ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 8, Issue 10, October 2021

Integrating Cognitive Computing with Machine Learning for Big Data Analysis in Marking Digital Essay Examination

[1] Oyebola Akande, [2]Olubukola D. Adekola, [3]Wumi Ajayi, [4]Yaw Mensah, [5]Ayokunle Omotunde, [6]Adesoji Adegbola, [7]Olawale Somefun, [8]Taiwo Adigun,[9]Samuel Abel
[1][7][9] Department of Computer Science, Babcock University, Ilisan-Remo, Ogun State, Nigeria
[2][3][4][6][8] Department of Software Engineering, Babcock University, Ilisan-Remo, Ogun State, Nigeria
[5] Department of Information Technology, Babcock University, Ilisan-Remo, Ogun State, Nigeria
[1] akandeo@babcock.edu.ng, [2]adekolao@babcock.edu.ng, [3]ajayiw@babcock.edu.ng, [4]mensahy@babcock.edu.ng, [5]omotundea@babcock.edu.ng, [6]soji.adegbola@gmail.com, [7]somefuno@babcock.edu.ng, [8]taiwofisayo2002@yahoo.com, [9]abels@babcock.edu.ng

Abstract: ***Manual marking of essay type questions challenges may be voluminous answer scripts voluminous, partly subjective marking even with marking guide etc. Currently, the automated essay-marking application did not integrate cognitive models which can reduce time-consuming activities such as marking essay-type examinations. This proposed research aimed to extend the previous automated essay-marking application with cognitive-based models and use a working system to show its practicability. This research, therefore, proposed a framework on which the extension can be based. The framework incorporates inference and reasoning engine, ingestion of data from multiple heterogeneous sources and data analytics on a computer using the neuromorphic chip. Practical implementation of this will help examination bodies like West Africa Examination Council (WAEC) in marking examinations by reducing the cost of marking, reducing the rate of human error, and simplify examination logistics. The system will be evaluated with the existing automated system running on a chip based on Von Neumann Architecture.***

Index Terms: ***Cognitive models, Von Neumann Architecture, Inference and Reasoning engine, Data analytics, Neuromorphic Chip***

I. INTRODUCTION

Big data research includes its analytics, security, infrastructure, technology, and mode of transfer. This proposed research is focusing on its analytics which is presently having an unprecedented impact from the integration of machine learning (or artificial intelligence techniques) and cognitive computing. Data analytics generally means finding actionable information from data using computational analysis from the field of mathematics, statistics and computer science(artificial intelligence (deep learning and machine learning) [1] Shivam G., 2018). Major techniques for data analysis has been from the field of artificial intelligence, but the current trend which is bringing in unimaginable insight, more precise predictions or precise decision making from data is cognitive computing. Cognitive computing can reduce the time-consuming activities of humans in analyzing big data.

Cognitive computing refers to systems that learn at scale, reason with purpose and interact with humans naturally. Rather than being explicitly programmed, they learn and reason from their interactions with us and their experiences with their environment [2](Kelly, 2015). Computing trends is from tabulating era to the programming era to the cognitive era. The programming era which leads to information systems is deterministic but cognitive is probabilistic, this implies it does not just provide the answer but recommendations. Cognitive computing can aid machine-learning algorithms to go through both structured and unstructured datasets that humans cannot go through even using a lifetime of reading and research. It combines the ability of humans to think (by weighing options) and the ability of a computer to read, analyze and search at high speed [3](IBM International Technical Support Organization, 2017).

The integration of cognitive computing with machine learning techniques into digital or automated or artificial intelligent systems [4](Computer Technology Association. 2016; [1]Shivam, et al., 2018) produces better results. Cognitive computing is currently applied to health care, customer care, travel planner, review of

insurance policies, speech recognition, machine translation etc.

II. PROBLEM STATEMENT

Another aspect of human life that is very important is education; therefore, this proposed research focuses on education, in the area of automatic marking of essay-type examination. Manual marking of essay type exams is always a burden to lecturers most especially when students are too many. In addition, manual marking can be subjective, based only on the knowledge of the examiners and only the marking scheme. To overcome part of these challenges researchers developed automated multiple-choice questions (MCQ) where students provide answers using a computer keyboard, marking done is automatically and the result can be generated almost immediately. But one of the weaknesses of this type of exam is that it is not the true test of students' knowledge or understanding because it encourages students to guess for an answer and gives room for cheating. Thus, researchers came up with automated marking of essay type examinations [5] Amalia, et al., 2019). Limitations of these applications are; marking is only done through marking guide, they could only mark text and not graphs or diagrams, they only use artificial intelligence techniques but do not include human reasoning model, the highest level of accuracy seen is 83.3% and the grade is solely done by machines. This is a good level of accuracy but a better one can be achieved with the integration of cognitive computing. In this proposed project, not only marking guide will be supplied to mark but also all course materials such as lecture notes, lecture videos, and recommended textbooks.

III. RELATED WORKS

Up till now, there are very few works on automated essay marking systems. Even though there are e-learning systems that allow essay questions and answer but marking is done just by comparing keywords with students' answers, but this is not found accurate most of the time. This is because there are many variations to an answer which may not have the same key but the same semantics. Thus, these systems compare only the linguistics and not the semantics.

[5]Amalia et. al., 2019 developed an automated essay evaluation system (AEE) using latent semantic analysis (LSA). The system can mark and score essay exams written in Bahasa. The Lecturer's key answers with respective grade values and students answers were taken as input, then preprocessed, after this LSA was used to find meaning by comparing semantic similarity. Each word and each query form a vector, cosine similarity then finds similarity between the document vector and query vector. If the document vector is much similar to the query vector, then the document (student's answer) is similar to the query vector (Lecturer's answer keys). Using LSA only allows the lecturer's answer keys and not the entire marking guide with the class instructional notes.

Project Essay Grader (PEG) was the first automated essay scoring system proposed by Ellis Batten Page in 1966 [6](Chun et. al., 2021), and first programmed in 1973 [7](Page, 2003). Metrics used were trins (i.e., writing style e.g., fluency) and proxes (i.e., observable component e.g., amount of vocabulary). These two are correlated to evaluate the writing of a student. The implementation was first done on predominantly slow mainframe computers, thus further research on PEG went on sleep mode. A reawakening of PEG came with the advent of microcomputers and the internet. The internet provides a platform for submitting work for grading. At this time linguistic were applied to the workings of PEG and larger datasets were taken as input.

IntelliMetric™ is the first AES using machine learning and artificial intelligence techniques. It is also based on quantum reasoning technology [8]((Elliot, 2003) & [9] (Shermis Mark D.). It is a proprietary AES proposed by Vantage learning in the year 1998 [6] (Chun et. al., 2021). It uses multiple reasoning engines which are using different statical models and then it compares results before bringing out the final one. Till now validity is ongoing research on IntelliMetric™ but in this current research each score result will always come out indicating confidence level. Although it is very close to this current research in the use of technologies, but up till now it could only mark essay answers and it is not explicitly stated that it could mark graphs, and diagrams which can be part of essay answers in some examination.

Intelligent Essay Assessor™ (IEA) was proposed and developed by [10] Landauer, T. K., & Dumais, S. T. in 1997 for the use of marking students' essays at the University of Colorado. It was later acquired by Pearson Educational Technology in 1998 [11] (Kaja Zupanc and Zoran Bosnic´, 2015). Scores in IEA is compared with human graders instead of by correlation. IEA uses both LSA and NLP techniques. LSA is used for the semantic, that is, to evaluate the content quality while NLP is for

ISSN (Online) 2394-6849

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 8, Issue 10, October 2021

extracting essay attributes such as lexical, grammatical, mechanical, stylistic and organizational structure [6] (Chun et. Al., 2021). IEA can be trained on few datasets like 100, this is unlike other AEEs that require as much as 300 or 500 pre-scored essays to be trained [12] (Dikli, 2006). IEA is a web-based scoring service on knowledge analysis technologies KATech web servers [13] (Foltz, Peter W. et. al., 1999). For now, it cannot be deployed on a stand-alone desktop because it consumes a lot of computer resources. The current proposed work in this research is to develop an AEE that can work on the desktop.

Reference [14] Al-Shalabi, 2016, proposed a web-based AEE for scoring Arabic essays. The automated system uses stemming (heavy stemming and light stemming) and Levenshtein edit operations. From the teacher's interface, questions and correct answers are stored in the database, then from the student's interface, each question is retrieved with a form to input the answer. After filling the form, AEE compares students' answers with the correct answer to determine their similarity. Heavy stemming is done first and similarity is found by using edit distance, this is followed by light stemming and similarity is found again using edit distance.

The Electronic Essay Rater® (E-rater) takes in unedited corpora, unlike other feature analysis tools that use edited corpora [15](Burstein, 2003) . It was built and first used by Educational Testing Service in the year 1999 [12](Dikli, 2006). It consists of three NLP-based modules which are syntactic, discourse and topical. Syntactic module works on the syntactic feature, discourse works with a conceptual framework to identify features such organization of ideas and discourse-based relationship features. The last module which is the topical module identifies topical content features and vocabulary features. These were the features a human grader uses to score an essay. The result of each module is taken as input into the model building which scores the essay. Final scoring is done by an e-rater and a human. E-rater is primarily working on pure essay writing, but the proposed work will be for any type of domain essay answers e.g., physical and engineering sciences, education, social sciences, biological sciences etc.

IV. METHODOLOGY

The development process will follow the Prototyping development model, where a kind of prototype will be first built to show high-level functions of the application. Figure I below is the framework of the proposed cognitive analysis model. Firstly, is the identification of the data needed and their respective format. The data can be stored physically or logically with the use of relational databases, NoSQL or text copra. The data includes lecture notes, lecture videos, and recommended textbooks.

The data in various repositories will be extracted and preprocessed by cleaning and compression. The preprocessed data will undergo knowledge representation using first-order logic. The first-order logic will be used to build inference and reasoning engines.

The preprocessed data will be fed into the machine learning framework while the inference engine will be integrated into the cognitive process. The cognitive process of analysis will use an artificial neural network, knowledge graph, and natural language processing on a neuromorphic chip. The result will be extracted from the cognitive process and visualized. The result will lastly be compared with the existing intelligent marking system.

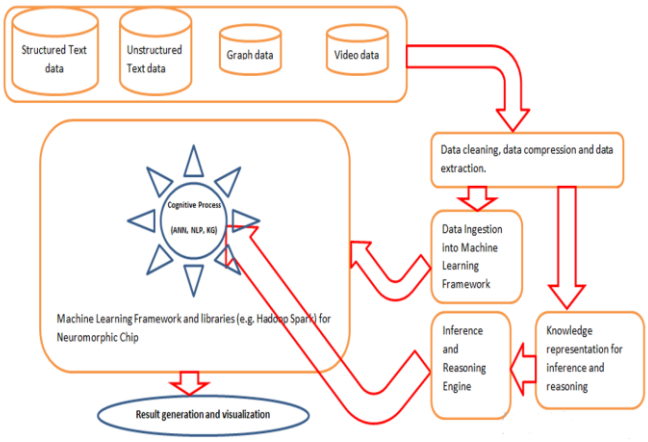

Figure I: Framework for Cognitive based Essay Exam Marking System

ANN: Artificial Neural Network, NLP: Natural Language Processing, KG: Knowledge Graph

Future Work

Implementation of this framework and its practical application in Universities and secondary schools.

CONCLUSION

To provide an intelligent essay marking application that will reduce social vices in grading students in tertiary institutions in Nigeria. This application will be useful for lecturers in the universities and will reduce the use of

paper during examinations. This application can also help an organization like West Africa Examination Council in marking and reduce the cost of marking, reduce the rate of human error, and simplify examination logistics.

REFERENCES

[1] Shivam G., Arpan Kumar K., Abdullah B., Wassan Al-Khowaiter. (2018). Big data with cognitive computing: A review of the future. International Journal of Information Management, 12.

[2] Kelly, J. (2015). *Computing cognition, and the future of knowing*. IBM Research.

[3] IBM International Technical Support Organization. (2017). *Building Cognitive Applications with IBM Watson Services: Volume 1 Getting Started.* Poughkeepsie: IBM Redbooks.

[4] Computer Technology Association. (2016). Ginni Rometty Keynote: CES 2016 (YouTube).

[5] Amalia A., Gunawan D., Fithri Y. and Aulia I. (2019). Automated Bahasa Indonesia essay evaluation with latent semantic analysis. *Journal of Physics*, 1-9. doi:10.1088/1742-6596/1235/1/012100

[6] Chun Then Lim, Chih How Bong, Wee Sian Wong and Nung Kion Lee. (2021). A Comprehensive Review of Automated Essay Scoring (AES) Research and Development. *Pertanika J. Sci. & Technol., 21*(3), 1875-1899.

[7] Page, E. B. (2003). Project Essay Grade: PEG. In S. &. M. D., & J. B. Mark D. Shermis (Ed.), *Automated Essay Scoring: A Cross-Disciplinary Perspective* (pp. 43-54). Mahwah, New Jessey: Lawrence Erlbaum Associates.

[8] Elliot, S. (2003). IntelliMetric: From here to Validity. In S. Elliot, & S. &. M. D. (Ed.), *Automated essay scoring: A cross-disciplinary perspective* (pp. 71-86). Mahwah, New Jessey: Lawrence Erlbaum Associates, Inc.

[9] Shermis Mark D. & Burstein Jill C. (2003). *Automated essay scoring: A cross-disciplinary perspective.* Mahwah: Lawrence Erlbaum Associates.

[10] Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240.

[11] Kaja Zupanc & Zoran Bosnic´. (2015). Advances in the Field of Automated Essay Evaluation. *Informatica, 39*, 383-395.

[12] Dikli, S. (2006). Automated Essay Scoring. *Turkish Online Journal of Distance Education-TOJDE, 7*, 49-62.

[13] Foltz, Peter. W., Laham, Darrel., & Landauer, Thomas. K., (1999). The Intelligent Essay Assessor: Applications to Educational Technology. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 939-944.

[14] Al-Shalabi, E. F. (2016). An Automated System for Essay Scoring of Online Exams in Arabic based on Stemming Techniques and Levenshtein Edit Operations. *IJCSI International Journal of Computer Science Issues, 13*(5), 45-50.

[15] Burstein, J. (2003). The E-rater® Scoring Engine: Automated Essay Scoring With Natural Language Processing. In J. Burstein, & S. &. Mark. D. (Ed.), *Automated essay scoring: A cross-disciplinary perspective* (pp. 113-122). Mahwah, New Jessy: Lawrence Erlbaum Associates Publishers.