IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# PERFORMANCE ANALYSIS ON STUDENT FEEDBACK USING SUPERVISED MACHINE LEARNING

*Mrs. S. Steffi Nivedita Ballari Institute of Technology & Management*

R Padmapriya *Ballari Institute of Technology & Management*

Sheetal V *Ballari Institute of Technology & Management*

Sushma Pallavi V *Ballari Institute of Technology & Management*

Vibha Hiremath *Ballari Institute of Technology & Management*

## ABSTRACT

Students feedback is an effective tool for teacher's evaluation resulting in faculty development. It investigates opinion mining by means of supervised learning techniques to search out the emotion of the students with their different perspective characterized choices of instructing and improve the quality of teaching. The proposed project is designed with a feature that only the students with registered mail having 85 percentage of attendance can give the feedback. The system maintains the feedback given by the students which is confidential and it presents in the form of pictorial representation. The feedback is also sent to the individual faculty member which informs about their strengths and weaknesses in a confidential manner.

.

## 1.INTRODUCTION

This project is exclusively for the Ballari institute of technology and management students. Here students will give feedback on faculty's performance in an effective evaluation tool resulting in faculty development. The proposed project is designed with a feature that only the students with registered mail having 85 percentage of attendance can give the feedback. The system maintains the feedback given by the students which is confidential and the presents it in the form of pictorial representation. The feedback is also sent to the individual faculty member which informs about their strengths and weaknesses in a confidential manner. This estimates the effect of various machine learning methods and algorithms. Algorithms are applied in generating predictive methods, which are decision trees, naïve bayes classifier and k- nearest neighbour classifier and support vector machines for the prediction of faculty performance. This analysis also estimates the increase caused by feature engineering, which leads to transforming the dataset to present it more fitting for machine learning.

## I. LITERATURE SURVEY

**In paper [1]** This paper explores opinion mining using supervised learning algorithms to find the polarity of the student feedback based on pre-defined features of teaching and learning. The study conducted involves the application of a combination of machine learning and natural language processing techniques on student feedback data gathered from module evaluation survey results of Middle East College, Oman. In addition to providing a step-by-step explanation of the process of implementation of opinion mining from student comments using the open-source data analytics tool Rapid Miner, this paper also presents a comparative performance study of the algorithms like SVM, Naïve Bayes, K Nearest Neighbor and Neural Network classifier. The data set extracted from the survey is subjected to data preprocessing which is then used to train the algorithms for binomial classification. The trained models are also capable of predicting the polarity of the student comments based on extracted features like examination, teaching etc. The results are compared to find the better performance with respect to various evaluation criteria for the different algorithms.

**In paper [2]** This paper explores opinion mining using supervised learning algorithms to find the polarity of the student feedback based on pre-defined features of teaching and learning. The study conducted involves the application of a combination of machine learning and natural language processing techniques on student feedback data gathered

from module evaluation survey results of Middle East College, Oman. In addition to providing a step-by-step explanation of the process of implementation of opinion mining from student comments using the open-source data analytics tool Rapid Miner, this paper also presents a comparative performance study of the algorithms like SVM, Naïve Bayes, K Nearest Neighbor and Neural Network classifier. The data set extracted from the survey is subjected to data preprocessing which is then used to train the algorithms for binomial classification. The trained models are also capable of predicting the polarity of the student comments based on extracted features like examination, teaching etc. The results are compared to find the better performance with respect to various evaluation criteria for the different algorithms.

**In paper [3]** The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia

**In paper [4]** Data mining involves the searching of large information of the data or records to discover patterns and utilize these patterns in the prediction the future events. In most educational sectors such as high schools, polytechnics and universities; classification technique is a vital analytical mechanism in prediction of various levels of accuracy. Classification is one of the methods in data mining for categorizing a particular group of items to targeted groups. Main goal of classification is to predict the nature of an items or data based on the available classes of items. Construction of the classification model always defined by the available training data set. In this paper we will only discuss about the classification algorithms, although there are different types of algorithms available in data mining for the prediction of the future strategy for a business. The decision tree classification technique utilized in this work focused mainly on data of the student's performance obtained in a high school during a quiz using the KNIME tool.

**In paper [5]** In the academic industry, students' early performance prediction is important to academic communities so that strategic intervention can be planned before students reach the final semester. This paper presents a study on Artificial Neural Network (ANN) model development in predicting academic performance of engineering students. Cumulative Grade Point Average (CGPA) was used to measure the academic achievement at semester eight. The study was conducted at the Faculty of Electrical Engineering, University Technology MARA (UiTM), Malaysia. Students' results for the fundamental subjects in the first semester were used as independent variables or input predictor variables while CGPA at semester eight was used as the output or the dependent variable. The study was done for two different entry points namely Matriculation and Diploma intakes. Performances of the models were measured using the coefficient of Correlation R and Mean Square Error (MSE). The outcomes from the study showed that fundamental subjects at semester one and three have strong influence in the final CGPA upon graduation.

**In paper [6]** Data is an asset; it abounds and is everywhere! A key ingredient to the flourishing of an Academic's/ University is its data asset and therefore the way much of such an asset is used to comprehend useful insight into the quality of the education and techniques to boost the performance of the organization. the need to implement a sustainable decision model for the expansion of the scholar performance level of the institution, supported feedback received from the information gathered over time, cannot be overemphasized. In this, I'd be that focus on the thanks to build an awfully simple prediction model in Python, using the k-nearest neighbor's (KNN) algorithm for classification. This helps the administration to identify and train the students in the specified area in which the skills to be improved. it can be easily classified and can be understand the scores of the students' performance level and the accuracy is given up to 1.0 (100%) since we have used our own data set which is easy to clean, analyze, visualization, extraction of the features and the classification. If we perform the operation on the any other clumsy data set the accuracy level is about 0.97 (97%), hence it is the greatest advantage of our research.

**In paper [7]** This work presented two prediction models for the estimation of student's performance in final examination. The work made use of the popular dataset provided by the University of Minho in Portugal, which relate to the performance in math subject and it consists of 395 data samples. Forecasting the performance of students can be useful in taking early precautions, instant actions, or selecting a student that is fit for a certain task. The need to explore better models to achieve better performance cannot be overemphasized. Most of earlier work on the same dataset used K-Nearest Neighbor algorithm and achieved low results, while Support Vector Machine algorithm was rarely used, which happens to be a very popular and powerful prediction technique. To ensure better

comparison, we applied both Support Vector Machine algorithm and K-Nearest Neighbor algorithm on the dataset to predict the student's grade and then compared their accuracy. Empirical studies outcome indicated that Support Vector Machine achieved slightly better results with correlation coefficient of 0.96, while the K-Nearest Neighbor achieved correlation coefficient of 0.95.

**In paper [8]** Although the educational level of the Portuguese population has improved in the last decades, the statistics keep Portugal at Europe's tail end due to its high student failure rates. In particular, lack of success in the core classes of Mathematics and the Portuguese language is extremely serious. On the other hand, the fields of Business Intelligence (BI)/Data Mining (DM), which aim at extracting high-level knowledge from raw data, offer interesting automated tools that can aid the education domain. The present work intends to approach student achievement in secondary education using BI/DM techniques. Recent real-world data (e.g. student grades, demographic, social and school related features) was collected by using school reports and questionnaires. The two core classes (i.e. Mathematics and Portuguese) were modeled under binary/five-level classification and regression tasks. Also, four DM models (i.e. Decision Trees, Random Forest, Neural Networks and Support Vector Machines) and three input selections (e.g. with and without previous grades) were tested. The results show that a good predictive accuracy can be achieved, provided that the first and/or second school period grades are available. Although student achievement is highly influenced by past evaluations, an explanatory analysis has shown that there are also other relevant features (e.g. number of absences, parent's job and education, alcohol consumption). As a direct outcome of this research, more efficient student prediction tools can be be developed, improving the quality of education and enhancing school resource management.

**In paper [9]** The ability of predicting a student's performance is very important in university-level distance learning environments. The scope of the research reported here is to investigate the efficiency of machine learning techniques in such an environment. To this end, a number of experiments have been conducted using five representative learning algorithms, which were trained using data sets provided by the "informatics" course of the Hellenic Open University. It was found that learning algorithms could enable tutors to predict student performance with satisfying accuracy long before final examination. A second scope of the study was to identify the student attributes, if any, that mostly influence the induction of the learning algorithms. It was found that there exist some obvious and some less obvious attributes that demonstrate a strong correlation with student performance. Finally, a prototype version of software support tool for tutors has been constructed implementing the Naive Bayes algorithm, which proved to be the most appropriate among the tested learning algorithms.

**In paper [10]** This study tested a structural equation model to estimate the relationship between health behaviors, body mass index (BMI), and self-esteem and the academic achievement of adolescents. The authors analyzed survey data from the 2000 study of Youth in Iceland , a population-based, cross-sectional sample of 6,346 adolescents in Iceland. The model demonstrated good fit with chi-square of 2685 (n = 5,810, df = 180), p < .001, Comparative Fit Index value of .94, and a root mean square error of approximation of .049. Lower BMI, physical activity, and good dietary habits were all associated with higher academic achievement; however, health behavior was positively and robustly associated with greater self-esteem. Self-esteem was positively influenced both through physical activity (beta = .16) and the consumption of fruits and vegetables (beta = .14). In contrast, poor dietary habits negatively influenced self-esteem and academic achievement, and self-esteem was negatively influenced by increasing levels of BMI (beta = -.05).

# III.ALGORITHMS

1. KNN ALGORITHM:

The k-nearest neighbours (**KNN**) **algorithm** is a simple, supervised machine learning **algorithm** that can be used to solve both classification and regression problems. It's easy to implement and understand, but has a major drawback of becoming significantly slows as the size of that data in use grows.

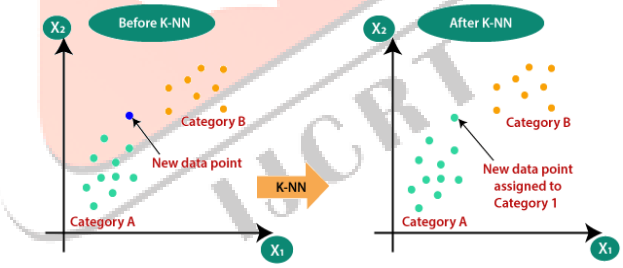



Fig1:KNN Algorithm

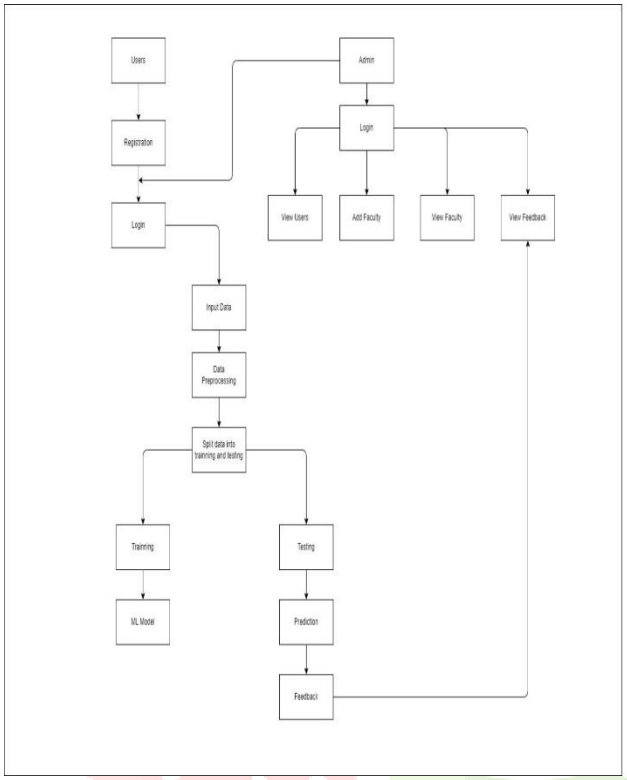
2. NAIVE BAYES ALGORITHM:

Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object

# IV. PROPOSED METHODOLOGY

To conduct performance analysis on students' feedback, you can follow a systematic methodology that includes the following steps:

Define Objectives: Clearly define the objectives of the performance analysis. Determine what you aim to achieve with this analysis, such as identifying areas for improvement, evaluating teaching methods, or measuring overall student satisfaction.




Select Metrics: Choose appropriate metrics to assess performance based on the objectives. Some common metrics for analyzing student feedback include overall satisfaction rating, specific aspect ratings (e.g., course content, instructor effectiveness), open-ended comments analysis (e.g., sentiment analysis, thematic analysis), and comparison of feedback across different courses or semesters.

Data Collection: Collect feedback data from students. This can be done through surveys, questionnaires, focus groups, or interviews. Ensure that the data collection method is appropriate for the objectives and provides a representative sample of the student population.

Data Pre-processing: Clean and prepare the collected data for analysis. This may involve removing incomplete or duplicate responses, anonymizing the data, and transforming qualitative data (e.g., comments) into a format suitable for analysis (e.g., coding or sentiment scoring).

Quantitative Analysis: Conduct quantitative analysis to examine numerical metrics. Calculate average ratings, perform statistical analysis (e.g., t-tests, ANOVA) to compare groups or time periods, and visualize data using charts or graphs to identify patterns and trends.

Qualitative Analysis: Analyze open-ended comments or qualitative data to gain deeper insights. Use techniques like sentiment analysis to determine the overall sentiment of comments, and employ thematic analysis to identify recurring themes or topics in the feedback.

Interpretation and Insight Generation: Interpret the analyzed data and generate meaningful insights. Relate the findings to the initial objectives and consider the implications for improving teaching practices or addressing student concerns. Look for actionable recommendations that can be implemented based on the analysis.

Report and Presentation: Summarize the analysis findings in a report or presentation format. Clearly communicate the objectives, methodology, results, and recommendations to relevant stakeholders, such as faculty members, administrators, or instructional designers.

Implement Improvements: Work collaboratively with relevant stakeholders to implement the identified recommendations and improvements based on the analysis. Monitor the impact of the changes over time and consider conducting regular performance analysis to track progress.

Continuous Improvement: Incorporate the feedback analysis process into an ongoing cycle of continuous improvement. Regularly revisit the objectives, metrics, and analysis methods to refine and enhance the performance analysis process.

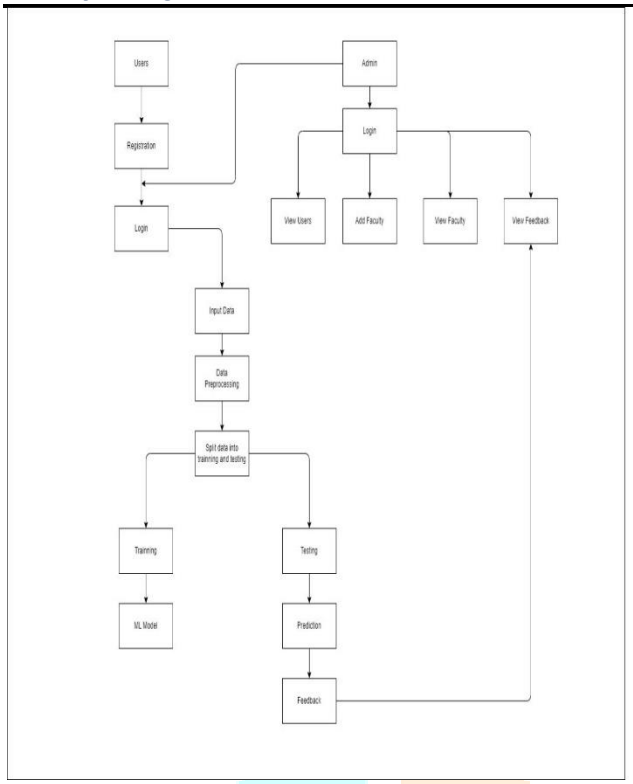
Remember, the proposed methodology can be adapted based on the specific context and resources available, but it provides a general framework for conducting performance analysis on students' feedback.

# V.SYSTEM ARCHITECTURE

Fig2: Flow Chart

# VI.PROPOSED SYSTEM

To design and develop a web application for analyzing student's feedback based on Sentimental Analysis using SVM Classifier.




## VII.OBJECTIVES

1. To enable students to give feedback.
2. To analyze student feedback.
3. To provide access the analyzed feedback to administration.
4. To display feedback in graphical representation.

## VIII.DESCRIPTION OF MODULES

### 1. System:

- Takes Dataset: The system allows users to upload the Dataset.
- Store Dataset: The System stores the dataset given by the user.
- Model selection: The system takes the data from the user and fed that data to the selected model.
- Model Predictions: The system takes the data given by the user and predict the output based on the given data.

### 2. User:

- Registration: If the user is new he/she will provide the details for the registration.
- Login: User can provide valid details for the login.
- Load Dataset: The user can load the dataset he/she want to work on.
- View Dataset: The User can view the dataset by clicking the view dataset module.
- Select model: User can select the model provided by the system for accuracy.
- Predictions: User can enter random values for prediction.
- Feedback: User can enter the feedback when the teachers performance is not satisfactory.

### 3. Admin

- Log in: Admin can enter the valid details for the login.
- View Users: Admin can see the users who used the application.
- View Remarks: Admin can see the remarks entered by the students.
- View Feedback: Admin can see the feedback given by the student.

## IX.RESULTS

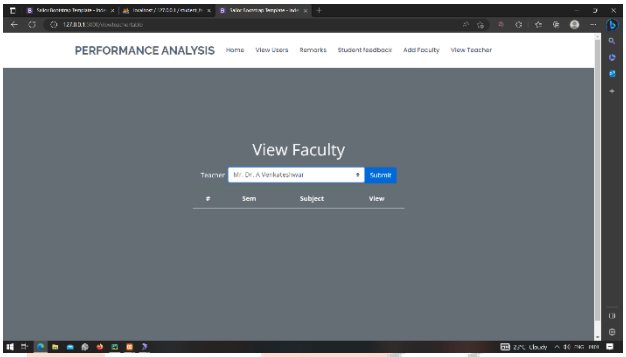



Fig:3 Viewing Faculty

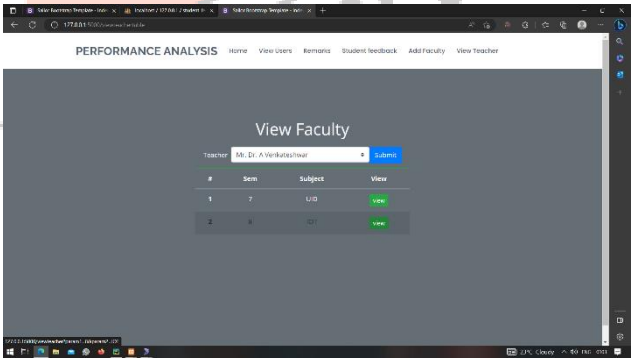



Fig4:View Subjects handled by each faculty

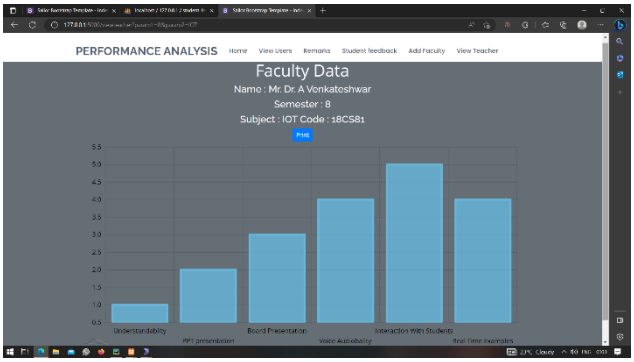



Fig5:Faculty data in pictorial form

## X.CONCLUSION

In this application, we have successfully created a ML models to estimate whether the Teacher performance is good or poor or fair. This is developed in a user friendly environment using Flask via Python programming. We noticed that out of Naïve Bayes Classifier, K Nearest Neighbors Classifier and Support Vector Classifier Naïve Bayes Classifier performs well with better accuracy.

## XI.REFRENCES

[1] V. Dhanalakshmi ; Dhivya Bino "Opinion mining from student feedback data using supervised learning algorithms", 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC).

[2] Sarpreet Kaur1 , Rasleen Deol," Students Feedback for Mining Their Opinions Using Supervised Learning Algorithm" International Journal of Engineering Science and Computing, June 2017.

[3] Vaibhav Jain, Evaluating and Summarizing Student's Feedback Using Opinion Mining, International Journal of - Scientific Research in Computer Science and Engineering.

[4] Anushree Chitriv,Prof.A.Thomas Analysis of Student Feedback by Using Data Mining Technique Helix 2018.

[5] Khan, Khairullah , "Mining opinion components from unstructured reviews: A review," Journal of King Saud University – Computer and Information Sciences , Vol. 26, 2016.

[6] Pang, B. and Lillian L. s.l, "Opinion mining and sentiment analysis : Foundations and trends in information retrieval", Vol. 2, 2016.

[7] Kumar Ravi, Vadlamani Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications", Published in Knowledge-Based Systems Vol. 89 , 14–46, 2017.S.Vijayarani,"Preprocessing Techniques for Text Mining –An Overview", International Journal of Computer Science & Communication Networks, Vol 5(1),7-16. 143–147

[8] S.Vijayarani , "Preprocessing Techniques for Text Mining - An Overview", International Journal of Computer Science & Communication Networks, Vol 5(1),7-16.

[9] Sunghwan MacKim and Rafael, "Sentiment Analysis in Student Experiences of Learning", Available at ResearchGate.com

[10] Trisha Patel, "Sentiment Analysis of Parents Feedback for Educational Institutes",International Journal of Innovative and Emerging Research in Engineering, Volume 2, Issue 3, 2017.