ISSN No:-2456-2165

Enhancing Academic Resource Evaluation in Computer Science and Engineering through Automated Assessment

Pranshu Jain; Riya Dubey; Pankhuri Deshmukh; Manas Tiwari; Dr. Mehjabin Khatoon
Department of Computer Science and Engineering
Dr. Akhilesh Das Gupta Institute of Professional Studies Delhi, India

Abstract:- Navigating the vast amounts of digital academic content on the Internet poses a formidable challenge. Addressing this, we have formulated an academic content evaluator that leverages machine learning algorithms - Decision Tree, SVM, Random Forest and RNN. This machine-learning approach is fueled by citation rates, authorship details, and content analysis. This paper explores the model's transformative potential, delving into its features, algorithms, and the evolving landscape of academic content assessment.

Keywords:- *Academic, Computer Science, Content Evaluation, Resource Evaluation, Quality Assessment.*

I. INTRODUCTION

In the ever-growing field of Computer Science, access to reliable and high-quality academic resources is essential, both for students and scholars. The exponential growth of digital content on the web has led to an overwhelming abundance of articles and textual materials, which makes it very challenging to identify trustworthy and valuable sources. Moreover, manually evaluating the vast amount of available resources in search of good-quality materials can become a time-consuming and inefficient endeavor.

Recognizing the need for an accurate and fast assessment tool, we have formulated the Academic Content Evaluator (ACE), an innovative system which uses artificial intelligence and machine learning for the assessment of academic materials. This paper delves into a comprehensive analysis of the ACE system, exploring the merits and demerits of its approach.

With a dataset comprising over 20,000 samples from Wikipedia articles, ACE employs a diverse set of features, including title, word count, vocabulary size, number of links, latest update date and Flesch-Kincaid readability score. The machine learning algorithms at its core, including Support Vector Machine, Random Forest, Recurrent Neural Network and Decision Tree, enable ACE to discern patterns and complexities within academic materials.

The holistic approach of ACE, from its dataset analysis to future endeavors, positions it as a pioneering solution at the intersection of AI, education, and technology, promising a transformative impact on academic resource evaluation.

II. LITERATURE REVIEW

Our academic resource evaluator aims to address the pressing need for automated assessment of educational resources, particularly in the context of digital libraries.

The past research [Wetzler 2009] involves a mixed-method study, combining data from previous studies of resource quality with insights from digital library quality experts. 12 key dimensions of resource quality have been identified, such as presence of pertinent data, general quality of the set-up, use of relevant graphics, pedagogical support, etc.

It has been established that defining a benchmark for quality of resources on the Internet can be used for automating the quality assessment of resources and so, a new approach using deep learning for automatic representation of features was introduced [Wetzler 2009]. The authors propose using Doc2Vec for converting textual documents into fixed-length numerical vectors and Deep Neural Networks (DNN) for classifying Wikipedia articles based on these representations.

In a project for UKRI (UK Research and Innovation, researchers explored the potential use of artificial intelligence (AI) in the Research Excellence Framework (REF). The authors explored two common AI approaches: knowledge-based and machine learning. Given the absence of a knowledge base related to article quality, they opted for a machine-learning approach. The inputs for the machine learning system included citation rates, authorship information, journal names, article length, abstract readability,and textual content.[1]

In reference to the quality assessment of Wikipedia articles, the present approaches of classification rely on manually predefined feature sets, and the paper introduces a novel approach using deep learning for automatic feature representation [Ignat 2016]. They discuss existing approaches that rely on specific feature sets, such as article length, textual,review, and network features, as well as those considering author-related features like reputation and interaction.

Table 1: Wiki Project article quality grading scheme

Label	Requirements
B	The article must not have any big problems andshould be mostly complete. However, it needs more work to attain better status.
C	Article is meaningful but incomplete, or containing immaterial information. It must have reliable references but there can still be problemsin it, or need of a large number of improvements.
Start	Article is incomplete, or under development. Itmay not have reliable references.
Stub	It is very basic in nature and does not meet anyof the required criteria.

III. METHODOLOGY

The Academic Content Evaluator (ACE) is a machine learning model that evaluates the quality and relevance of academic content in the context of computer science. The primary dataset consists of 20,000 samples from Wikipedia articles, categorized into six academic grades(O,A,B,C,D,E). Relevant features such as title, number of words, vocabulary size, number of links, Flesch-Kincaid readability score, and latest update date are extracted from each article.

This ACE utilized the Wikipedia API to collect data on articles in the Computer Science and Software categories. The dataset was categorized into Stub, Start, C, B, A, GA, FM and FA, each assigned a label based on its quality and completeness.

Table 2: This ACE utilized the Wikipedia API to collect data on articles in the Computer Science and Software categories

Label	**Requirements**
FA	This status is assigned when the article has beenreviewed by impartial reviewers from featured article candidates and has passed the review.
A	Article must be complete and properly organized, and must be reviewed by impartial reviewers from a WikiProject or another place. Attainment of Good article status is not required.
GA	The article must be checked by impartial reviewers (one or more) from WP: Good articlenomination.

The data was labeled according to these classifications, providing the ground truth for training machine learning models. To handle limited samples in FA and A classes, FA and FM were combined into a single category (F) and GA and A into another category (A). The dataset used for training was over 20,000 samples, ensuring diversity in content types and allowing the machine learning models to generalize well to different scenarios. The selected features for the machine learning model included article title, word count, vocabulary size, links count, readability score, and last edited date. This dataset acts as the basis for the development of the machine learning models in the context of the ACE project.

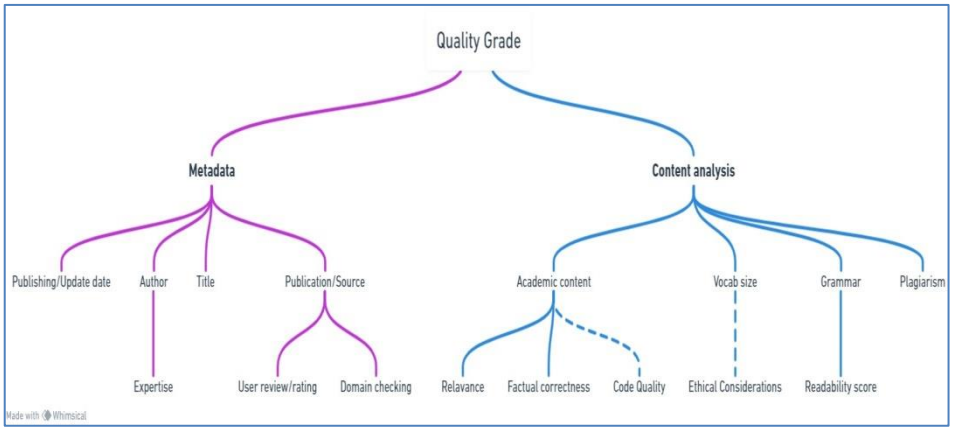

Fig. 1: Tree structure of the assessment grid

A dataset with over 20,000 samples has been used to train the model. Each sample corresponds to an article from the Computer Science and Software categories, covering a wide range of topics within the field. The primary goal of the ACE is to learn patterns and relationships within the features extracted from each article and their corresponding assessment labels. The dataset's size contributes to the diversity in data, allowing the model to accurately generalize different articles and assessment classes. Many different machine learning algorithms, such as Decision Tree, Random Forest, Recurrent Neural Network and Support Vector Machine were trained using this dataset, to capture and understand the complexities present in different types of articles. In summary, the extensive dataset of over 20,000 samples from Wikipedia forms the backbone of the machine learning model training process, providing the necessary breadth and depth for the model to develop a detailed comprehension of academic content in the domain of computer science.

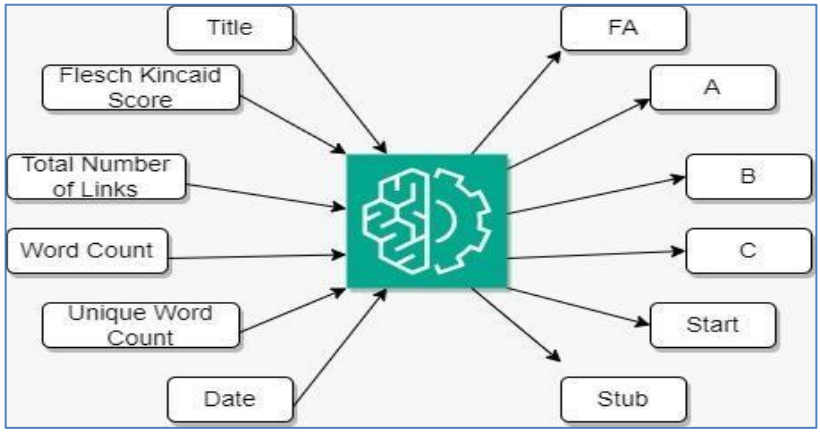

Fig. 2: ML Model makes deductions from surface-levelinformation

IV. RESULT ANALYSIS

This research analyzes the performance of four machine learning models trained on a dataset of approximately 20,000 Wikipedia articles, focusing on academic grades.

The machine learning algorithms we used include Support Vector Machine, Random Forest, Recurrent Neural Network and Decision Tree. The training dataset encapsulates a diverse set of academic content, highlighting the challenges of class imbalance.

The Decision Tree model, with a commendable accuracy of 65.04%, exhibits a balanced trade-off between precision and recall, making it an efficient tool for academic gradeprediction.

The Random Forest model, an ensemble of Decision Trees, attains an accuracy of 65.90%, demonstrating an equilibrium between precision and recall.

The Support Vector Machine (SVM) model, with an accuracy of 63%, demonstrates resilience in handling complex decision boundaries, making it suitable for tasks with intricate feature relationships.

The Recurrent Neural Network (RNN) model, designed for sequential data, achieves an accuracy of 62%, showcasing its competitiveness. The findings underscore the significance of ensemble methods, especially Random Forests, in handling the dataset's diversity. Interpretability, resilience to high-dimensional spaces, and sequential dependency capturing are key attributes, guiding future optimizations for enhanced predictive accuracy in academic resource evaluation systems.

V. CONCLUSION AND FUTURE WORKS

The paper discusses the gaps in research on Open Educational Resources (OER) adoption and presents a new perspective on Wikipedia article classification using deep learning techniques. The results suggest that this approach can improve the accuracy of automatic quality assessment of Wikipedia articles. The proposed architectural framework harmonizes a sophisticated machine learning model for academic resource evaluation, aligning with contemporary educational demands and demonstrating a significant stride in technology integration into academic practices. The website's architecture represents an innovative approach to resource evaluation, positioning it as a pioneering solution in the ever-evolving educational technology landscape.

- **User Feedback Integration:** Implementing a dynamic user feedback system, ACE will prioritize real-time updates, effective data monitoring, and version control, ensuring relevance, temporal considerations, and ethical data use.
- **Multi-Language Support:** There can be support for multiple languages, focusing on Indian languages, through language identification, translation services, community involvement, and domain-specific vocabulary.
- **Incorporate Multimedia Evaluation:** The ML model can be extended for evaluating multimedia content, adapting to the evolving nature of educational resources.

In summary, ACE is poised for continuous improvement, embracing adaptability, inclusivity, and cutting-edge capabilities in academic resource evaluation.

REFERENCES

[1]. El Mhouti A., Nasseh A. and Erradi M. (2013). Development of a Tool for Quality Assessment of Digital Learning Resources. International Journal of Computer Applications, 64(14).
[2]. Al Abri, M., & Dabbagh, N. (2018). Open Educational Resources: A Literature Review. Journal of Mason Graduate Research, 6(1), 83-104. George Mason University. ISSN: 2327-0764.
[3]. Wetzler, Philipp & Bethard, Steven & Butcher, Kirsten & Martin, James & Sumner, Tamara. (2009). Automatically assessing resource quality for educational digital libraries. WICOW'09 - Proceedings of the 3rd Workshop on Information Credibility on the Web, Co-located with WWW 2009. 3-10. 10.1145/1526993.1526997.
[4]. Quang-Vinh Dang, Claudia-Lavinia Ignat. Quality Assessment of Wikipedia Articles: A Deep Learning Approach. ACM SIGWEB Newsletter (ACM Digital Library), Association for Computing Machinery

(ACM), 2016, <10.1145/2996442.2996447>.
[5]. Shen, A., Qi, J. and Baldwin, T., 2017, December. A hybrid model for quality assessment of Wikipedia articles. In Proceedings of the Australasian Language Technology Association Workshop 2017 (pp. 43-52).
[6]. Mike Thelwall, Kayvan Kousha, Mahshid Abdoli, Meiko Makita, Emma Stuart, Paul Wilson, Jonathan Levitt article on Can artificial intelligence assess the quality of academic journal articles in the next REF?
[7]. J. E. Blumenstock. Size matters: word count as a measure of quality on Wikipedia. In Proc. of WWW, pages 1095–1096, 2008.
[8]. DANG, Q. V. AND IGNAT, C. 2016. Quality assessment of wikipedia articles without feature engineering. In JCDL. ACM, 27–30.
[9]. HU, M., LIM, E., SUN, A., LAUW, H. W., AND VUONG, B. 2007. Measuring article quality in Wikipedia models and evaluation. In CIKM. ACM, 243–252
[10]. LI, X., TANG, J., WANG, T., LUO, Z., AND DE RIJKE, M. 2015. Automatically assessing wikipedia article quality by exploiting article-editor networks. In ECIR. Lecture Notes in Computer Science, vol. 9022. 574– 580.
[11]. LE, Q. V. AND MIKOLOV, T. 2014. Distributed representations of sentences and documents. In ICML. JMLR Workshop and Conference Proceedings, vol. 32. JMLR.org, 1188–1196.