iConclusions https://iconclusions.com/ Conclusions in Medicine

Opinion mining of student regarding educational system using online platform

Muhammad Irfan,[1, a)] Anwar khan,[1] Muhammad Usman,[1] and Muhammad Waqas Haider[1]

Department of Computer Sciences, The University of Lahore, Sargodha, Pakistan

ABSTRACT: Covid-19 is a virus that is spread rapidly around the world. World Health Organization announced social distancing to control the spread of that virus. All institutions are closed in Pakistan. Education was also effecting with this shutdown. In the age of computing, social computing has emerged as a means of sharing knowledge, conveying ideas, and forming academic discussion groups, to name a few. Social websites or apps are also used for online study due to some critical situation as if nowadays we are facing many problems due to covid-19. Due to the covid-19 educational system is disturbed for that purpose we are introducing a different online platform for delivering knowledge and continue the educational system Many data mining techniques are applied to social network data for online analysis due to a large number of users and widespread use. This paper describes a method for extracting and analyzing master's student comments from the online survey that which platform is better for online study and also giving the opinion about most used platform. The proposed technique is implemented using different models or algorithms. By providing various preform as and analyzing various student opinions, the said system may assist the administration in improving the learning environment.

Received: 17 December 2024
Accepted: 23 March 2025
DOI: https://doi.org/10.71107/kx23gt95

I. INTRODUCTION

For academic intervention, data mining is a powerful method. Educational Machine Learning is the term for research in the context of education. Educational Data Mining is concerned with the creation of new methods for discovering information from educational databases and applying it to educational decision-making. It is important to obtain an education. No one should doubt its importance since it is a globally accepted phenomenon[1,2]. It has an effect on people's lives. A well- educated person can weather the challenges of life[3]. People's conduct is improved by education. COVID-19's global spread contributed in the cancellation of classes for over 850 million students around the world, undermining schools' initial teaching plans in these countries and regions[4,5].

The author in study[3], evaluated the online portion of a textured degree program for practitioners using a formal self-assessment and peer review using an instrument systematically designed according to Moore's principles of transaction's distance, and he discovered that tweaking a few course elements improved the layout, conversation, and autonomy of student learning. A virtual reality tour-guiding network was created, and some students from a Pakistani technical university took part in the research[6–10]. The results revealed their learning efficacy and acceptance of technology in the educational system.

This graph (Fig. 1) represents that how much the occurrence of online study in the next years with different platforms, which perform better. This classification can be used by developers to choose online platforms based on the setting and context of use. Lastly, research gaps have been identified for future researches.

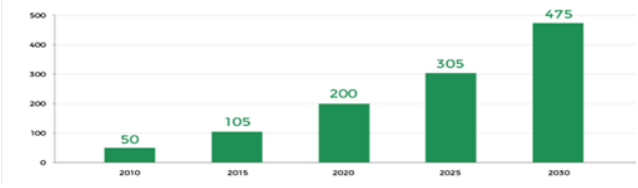

FIG. 1: Yearly Chart.

[a)] Electronic mail: Muhammad.irfan1@cs.uol.edu.pk

II. THEORETICAL FRAMEWORK

Constructivist theory and teaching paradigms are congruent with the literature analysed, as well as the issues and potential solutions that address student impressions of on-line courses[11]. Identified 16 constructivism features, three of which are highlighted in this research study:

1. Guides, monitors, mentors, tutors, and facilitators are all roles that teachers play.

2. Student major issue, relatively high reasoning abilities, and deep comprehension are emphasized.

3. Individual contexts, social negotiation, teamwork, and practise all play a role in learning.

In e-learning environments[12], adapted[11] characteristics into an advanced model of constructivism. The three components that have been determined are: Collaboration, teamwork, various interpretations of concepts, and social negotiations are all part of the design of learning activities.

1. Mentoring, acknowledging, giving input, and evaluating student learning are all roles that instructors play.

2. Learning assessment may be done by the mentor, in conjunction with the student alternatively, the student himself.

Educational technologies are frequently utilised to simplify the learning process by creating a conducive learning environment and assessing student progress[13]. As a result, A constructivist approach to online education will promote more creativity in the classroom while avoiding the use of technology to measure student success. (for example, presenting knowledge and providing drills and practise). Technology is more useful in theory when students use it to identify a problem and then use the appropriate tools to help them understand it[14].

Teamwork (amongst teachers and students, as well as between students and other students), online content methods, and educational methodology frameworks are three areas that need particular attention and are discussed in this report are based on these distinctive nature of online courses and e-learning environments, as well as constructivist theory components.

III. METHODOLOGY

A. PROPOSED METHODOLOGY

In recent studies, machine learning has proven to be extremely useful in offering better solutions and data prediction in less time. We can make predictions based on historical data thanks to machine learning[15,16]. Machine learning is another term for it (ML) enhancing our ability to use forecasts. We will describe the proposed architecture in this section. We defined our work's literature in the previous section, then created a comparison table of previous work and developed an existing architecture. The following is how we split the proposed architecture:

- Data Definition
- New Variable
- Proposed Techniques High Level Architecture
- Proposed Methodology Detailed Description

B. DATA COLLECTION

The data for this thesis was gathered from Pakistani students using an online Google form. The purpose of the study was explained in the email to the student recipients, which also included an informed consent statement (Fig. 2). The authorization form also contained the the name of the researcher and contact details for the researcher's counsellor. Students may use this information to contact either the researcher or the researcher's academic advisor if they had any questions or concerns regarding the survey. The declaration also emphasizes that participation in the study is voluntary. When survey participants started a survey, they were told that they could stop at any time and that there would be no consequences if they did. The researcher will not be gathering or receiving any identifying in-formation, according to the email address. The university's Statistical Consulting Center was in charge of this account.

IV. RESULTS

A. DATA ANALYSIS

The research question did not differentiate between the four classes because it was designed to measure student barriers to virtual classrooms irrespective of whether they were able or planning to take an online study in the future. The aim of this question will be to identify any potential limitations that almost all students could face when deciding whether to take an online course.

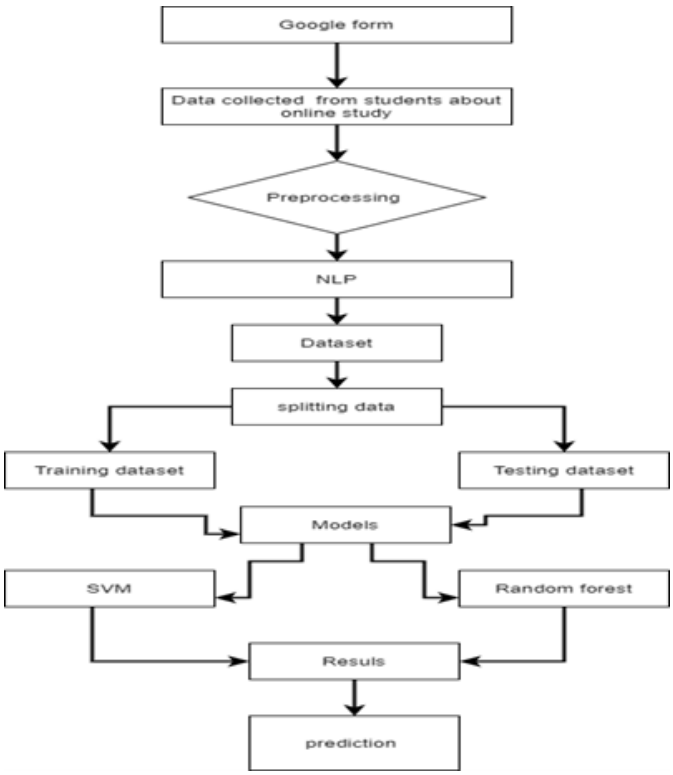

FIG. 2: Proposed methodology.

	Timestamp	Gender	First Education Language?	Are you studying in collage or University ?	University or collage Name?	Departmen: ?	What device do you use for online classes?	Which platform used for online classes?	During class do you face Electricity Problem?	During class do you face internet Problem?	During online class do you face noise disturbance at home?	Are you satisfied with your online tutor?	Do you have sufficient computer skills for doing online class?	Do yo comfortat communic in online w.r.t phy cl
0	3/28/2021 2:59	Male	English	University	IUB	Computer science	Laptop	University LMS	4.0	2.0	3.0	1.0	2.0	
1	3/28/2021 3:01	Male	English	Collage	Muslim Collage	Computer science	Smart phone	Zoom	4.0	2.0	3.0	2.0	3.0	
2	3/28/2021 5:59	Male	Urdu	University	Air university	Computer science	Smart phone	Zoom	1.0	1.0	1.0	1.0	1.0	
3	3/28/2021 6:03	Male	English	University	IUB	Computer science	Laptop	University LMS	5.0	4.0	5.0	5.0	5.0	
4	3/28/2021 6:05	Male	English	University	IUB	Computer science	Smart phone	Zoom	NaN	NaN	2.0	3.0	4.0	

FIG. 3: Original Data set.

A research Question All of the respondents who would not have been able to choose whether to take an online course must be considered. In improving the 34 validity of the data for the first research question, response frequencies were calculated and evaluated.

The research question seeks to determine whether there is a substantial gap between students who choose or plan to take an online course in the future and those who do not. A statistical comparison was made using several Mann-Whitney U Tests. This approach was cho-

	Are you studying in collage or University ?	University or collage Name?	Department ?	What device do you use for online classes?	Which platform used for online classes?	During class do you face Electricity Problem?	During class do you face internet Problem?	During online class do you face noise disturbance at home?	Are you satisfied with your online tutor?	Do you have sufficient computer skills for doing online class?	Do you feel comfortable to communicating in online class w.r.t physical class?	Are you able to manage study time effectively and can easily complete assignments on time?	Do you think online classes are saving your time?	in
0	1	35	38	1	6	4.0	2.0	3.0	1.0	2.0	1.0	4.0	2.0	
1	0	46	38	2	9	4.0	2.0	3.0	2.0	3.0	2.0	3.0	4.0	
2	1	5	38	2	9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	
3	1	35	38	1	6	5.0	4.0	5.0	5.0	5.0	5.0	5.0	5.0	
4	1	35	38	2	9	3.0	1.0	2.0	3.0	4.0	4.0	5.0	4.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
228	1	44	3	2	6	4.0	2.0	2.0	1.0	4.0	4.0	2.0	4.0	
229	1	54	66	2	1	4.0	5.0	5.0	1.0	3.0	2.0	3.0	2.0	
230	1	10	1	2	6	3.0	5.0	3.0	3.0	3.0	4.0	3.0	2.0	
231	1	54	51	2	9	5.0	1.0	1.0	1.0	2.0	1.0	4.0	1.0	
232	1	51	51	2	2	2.0	1.0	1.0	1.0	2.0	1.0	1.0	2.0	

233 rows × 24 columns

FIG. 4: preprocessing the data set.

```
df.describe()
```

	During class do you face Electricity Problem?	During class do you face internet Problem?	During online class do you face noise disturbance at home?	Are you satisfied with your online tutor?	Do you have sufficient computer skills for doing online class?	Do you feel comfortable to communicating in online class w.r.t physical class?	Are you able to manage study time effectively and can easily complete assignments on time?	Do you think online classes are saving your time?	As a student, do you enjoy working independently?	During class do you face voice problem?	During class do you face streaming problem?	Dur do comm t
count	228.000000	229.000000	229.000000	229.000000	229.000000	228.000000	230.000000	230.000000	231.000000	227.000000	226.000000	22
mean	2.758772	2.550218	2.633188	2.851528	3.331878	2.771930	3.182609	3.000000	3.380952	2.577093	2.544248	
std	1.346774	1.338814	1.464773	1.604478	1.499674	1.601404	1.507377	1.640109	1.510077	1.362261	1.278988	
min	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	
25%	1.000000	1.000000	1.000000	1.000000	2.000000	1.000000	2.000000	1.000000	2.000000	1.000000	1.000000	
50%	3.000000	2.000000	2.000000	3.000000	4.000000	3.000000	3.000000	3.000000	4.000000	3.000000	2.000000	
75%	4.000000	3.000000	4.000000	5.000000	5.000000	4.000000	5.000000	5.000000	5.000000	4.000000	3.000000	
max	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	

FIG. 5: Description of data set.

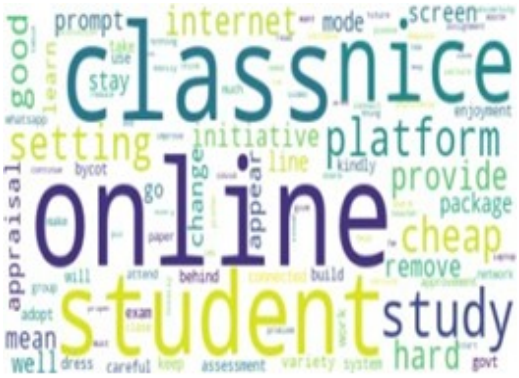

FIG. 6: Opinion analysis of students for conducting online classes.

sen because it allows for comparisons between groups of various sizes. This approach was used to examine the variations between students who preferred traditional environments and students who chose to take an online class. Data analysis have some steps:

- Collect data (Google form)
- Preprocessing the data
- Model development (NLP)
- Then splitting the data (testing and training)
- Applying Models (SVM, RF)
- Compare and evaluation of Outcomes

B. ORIGINAL DATA

First we visualize (Fig. 3) the data which we collected from the source of Google online form. In which students give their opinions about the online study from different platform.

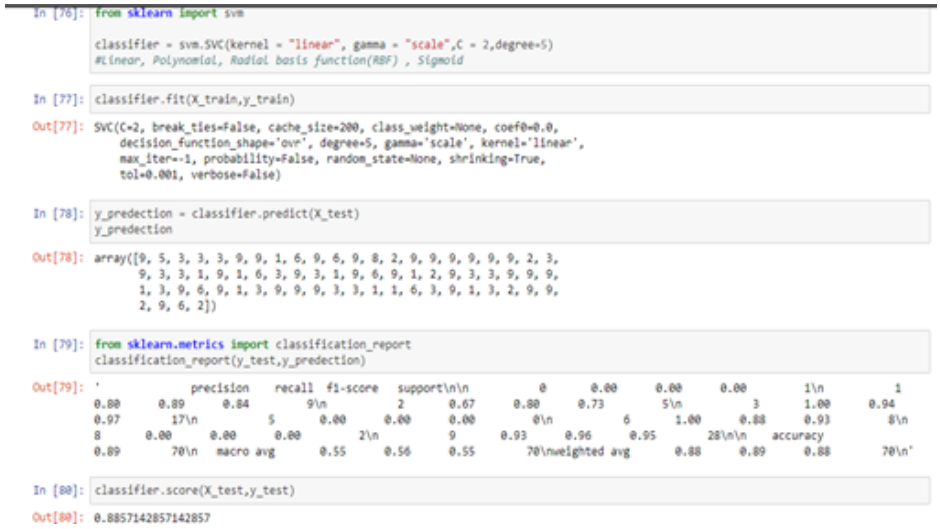

FIG. 7: Preferring Online classes.

FIG. 8: Preferring Online classes.

C. PREPOSSESSING THE DATA

We display (Fig. 4) all of those things which we doing on my project. First step is conducting the data from students. Then we have to preprocess our data because some students missing the questions and some are not correctly giving authentic answer of the question. That is why we preprocess the data into integer values from text and overcome the missing values. Description of data set is shown in Fig. 5.

D. OPINION ANALYSIS

Now we are going to show (Fig. 6) the graphical representation of the Students opinions in Table I.

TABLE I: Opinion Analysis.

Opinion	Comments
Positive	Good, Appraisal, Initiative, well, Appear,
Negative	Nice, Enjoyment, Learn
Heutral	Hard, Remove, Go, Cheap, Bycot,

E. SVM

Multiple classification problems, a support vector machine (SVM) is a supervised machine learning model that uses classification algorithms. SVM models will categories new text after being given sets of labelled training data for each group. Therefore, you are attempting to solve a text classification problem (Fig. 7).

F. RANDOM FOREST

Random forests, also known (Fig. 8) as random decision forests, are an ensemble learning method for classification, regression, and other tasks that works by training a large number of decision trees and then outputting the class that is the mode of the classes (classification) or the mean/average prediction (regression) of the individual trees. We are showing the comparison of RF and SVM by Table II.

TABLE II: Our Model Prediction.

Sr. No	Models	Accuracy	Precision	Recall	F1-Score
1	SVM	0.8	0.80	0.89	0.84
2	RF	0.6	0.72	0.75	0.78

V. CONCLUSION

The conclusion of this research is to look at the challenges that students face while taking online classes. We are summarizing a big problem and then posing some significant questions. Taking a survey from Pakistani students using a Google form Collaboration and engagement are significant obstacles to taking an online class because of this. Students tend to communicate with an instructor through asynchronous communication tools for course topics involving collaboration among students in a small group to complete an assignment, discussion of course content, and communication with an instructor (e.g., email, text messages, and discussion board). A new variable was developed using an existing variable to get a more accurate predicted value. We use Natural Language Processing (NLP) to predict various points of view and help us refine the framework. We used various models (SVM AND RF) to predict and determine which model is most efficient in analyzing the data sets. Then we also conclude that SVM give the best accuracy as compared to random forest. The authenticity of online study was determined through models using various techniques. For future jobs, social media data and all schools and colleges' data or survey can be combined with this data to predict current time. Financial news may also be taken into account for additional criteria in order to predict better outcomes. This study has some limitations, including the relatively small sample size and the focus on a single country (Pakistan). Future studies could expand the sample size and include students from different countries to improve the generalizability of the results. Future research could explore the integration of more advanced NLP techniques, such as deep learning models, to further improve sentiment analysis accuracy.

DECLARATION OF COMPETING INTEREST

The authors have no conflicts to disclose.

REFERENCES

[1] K. Mouthami, K. N. Devi, and V. M. Bhaskaran, "Sentiment analysis and classification based on textual reviews," in *2013 International Conference on Information Communication and Embedded Systems (ICICES)* (IEEE, 2013) pp. 271–276.

[2] G. Vinodhini and R. Chandrasekaran, "Sentiment analysis and opinion mining: a survey," International Journal **2**, 282–292 (2012).

[3] F. Kainat, S. Ilyas, N. Jabeen, and A. Hussain, "Advancements in layered double hydroxide (ldh) nanosheets: Preparation techniques and optoelectronic applications," in *Layered Nanomaterials for Solution-Processed Optoelectronics* (CRC Press, 2025) pp. 295–317.

[4] N. Jabeen, A. Hussain, I. Ahmad, and J. Ali, "Mxenes-based hybrid for electrochemical sensing application," in *MXenes as Emerging Modalities for Environmental and Sensing Applications* (Elsevier, 2025) pp. 203–215.

[5] N. Tanwani, S. Kumar, A. H. Jalbani, S. Soomro, M. I. Channa, and Z. Nizamani, "Student opinion mining regarding educational system using facebook group," in *2017 First International Conference on Latest Trends in Electrical Engineering and Computing Technologies (INTELLECT)* (IEEE, 2017) pp. 1–5.

[6] V. Kagklis, A. Karatrantou, M. Tantoula, C. T. Panagiotakopoulos, and V. S. Verykios, "A learning analytics methodology for detecting sentiment in student fora: A case study in distance education," European Journal of Open, Distance and E-learning **18**, 74–94 (2015).

[7] A. Abdelrazeq, D. Janßen, C. Tummel, S. Jeschke, and A. Richert, "Sentiment analysis of social media for evaluating universities," in *Automation, Communication and Cybernetics in Science and Engineering 2015/2016* (Springer, 2016) pp. 233–251.

[8] W. Jie, L. Hai-yan, C. Biao, and Z. Yuan, "Application of educational data mining on analysis of students' online learning behavior," in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)* (IEEE, 2017) pp. 1011–1015.

[9] E. P. Yildiz, M. Tezer, and H. Uzunboylu, "Student opinion scale related to moodle lms in an online learning environment: Validity and reliability study," International Journal of Interactive Mobile Technologies (iJIM) **12**, 97–108 (2018).

[10] R. O. Bustillos, R. Z. Cabada, M. L. B. Estrada, and Y. H. Pérez, "Opinion mining and emotion recognition in an intel-

ligent learning environment," Computer Applications in Engineering Education **27**, 90–101 (2019).

[11]V. Kagklis, A. Karatrantou, M. Tantoula, C. T. Panagiotakopoulos, and V. S. Verykios, "A learning analytics methodology for detecting sentiment in student fora: A case study in distance education," European Journal of Open, Distance and E-Learning **18**, 74–94 (2015).

[12]R. O. Bustillos, R. Z. Cabada, M. L. B. Estrada, and Y. H. Pérez, "Opinion mining and emotion recognition in an intelligent learning environment," Computer Applications in Engineering Education **27**, 90–101 (2019).

[13]V. Dhanalakshmi, D. Bino, and A. M. Saravanan, "Opinion mining from student feedback data using supervised learning algorithms," in *2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC)* (IEEE, 2016) pp. 1–5.

[14]E. P. Yildiz, M. Tezer, and H. Uzunboylu, "Student opinion scale related to moodle lms in an online learning environment: Validity and reliability study," International Journal of Interactive Mobile Technologies (iJIM) **12**, 97–108 (2018).

[15]T. Chen, L. Peng, B. Jing, C. Wu, J. Yang, and G. Cong, "The impact of the covid-19 pandemic on user experience with online education platforms in china," Sustainability **12**, 7329 (2020).

[16]R. Yang, "Machine learning and deep learning for sentiment analysis over students' reviews: An overview study," (2021), unpublished manuscript.