A

Synopsis

on

**ResearchReviewX**

Submitted in partial fulfillment for

**B. Tech Project-I**

Submitted by

**Pratap Shingane(2020BTEIT00050)**

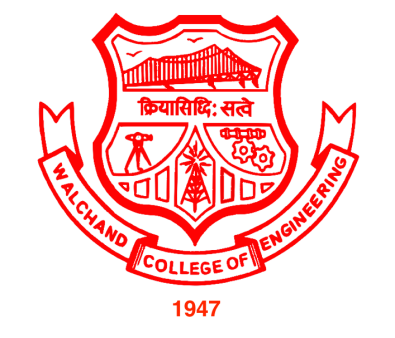
**Chinmayee Hankare(2020BTEIT00008)**

**Rajeshwari Joshi(2020BTEIT000033)**

Under the guidance of

**Dr. A. J. Umbarkar**





Department of Information Technology,

**Walchand College of Engineering, Sangli.**

Maharashtra, India. 416415

**Synopsis of Major Project-I**

1. **Introduction:** Automated web application for objective research paper quality assessment.
2. **Relevance/Importance of proposed work:**

* Significantly relevant in the academic and research domain.
* Addresses the need for an efficient and reliable assessment system.
* Provides researchers and readers with valuable insights for informed decisions. Streamlines the evaluation process using data preprocessing and analysis.
* Utilizes machine learning algorithms to enhance the quality of assessments.
* Contributes to maintaining the integrity of scholarly literature.

1. **Problem Statement:** The challenge is to create an automated web application that objectively evaluates the quality of research papers. The system should perform data preprocessing, analyze the papers based on predefined criteria, and generate quality ratings, enabling researchers and readers to make informed decisions about the reliability and significance of the literature they encounter.
2. **Objective:**

* Develop an automated web application for objective research paper quality assessment.
* Implement data preprocessing techniques, including text cleaning and Tokenization.
* Analyze papers using predefined parameters like verb analysis, readability, and language complexity.
* Generate quality ratings through machine learning algorithms for informed decision-making.
* Ensure user-friendliness, scalability, and integration of image-to-text conversion and paragraph separation modules.

1. **Technology Stack:**

Front-end: Streamlit

Backend: Python

Libraries: Tesseract OCR, OpenCV, Scikit-learn, TensorFlow, PyTorch

1. **Methodology:**

The methodology for the research paper quality assessment project involves several key steps. First, we will develop a web-based application using Streamlit for the frontend and Python for the backend to provide a user-friendly interface for researchers and readers to upload their papers. Next, we will implement data preprocessing techniques, including text cleaning, lowercasing, tokenization, stopword removal, and stemming, to prepare the text data for analysis. The papers will then undergo an objective evaluation using predefined parameters, such as verb analysis, readability scores, and language complexity, to quantify their quality. Machine learning algorithms, such as regression models or decision trees, will be employed to analyze the data and generate quality ratings. Additionally, we will integrate modules for image-to-text conversion and accurate paragraph separation to handle diverse file formats. The system's performance and accuracy will be validated through extensive testing, ensuring the reliability of the quality assessment process.

1. **References:**
2. Cai, L and Zhu, Y 2015 The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. Data Science Journal, 14: 2, pp. 1-10, DOI: <http://dx.doi.org/10.5334/dsj-2015-002>
3. S. Juddoo and C. George, "A Qualitative Assessment of Machine Learning Support for Detecting Data Completeness and Accuracy Issues to Improve Data Analytics in Big Data for the Healthcare Industry," 2020 3rd International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM), Balaclava, Mauritius, 2020, pp. 58-66, doi: 10.1109/ELECOM49001.2020.9297009.
4. Thelwall, M., Kousha, K., Wilson, P., Makita, M., Abdoli, M., Stuart, E., Levitt, J., Knoth, P., & Cancellieri, M. (2023). Predicting article quality scores with machine learning: The U.K. Research Excellence Framework. Quantitative Science Studies, 4(2), 547–573. https://doi.org /10.1162/qss\_a\_00258
5. Systematic literature review of machine learning methods used in the analysis of real-world data for patient-provider decision making (Brnabic and Hess BMC Med Inform Decis Mak <https://doi.org/10.1186/s12911-021-01403-2> )
6. Outcome of project: Paper publication, real world problem solving

**Name and Signature**

1. Pratap Shingane(2020BTEIT00050)
2. Chinmayee Hankare(2020BTEIT00008)
3. Rajeshwari Joshi(2020BTEIT00033)

|  |  |  |
| --- | --- | --- |
| Name and Signature  Guide | Name and Signature  Panel Member | Name and Signature  HoD |