



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume:** 12    **Issue:** V    **Month of publication:** May 2024

**DOI:** <https://doi.org/10.22214/ijraset.2024.60649>

[www.ijraset.com](http://www.ijraset.com)

Call:  08813907089

E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)

# Predicting the Students Performance Using WUM and Machine Learning Technique

Chandan M<sup>1</sup>, Amrith<sup>2</sup>, Athish Raj Mohan<sup>3</sup>, Chandan M V<sup>4</sup>, Prashanth Kumar K N<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup>Department of Computer Science & Engineering, Bangalore Institute of Technology, Bengaluru, India

**Abstract:** This project aims to develop a predictive model for assessing students' academic performance by integrating web usage mining (WUM) and machine learning (ML) techniques. The model leverages user-provided data such as gender, parental profession, reading score, and writing score, along with behavioural insights from online educational platform interactions. The methodology includes data preprocessing, feature extraction, and adaptive algorithm selection to enhance prediction accuracy and adaptability across various student profiles. The integration of WUM provides a robust framework for educational institutions to tailor interventions and support systems effectively.

**Keywords:** Web Usage Mining (WUM), Machine Learning (ML), Predictive Model, Educational Data Analysis, Adaptive Algorithm Selection

## I. INTRODUCTION

In the rapidly evolving landscape of education, understanding and predicting students' academic performance is crucial for educators and institutions seeking to provide targeted support and interventions. Traditional methods of assessment often fall short in capturing the diverse factors influencing student success. Education has increasingly embraced digital platforms for learning, with students engaging in a variety of online activities. Web usage mining, a process of extracting valuable patterns and information from users' interactions with web applications, offers a unique opportunity to gain insights into student behaviour. By combining this with machine learning algorithms, which have demonstrated prowess in predictive modelling, we aim to create a robust and adaptive system for predicting academic outcomes. Our model considers key demographic factors such as gender and parental profession, along with academic indicators like reading and writing scores, as input features. The integration of web usage mining enriches this dataset with behavioural patterns obtained from students' interactions with online educational resources. What sets this research apart is the dynamic selection of the most suitable machine learning algorithm for each prediction instance, ensuring adaptability to the evolving nature of educational data.

## II. OBJECTIVES

- 1) To Accomplish a robust predictive model by integrating Web Usage Mining features with relevant data.
- 2) To minimize data noise and inconsistencies by preprocessing the collected dataset.
- 3) To build a web usage mining pipeline that extracts relevant features from log files and web data.
- 4) To design a system for predicting students' performance using novel clustering algorithms.
- 5) To implement a linear regression and other models, utilizing the best among them to cluster features and predict the score with high accuracy

## III. LITERATURE SURVEY

This paper [1] provides a general idea of web mining and its types. The review paper mainly focuses on the web usage mining and methods utilized in it. Also it gives a review on the application of web usage mining.

This paper [2] was concerned with the discovery of the customers' access patterns when browsing the Internet, By breaking down information source from the Internet and incorporating it with information from conventional information stores a degree of information beforehand unrealistic can be reached. The Outcome from this paper is that, using Web Sessions personalized web pages are identified and find some important information of user from the raw access logs and this information has been get to improved E-commerce sites.

This paper [3] proposes a system which aims at identification and analysis of learning styles and emotional behaviour of users based on the techniques of web usage mining where web server logs are used for data capturing as they have many advantages over other techniques. The logs are further preprocessed and used for extraction of desired results using different machine learning methodologies.

This paper [4] proposes the strategies for discarding the old sessions for storing new sessions. It chooses the sessions which are going to be kept for prediction so that the prediction time is reduced with improved accuracy.

In this paper [5] the study highlights the potential of integrating Machine Learning (ML) into User Experience (UX) design, with most UX designers lacking prior ML expertise. Addressing this gap through specialized education and collaborative efforts can lead to more innovative, user-centric applications, improved user feedback, and significant time and cost savings in the early design stages. Limited understanding in design practice and incomplete and early-stage research.

In this paper [6] the proposed system aims to prioritize User Experience (UX) as a central element in any business solution, including web applications, websites, and mobile apps. It acknowledges that a good solution not only addresses the exact problem but also ensures that end users find it user-friendly. The system focuses ongoing user feedback and delivering self-understandable User Interfaces to enhance productivity and customer satisfaction. Lack of appropriate name for UI Element and varying perceptions of developer and customer.

In this paper [7] provides a general idea of presenting various machine learning (ML) algorithms to predict and analyse the current user behaviours. The main objective of this work is discriminating and classifying the close group to which user is most interested.

In the paper [8] Web usage mining has seen a quick expansion in interest from both the examination and practice perspective in networks, considering its application potential. This paper provides a comprehensive taxonomy, counting research endeavours of the work in the same field. An up-to-date survey is also given on the ongoing work.

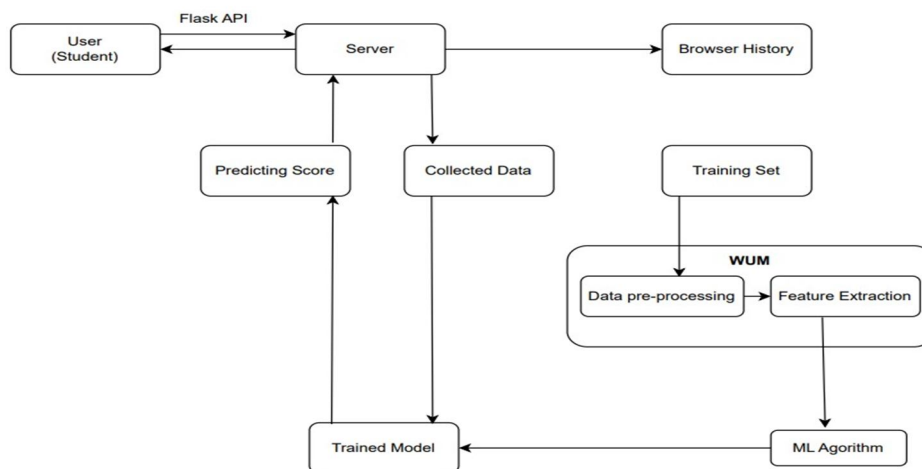
In this paper [9] provides general idea of the behaviour classification-based e-learning performance (BCEP) prediction framework, which selects the features of e-learning behaviours, uses feature fusion with behaviour data according to the behaviour classification model to obtain the category feature values of each type of behaviour, and finally builds a learning performance predictor based on machine learning.

In the paper [10] the main idea is to combine Simple K-means clustering and the Apriori association rule algorithm to offer students highly effective and user-friendly course recommendations. Moreover, the system provides a rich set of association rules, enabling a deeper understanding of students' preferences and enabling informed course selections.

#### IV. PROPOSED SYSTEM

- 1) The proposed system for predicting students' performance integrates Web Usage Mining (WUM) and Machine Learning techniques to create a robust and adaptive framework. It utilizes demographic features such as gender, parental profession, reading score, and writing score, combined with behavioural insights obtained from web usage patterns.
- 2) The system employs dynamic algorithm selection to continuously optimize the choice of machine learning algorithms, ensuring adaptability to diverse datasets.
- 3) Through iterative model training and refinement, the system aims to enhance prediction accuracy, providing valuable insights for educators and institutions to tailor interventions and support systems effectively.
- 4) Machine Learning Classifiers such as Random Forest, Logistic regression to be used for performance prediction.

#### V. BLOCK DIAGRAM



## VI. COMPONENT DESIGN

- 1) *Data pre-processing*: The quality of e-learning behaviour data directly affects the accuracy of predictive models. Therefore, the first step is to clean the e-learning behaviour data obtained from the e-learning platform. The method should be selected according to the real situation of the data to manage missing values, duplicate values, and abnormal values.
- 2) *Feature selection*: Feature selection can select relevant features that are beneficial to the training model from all features, thereby reducing the feature dimension and improving the generalizability, operating efficiency and interpretability of the model. This framework uses the variance filtering method to perform feature selection on standardized e-learning behaviour data.
- 3) *Feature fusion*: This classifies core learning behaviours according to specific rules, constructs a collection of behaviour categories, and then performs feature fusion to obtain the category feature value of each type of e-learning behaviour.
- 4) *Model training*: In the model training session, classic machine learning methods such as SVC, Naive Bayes, Random Forest are selected, and the e-learning behaviour category feature value set is used as the feature data to train the e-learning performance prediction model. After many iterations, the best e-learning performance prediction model is selected to predict the e-learning performance of e-learners.

## VII. CONCLUSION

- 1) Testing the proposed approach on larger and more diverse datasets to evaluate its performance in different scenarios
- 2) Investigating the potential of the approach for other types of data analysis tasks, such as classification or clustering
- 3) Exploring the use of other optimization algorithms or machine learning techniques to improve the efficiency and accuracy the feature selection and prediction process.
- 4) Develop an interactive and user-friendly system that implements the proposed approach and can be used by students and learners.

## REFERENCES

- [1] Dr. Rajesh K Shukla, Prachi Sharma, Noopur Samaiya, Monika Kherajani, "Web Usage Mining- A study of Web Data Pattern detecting and Methodologies and its Applications" International Journal on Computer Engineering and Intelligent Sstems, vol.12, issue 1, September 2020.
- [2] Jayanti Mehra, "Web Personalization Using Web Session for Web Usage Mining", International Journal Conference Computer Engineering and Intelligent Systems, vol.13, issue 1, September 2020.
- [3] Snehal Rathi , Yogesh Deshpande , Shashidhar Nagaraj, Ankita Narkhede, Radhika Seiwani , Varad Takalikar , "Analysis of User's Learning Styles and Academic Emotions through Web Usage Mining" -2021 International Conference on Emerging Smart Computing and Informatics (ESCI), March 2021.
- [4] Shivani Yadao, A. Vinaya Babu, Midhunchakkaravarthy Janarthan, Amiya Bhaumik, "Web usage Mining: A Comparison of WUM Category Web Mining Algorithms", in Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp 1020-1024.
- [5] M. Pooja Bharti, J. Tushar Raval, "Improving Web Page Access Prediction using Web Usage Mining and Web Content Mining", in 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), 2019.
- [6] Khairil Imran Ghauth, Abdallah M. H. Abbas, Khairil Imran Ghauth, Choo-Yee Ting "User Experience Design Using Machine Learning: A Systematic Review" Institute of Electrical and Electronics Engineers (IEEE), January 2022.
- [7] Sumaiya PK "Enhancing User Experience using Machine Learning" International Journal of Engineering Research & Technology, 2018.
- [8] Ashwini, K Viswvardhan Reddy "Predicting the User Behaviour Analysis using Machine Learning Algorithms", International Research Journal of Engineering and Technology, 2020.
- [9] Feiyue Qiu, Guodao Zhang, Xin Sheng "Predicting students' performance in e-learning using learning process and behaviour data", Scientific Reports, 2022 [10] S.B. Aher, L.M.R.J. Lobo, "Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data", Elsevier, 2013.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)