

RESEARCH PAPER

Online Student Feedback System Using Machine Learning

Haider Abdula Haddad

Department of Computer Science & Information Technology, Salahaddin University, Erbil, Kurdistan Region, Iraq

ABSTRACT:

In order to develop plans to enhance the teaching experience, student feedback data analysis is a very good tool to enhance the relationship between teachers and students.

This research is to present an analytical model for data from student feedback systems to improve the quality of teaching in academic institutions and universities. The developed system in this research uses the lexical analysis algorithm Support Vector Machine (SVM), which has the best accuracy and is one of the machine learning algorithms that can provide textual feedback and useful insights into the overall quality of teaching to improve teaching performance.

In this research an online system for student feedback was created. The system is used to get feedback from students about teachers and their methods of teaching. The system uses a large database to collect a large dataset from all students at different colleges at the university level. The system administrators include staff on the college levels from all colleges. All students will be provided with unique usernames and passwords to log in to the system.

Among the main tasks for the system administrator is to create classes and to create feedback questions that are designed in two questionnaire forms. The first questionnaire form is about academic questions that are related to the quality of teaching the academic subject. The second questionnaire form is the questions that are related to general education for students. The textual analysis in this system is provided using the (SVM) lexical analysis algorithm, which has the best accuracy but it requires more training time for large data sets to classify the text.

The student feedback system developed and used in this research proved to be an excellent tool to improve the academic and educational status of the university. It also helps reduce manual labor in collecting, storing and analyzing feedback data. This system is an efficient way to provide qualitative feedback to teachers that improves student-learning performance.

KEY WORDS: Feedback system, (SVM) algorithm, machine learning, university dataset, online system

DOI: <http://dx.doi.org/10.21271/ZJPAS.35.3.7>

ZJPAS (2023) , 35(3);78-85

1. INTRODUCTION:

All higher education institutions strive to improve the learning environment and academic level of all subjects taught in all the different colleges and departments. One way to achieve this goal is to collect student feedback on all the academic and educational activities in the university. The student feedback data collected can be analyzed in order to improve academic performance and provide better academic experience for students. This process will provide a great opportunity to improve the performance of teachers and also develop better curriculum for the subjects across all the programs (Lavanya, 2022).

The data collection system for the student feedback is an online system that is administered by the college staff and provides the students with log in IDs and password to provide security and privacy in answering the questions. Different kind of general and academic questionnaires could be developed and administered in different times of the academic year. The data is collected and stored in secure databases which could be accessed later on for analysis. The questions could be designed in a way to measure student responses using scales from 1 to 5 as a measurement for the level of teaching and academic level of the class. Some questions also could be designed to solicit

* Corresponding Author:

Haider Abdula Haddad
E-mail: Haider.haddad@su.edu.krd

Article History:

Received: 22/02/2023

Accepted: 02/03/2023

Published: 15/06/2023

textual answer from students to describe academic and educational levels or any problems that the students face in the university. All the responses could be stored in databases to be analyzed in a later time.

The feedback responses could be analyzed using machine learning algorithms and statistical methods to calculate different statistical features such as mean and standard deviation. Knowledge could be inferred from the data to be used to improve the academic performance. Knowledge could be represented using tables, graphs, and visualization methods. Deep learning methods could be used also on the data to carry out some data mining.

This system and the process is a very useful methodology that can be used by universities to improve their performance. The process could be automated by using online tools and in arrangement with the academic calendar to optimize the process and maximize the benefit to the academic institution. These programs could be implemented on a national level and to be unified across all academic institutions in order to come up with national policies that can benefit the whole country. Many countries have come up with such unified plans and programs and made it mandatory on the academic institutions to carry out annual review of all the programs they offer. The information gained from the feedback data has been used to develop a better higher education plans for the whole country not just one university (Kuo, et al.,2022).

2.Related Research:

Student feedback systems have been implemented in many universities and academic institutions and many research papers have been published on this subject. Each research paper explored the subject from different angle. In this section we present an overview of some research done in this area.

In paper (Lavanya, 2022), authors present their system of student feedback data collection using Google form. In their system, authors used questionnaires to collect data from students about topics such as curriculum, assessment, teaching learning process, placement, infrastructure and co-curricular activities. The data collected was analyzed using two algorithms. The algorithms were Naïve Bayes and Random Forest. The authors concluded that the model accuracy Naïve

Bayes gives best accuracy as 95% compared to Random Forest as 30%.

In paper (Kuo, et al.,2022), authors show how a student feedback system can be used for a specific purpose to aid students in their class in order to learn how to program in high end programming language. The system presented in the research divides the student into two groups; the experimental group and the control group. A system for both groups of students to upload and submit assignments and a code analysis and feedback improvement system were used. The system obtained feedback and the degree of improvement of homework programs; then, it distributed specially designed questionnaires to all students to directly obtain and quantify their feedback and perceived benefits of this system.

Papers (Bidve, et al., 2021) (Lwin, et al., 2020) (Kandhro, et al., 2020) (Asmita, et al., 2019) (Nasim, et al., 2017) represents research in student feedback systems that utilize systems to collect student feedback from students in universities and then analyze the data using some machine learning techniques. In paper (Bidve, et al., 2021) the emphasis of the data collection and analysis is textual feedback from students about the overall teaching quality. The system provided valuable ways to improve the teaching quality in the university. Some of the machine learning techniques used were (SVM), Random Forest, Naïve Bayes algorithm and lexical analysis out of which (SVM) has the best accuracy but takes more time in training for the large dataset.

The other research papers in this topic (Hatwar, et al., 2019) (Bhanukiran, et al., 2018) (Paulaa, et al., 2017) represent research in designing general student feedback systems and some analysis for the data collected. The research presented in these papers shows how academic and educational institutions can benefit from installing and using such system for improving quality of education in terms of teaching methodology, overall education experience, and curriculum.

3.Design of the system

The student feedback system developed in this research has some modules which they are put together to make the whole system. The system is designed to carry out two major functions: Data collection and Data Analysis. In the following section we explain the different modules of the

system. The system architecture is explained in

the following figure 1.

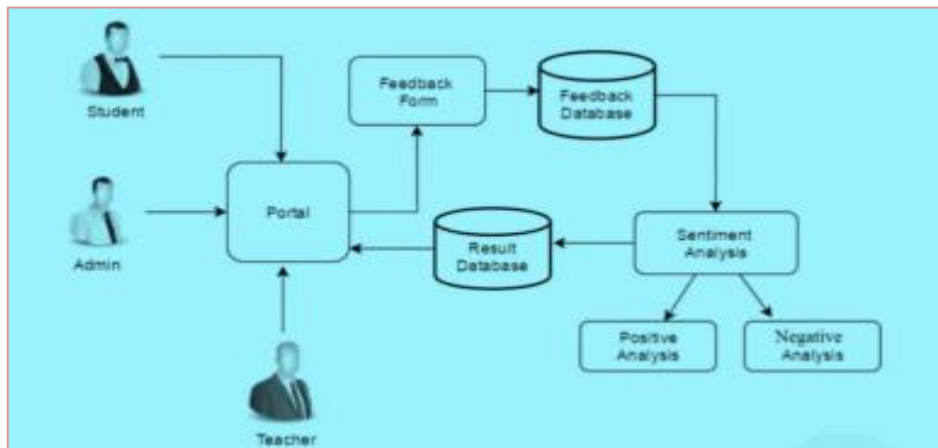


Figure 1: System Architecture

3.1 Data Collection Modules:

Module 1: Online system for student feedback.

Module 2: Log in module. This module allows system admins of colleges to set up log in IDs and Passwords for students.

Module 3: Questionnaire Form Builder. This module allows system admins to build and design different questionnaire forms. These forms can be used to collect responses from students.

Module 4: Data Collection Module. This module consists of the databases to store the answers from the questionnaires. The answers are obtained as a result of student feedback through this online system. The results can be also exported into spreadsheets.

3.2 Data Analysis Modules:

These modules are designed to use machine learning and statistical tools to analyze data and also use tools to visualize the data for presentation. This part of the system consists of the following modules:

Module 1: Data Understanding and Pre-Processing: In this module all the answers are prepared and subjected to an initial analysis. The collected student data sets are analyzed and evaluated using mean and standard deviation. Average and standard deviation for each question is calculated for evaluation purposes to determine how many students answered strongly satisfied, partially satisfied and dissatisfied. The purpose of the mean and standard deviation is to calculate student feedback responses to analyze and evaluate data sets to understand the data.

Module 2: Data Analysis Module: In this module, the machine learning techniques and statistical

analysis are employed to analyze the data and get useful information from the data. The information gathered in this step will help the university to improve the performance of the academic modules and the education process.

Module 3: Data Visualization Module: Data visualization is a powerful tool to represent information. Graphical representation of information provides a better understanding of the results of the data analysis. The graphical representation is used to illustrate the responses, statistical analysis, machine learning analysis, and conclusions.

4. Implementation of the system

The online student feedback system was implemented in the Salahaddin university which is the largest university in the Kurdistan region. The system is an online system for collecting student feedback data. The system consists of three main modules which they are the university module, the college module, and the student module. The users of the system include the super system administrator which has the overall control of the system and the college system administrator which has the control on the college level. These system administrators can add, delete, and authorize student and academics. If a student or a teacher tries to log in to the system they will use the ID and password issued to them by the system admin, then the system verify the users and their level of authority. The system then allows the users to fill out questionnaires and provide responses to specific questions. The responses will be collected and saved in the system database for later analysis. The flow of operations goes from designing forms to data collection, data

preparation, data analysis and then presenting results using different statistical and visualization

methods. This process is explained in figure 2 below.

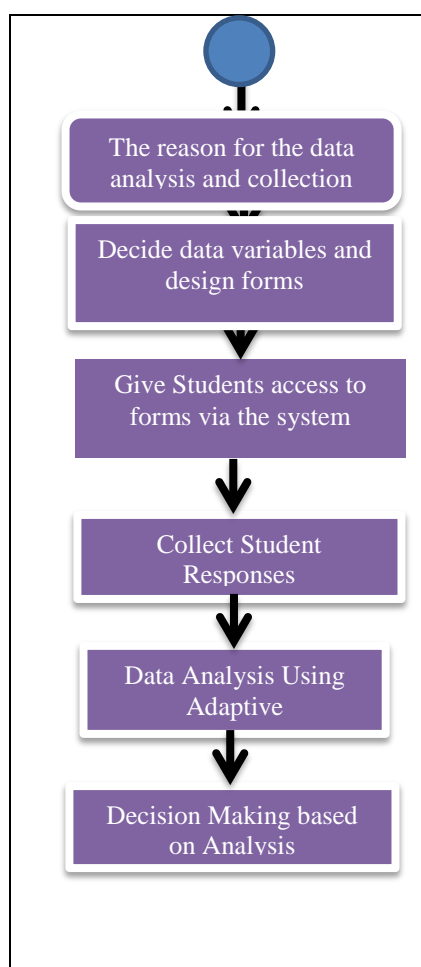


Figure 2: Algorithm of the data collection and analysis process

5.Data Collection

The data collected from the feedback system is a huge amount of data about the classes and the academic staff. The data includes answers from students about the aspects of each class. This huge amount of data can be stored in secure database and be analyzed later on. The data collected for this research is implemented on the case of Salahaddin University.

After the Super Admin enters the academic questions into the online system and sets the start and end times of the process, all registered students have official email addresses of the university, passwords are provided for all students who can easily answer the feedback questions at anytime and anywhere they have access to internet. After the students provide feedback, the data automatically appears in the dashboard which is seen by the super admin and then the system

can export the feedback results to separate spreadsheets in Microsoft Excel to analyze the data results obtained by the students.

The data collected is organized in a format where the students have the chance to score the performance of the teachers in three categories such as Poor, Average and Good performance.

6.Data Analysis

Data analysis in the system designed for this research can be performed using different methodology like statistical or machine learning methods. In order to reach higher accuracy in data analysis we chose to use a machine learning methodology.

Data classification problem has many applications in a wide variety of data mining applications especially when dealing with big data. These classification problems try to learn the relationship

between a set of feature or variables and a target variable of interest. Since many practical problems can be expressed as associations between feature and target variables, this provides a broad range of applicability of this model.

There are many common techniques in data classification. The most common methods used in data classification are decision trees, rule-based methods, probabilistic methods, (SVM) methods, instance-based methods, and neural networks. Among one of the best machine learning methodologies we selected Support Vector Machines to analyze the feedback data because it

is one of the best classification and fitting methods. (SVM) is a type of clustering algorithm that classifies data based on its characteristic. The (SVM) method is based on statistical learning theory. According to the limited data information, the (SVM) method tries to find the best compromise between complexity and the learning capacity. (SVM) can find the appropriate classification surface for the features. As you can see from figure 3, there is a separating line that can be considered the best separating line between data points. This separating line is called separating hyper plane.

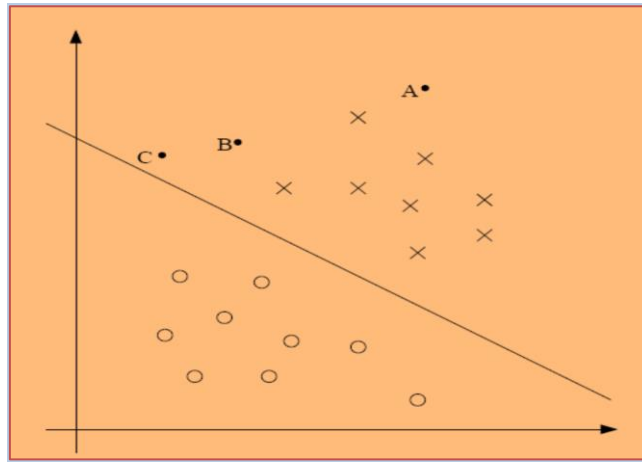


Figure 3: Support Vector Machine for classification

In order to implement the support vector machine to do the classification on the student feedback data the following steps are followed by (SVM):

- Import the dataset.
- Explore the data to figure out what they look like.
- Pre-process the data.
- Split the data into attributes and labels.
- Divide the data into training and testing sets.
- Train the (SVM) algorithm.
- Make some predictions.
- Evaluate the results of the algorithm

7- Classification and Results

The data classification method was implemented using (SVM) method on the student feedback data

using the steps explained in the previous section. The algorithm was implemented using MATLAB code. The results obtained from the classification of the data proved to be accurate and valid in terms of testing and training of the data set. Figure 4 and figure 5 show the obtained results in graphs. The student were surveyed about the performance of the teachers by asking them questions about the teaching, knowledge, using of technology, and assessment of the students. The classification from the data was to classify the teachers into three classes according to their academic performance from the responses to the questionnaires. The three classes from the analysis were Poor, Average, and Good.

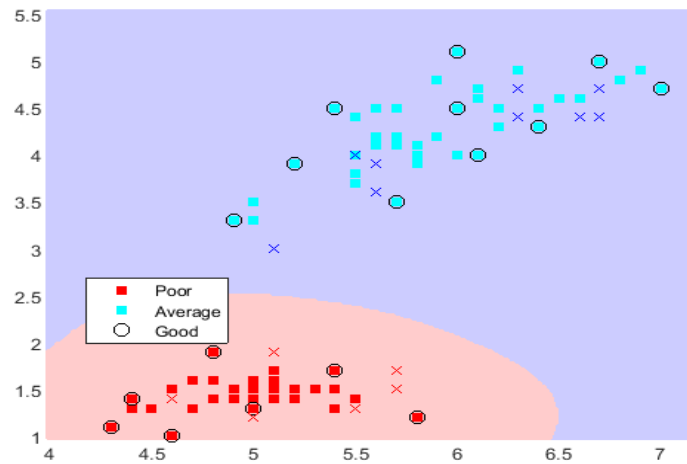


Figure 4: Classification results using (SVM) method

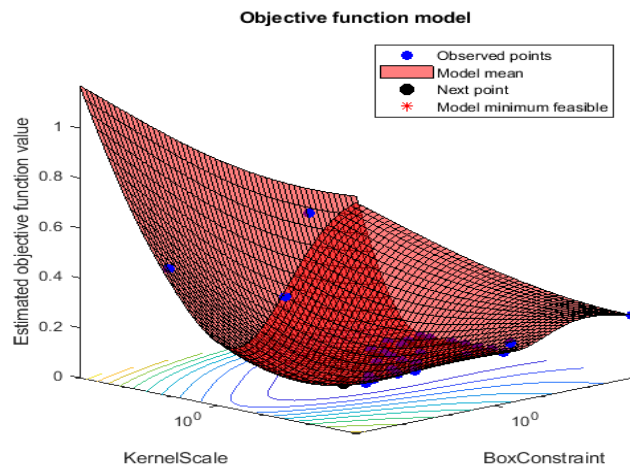


Figure 5: performance of the (SVM) method in fitting the objective function

A statistical data analysis was carried out on the student feedback data in major areas of the questions which are teaching quality, knowledge about subject, using technology, and fair assessment. The purpose of the statistical analysis was to see the performance of the analysis model

in dealing with the data. The results of the statistical analysis are shown in Figure 6 in the mentioned categories. From the statistical analysis we can observe the performance of the system according to different categories which is shown in figure 7.

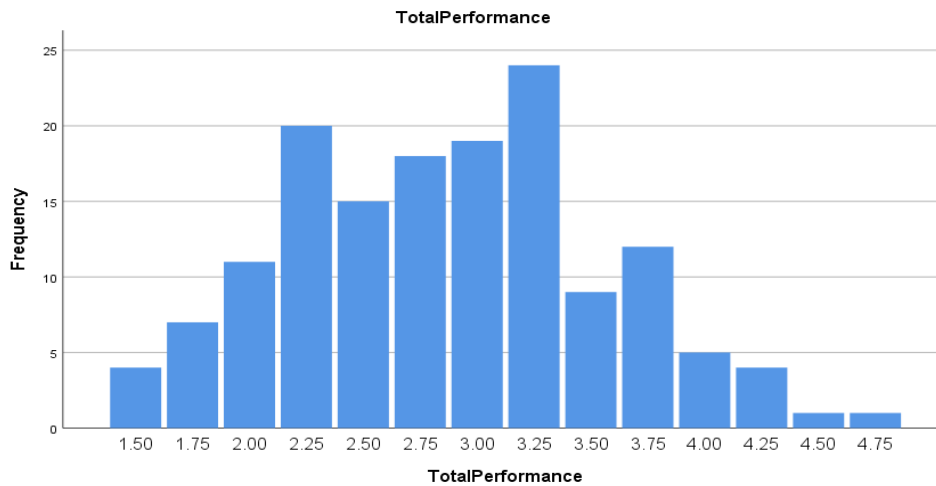


Figure 6: Analysis of the responses in major areas of teacher performance.

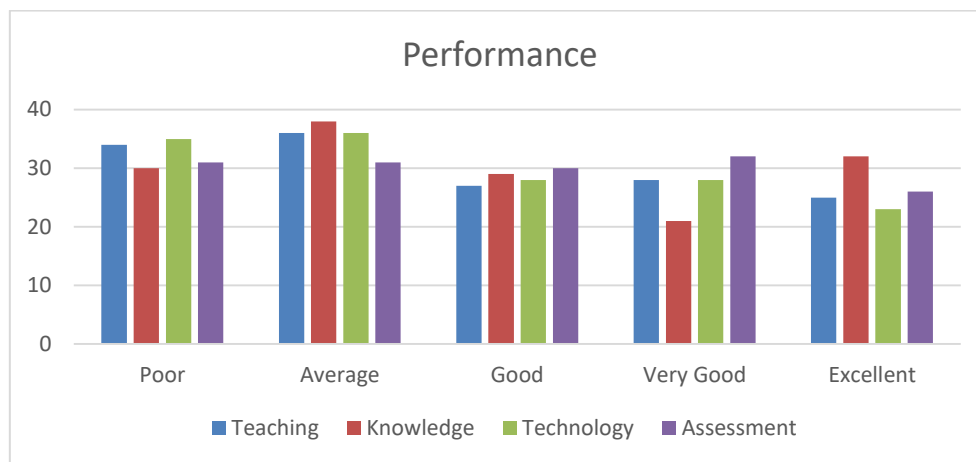


Figure 7: Performance according to different categories.

8. Discussion

From the obtained results of the classification that we observed we can clearly see that there are distinct classes of (Poor, Average, and Good) from the student feedback. These classes are the results of classifying the teaching staff data into three distinct classes. (SVM) method was successful in carrying out the classification. In this algorithm we divided the data into 80:20 set for training and testing set. From the figure 4 we can see the separate classes of (Poor, Average, and Good) shown in three different shapes and colors. This classification was followed with a prediction model for the final objective function as we see it on the figure 5. The figure shows the best fit objective function from the classes. When (SVM) method was implemented in this research it used the Radial Basis Function (RBF) kernel function for classification and it proved to be the best way

to provide clear separation in the hyperplane for the classes as seen in figure 4 and figure 5.

References

- Aggarwal, C. C., 2015. In: C. &. Hall/CRC, ed. *Data Classification Algorithms and Applications*. Yorktown Heights, New York, USA: CRC Press is an imprint of Taylor & Francis Group, an Informa business.
- A. N. & T. M., 2023. *CS229 Lecture Notes*. [Online] Available at: https://sgfin.github.io/files/notes/CS229_Lecture_Notes.pdf
- Asmita, S., Anuja, D. & D., A., 2019. Analysis of Student Feedback Using Deep Learning. *International Journal of Computer Applications Technology and Research*, 8(5), pp. 161-164.
- Bhanukiran, G. et al., 2018. Student Feedback System. *IRE Journals*, 1(10).

- Bidve, V. S., Patwari, P., Raut, R. & Kadam, A., 2021. Student Feedback System Using Machine Learning. *International Journal of Engineering Research in Computer Science and Engineering*, 8(8).
- Hatwar, N. et al., 2019. Online Student Feedback System. *Journal of Innovation in Engineering and Science*, 4(5).
- Kandhro, I. A. et al., 2020. Performance Analysis of Hyperparameters on a Sentiment Analysis Model. *Engineering, Technology & Applied Science Research*, 10(4), pp. 6016-6020.
- Kuo, J.-Y., Lin, H.-C., Wang, P.-F. & Nie, Z.-G., 2022. A Feedback System Supporting Students Approaching a High-Level Programming Course. *Applied Sciences*, 12(14).
- Lavanya, S., 2022. Application Of Student Feedback Using Machine Learning Model. *ijcrt*, 10(7).
- Lwin, H. et al., 2020. *Feedback Analysis in Outcome Base Education Using Machine Learning*. s.l., s.n.
- Nasim, Z., Rajput, Q. & Haider, S., 2017. Sentiment Analysis of Student Feedback Using Machine Learning and Lexicon Based Approaches. *IEEE*, Volume 4.
- Paulaa, F., Alexandera, B. & Georga, H., 2017. *Feedback Based Learning Through Online Feedback Systems in Higher Education*. s.l., Polytechnic University of Valencia Congress.