughly. Now make the required amendments in your paper. If you are not confident about any review comment, then don't forget to get clarity about that comment. And in some cases there could be chances where your paper receives number of critical remarks. In that cases don't get disheartened and try to improvise the maximum.

This completes the entire process required for widespread of research work on open front. Generally all International Journals are governed by an Intellectual body and they select the most suitable paper for publishing after a thorough analysis of submitted paper. Selected paper get published (online and printed) in their periodicals and get indexed by number of sources.

V. CONCLUSION

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

APPENDIX

Appendixes, if needed, appear before the acknowledgment.

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments.

REFERENCES

- [1] Z. K. a. V. Ng, "Automated Essay Scoring: A Survey of the State of the Art," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 2019.
- [2] M. & B. J. & H. D. & Z. K. Shermis, "Automated Essay Scoring: Writing Assessment and Instruction," *International Encyclopedia of Education*, pp. 20-26, 2010
- [3] Calfee, "To grade or not to grade," *IEEE Intelligent Systems* 15(5), p. 35-37, 2000
- [4] B. L. Sevcikova, "Human versus Automated Essay Scoring: A Critical Review," *Arab World English Journal*, pp. 157-174, June 2018.
- [5] L. S. Larkey, "Automatic essay grading using text categorization techniques," in *In Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '98) ACM*, 1998
- [6] Ramesh, D., Sanampudi, S.K. An automated essay scoring systems: a systematic literature review. *Artif Intell Rev* (2021). <https://doi.org/10.1007/s10462-021-10068-2>
- [7] Ajay HB, Tillett PI, Page EB (1973) Analysis of essays by computer (AEC-II) (No. 8-0102). Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education, National Center for Educational Research and Development
- [8] Foltz PW, Laham D, Landauer TK (1999) The Intelligent Essay Assessor: Applications to Educational Technology. Interactive Multimedia Electronic

Journal of Computer-Enhanced Learning, 1, 2, <http://imej.wfu.edu/articles/1999/2/04/index.asp>

[9] Rudner, L. M., & Liang, T. (2002). Automated essay scoring using Bayes' theorem. *The Journal of Technology, Learning and Assessment*, 1(2)

[10] Rudner, L. M., Garcia, V., & Welch, C. (2006). An evaluation of IntelliMetric™ essay scoring system. *The Journal of Technology, Learning and Assessment*, 4(4)

[11] Powers, D. E., Burstein, J. C., Chodorow, M., Fowles, M. E., & Kukich, K. (2002). Stumping e-rater: challenging the validity of automated essay scoring. *Computers in Human Behavior*, 18(2), 103–134

[12] Dong F, Zhang Y, Yang J (2017) Attention-based recurrent convolutional neural network for automatic essay scoring. In: *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)* p 153–162

[13] A. R. Lahitani, A. E. Permanasari and N. A. Setiawan, "Cosine similarity to determine similarity measure: Study case in online essay assessment," *2016 4th International Conference on Cyber and IT Service Management*, 2016, pp. 1-6, doi: 10.1109/CITSM.2016.7577578.

[14] Chen, H.; He, B. Automated essay scoring by maximizing human-machine agreement. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, DC, USA, 18–21 October 2013; pp. 1741–1752.

[15] Jacob Devlin, Ming-Wei Chang, Kenton Lee & Kristina Toutanova, Google AI Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", 2019 ([arXiv:1810.04805v2](https://arxiv.org/abs/1810.04805v2))

[16] Jiawei Liu, Yang Xu, Yaquang Zhu, Automated Essay Scoring based on Two-Stage Learning, [arXiv:1901.07744v2](https://arxiv.org/abs/1901.07744v2)

AUTHORS

Agastya Seth – B.Tech (Electronics & Communication), Software Engineer, Cadence Design Systems, Noida, India, agastya.seth@gmail.com

Dr Rajeev Singh – PhD, Dean Department of Computer Science, Shiv Nadar University Delhi NCR, Dadri, UP

Anurag Seth – M.E. (Computer Science), Executive Masters in International Business, CTO & Partner, Tembusu ICT Fund, Singapore and Chief Technology Advisor, LiqVid eLearning Pvt Ltd, Noida

Correspondence Author – Agastya Seth,
agastya.seth@gmail.com, +91-888-234-0002