

# An Efficient Data Mining Technique for Assessing Satisfaction Level with Online Learning for Higher Education Students

Mrs. M. Pragathi<sup>1</sup>, Sharanya Kaleri<sup>2</sup>, M.Sony<sup>3</sup>, Saleha Sulthana<sup>4</sup>

<sup>1</sup>Professor, Computer Science and Engineering, Sridevi Women's Engineering College Hyderabad, India

Email: <sup>1</sup>yeddula.pragathi@gmail.com

<sup>2</sup>Computer Science and Engineering, Sridevi Women's Engineering College, B.Tech IV Year Hyderabad, India

Email: <sup>2</sup>sharanya.7701@gmail.com

<sup>3</sup>Computer Science and Engineering, Sridevi Women's Engineering College, B.Tech IV Year Hyderabad, India

Email: <sup>3</sup>mukarasony@gmail.com

<sup>4</sup>Computer Science and Engineering, Sridevi Women's Engineering College, B.Tech IV Year Hyderabad, India

Email: <sup>4</sup>salehasulthana980@gmail.com

## Abstract

By working on understudies' scholarly execution, each instructive association endeavors to raise the general nature of schooling. For this reason Educational Data Mining (EDM) is a quickly developing field of exploration that utilizes the standards of Data Mining (DM) to assist instructive foundations with finding significant data about Student Satisfaction Level (SSL) with regards to Online Learning (Old). To get the ideal informational settings, a couple of practices have been taken a stab at using EDM to calculate students' approaches to acting. Thus, Feature Selection (FS) is regularly used to distinguish the least cardinality yet most pertinent subset of attributes. This study examines the similarity of the SSL model with FS approaches inside and out in light of the fact that the FS cycle fundamentally affects a fulfillment model's anticipated precision. In such manner, a web-based poll was utilized to at first gather a dataset of understudy assessments of Old courses. The display of covering FS approaches in DM and portrayal computations was surveyed to the extent that health values using this dataset. At long last, the nature of 11 covering based FS calculations and the k-NN and SVM as gauge classifiers is utilized to assess the decency of subsets with different cardinalities concerning expectation precision and the quantity of chosen highlights. Both the best strategy and the element subset with the best dimensionality were tracked down in the examinations. The ongoing review's discoveries plainly support the deeply grounded association between an expansion in expectation exactness and the presence of few qualities. The fundamental assortment of attributes may really help with the improvement of productive

educational drives, making FS very important for high-accuracy SSL expectation. Our work yields a component size decrease of up to 80% and order accuracy of up to 100% on the used continuous dataset.

**Keywords:** classification, feature selection, machine learning, and educational data mining.

---

## 1. Introduction

To guarantee that the instructive interaction proceeds, various whimsical instructive systems have been introduced. Online learning (OL), also known as learning in a coordinated or nonconcurrent environment using various Web-connected devices like mobile phones, PCs, etc., was one of the options. Understudies can review and speak with teachers and colleagues anyplace through these stages. Old has filled in popularity fairly as of late considering the way that it thinks about more noteworthy flexibility in general setting, faster review, more imperative receptiveness, more powerful permission to a greater collection and more vital proportion of information, and lower monetary utilizations. The most noticeable piece of the advancement was the usage of online stages and courses. Be that as it may, we keep on experiencing various deterrents and detours. Unfortunate Web association thwarts understudy learning around the world, regardless of the way that strong computerized foundation and stages are essential for the conveyance of online courses and commitment to information assortment. New development is normal for the two students and instructors to associate immaculately with dynamic and free learning. A solid assessment framework was expected to ensure the general nature of online schooling with regards to learning results. These days, quality is estimated by whether learning goals are met and close to home and social aspects are created. In this way, a gadget to separate the developing involvement with general, as well as the positions and associations of teachers, understudies, and showing materials in post-mechanized learning settings, is required. Universities and schools are severely constrained in their ability to meet learning objectives and improve Student Satisfaction Level (SSL), since these estimates inadvertently demonstrate the legitimacy of such foundations' previous organisational frameworks. Over the span of the review, SSL was a part of the general degree of encounters and seen execution related with instructive administrations. SSL is finished up somewhat by how students overview their educational experiences, organizations, and workplaces. As indicated by, SSL must be carried out when there is no inconsistency between what is expected and what is executed by the specialist co-op.

It is vital for note that Educational Data Mining (EDM) can possibly extraordinarily help the instructive examination process here. Accordingly, to create results that can be deciphered, the gathered information ought to be arranged and

appropriately analyzed. EDM techniques won't function also in the event that you don't pick the right methodology for the examination. Truly outstanding and huge data examination progressions is feature selection (FS). High layered data could have negative repercussions in applied models. With improved components and model management, two models are increasing the amount of time spent preparing. Particularly with highdimensional datasets including noisy, abundant, and inessential elements, FS is crucial in machine learning (ML) and data mining (DM). To limit the impact of noise and insignificant features while still producing amazing intended outcomes, FS tries to choose a selection of variables from the data sources that can more clearly portray the data. A significant problem in data disclosure, the acceleration of DM methods, and execution smoothing out is the choice of component subsets. FS has additionally demonstrated to be a productive and compelling strategy for planning high-layered information in various DM and ML circumstances. FS objectives could incorporate making models that are more precise, accelerating information mining, and making information that is straightforward.

## 2. Literature Review

### **Consequences of people mobility limitations under COVID-19 in terms of the economy and society:**

This article examines the financial cost of safety measures implemented to halt the emerging coronavirus illness 2019 (COVID-19) among Italians. We model convenience limitations as a regular strange interruption to the economy by utilizing an extensive variety of human development. Our discoveries demonstrate that lockdown systems make two impacts: As a matter of some importance, they greater affect versatility in monetarily prosperous locales; Second, they partition the populace in two ways: Portability diminished more pronouncedly in regions with more prominent variations and lower per capita compensation. We stress the requirement for planning systems that emphasis on decreasing destitution and imbalance while additionally considering these results.

Various public legislatures have executed safety efforts to lessen the pace of transmission because of the coronavirus disease 2019 (COVID-19) episode. Utilizing broad investigation of Facebook-gave close steady Italian improvement information, we inspect the monetary impacts of prison systems on people and neighborhood assemblies. The improvement shift is shown as an external interference, equivalent to a trademark calamity. Italians are affected by advancement limits in two ways. To start, we find that the determination greater affects districts with bigger monetary plans. Second, since individuals with restricted portability and lower per capita pay live in regions with more noteworthy differences, we track down proof of isolation. Our discoveries exhibit the unprecedented hardships of detainment as well as the cultural expenses: From

one perspective, the crisis lessens the money related pay of public and neighborhood state run organizations. To shield society's most powerless people and cutoff the conclusion provoked improvement in destitution and awkwardness, a broad money related liability is fundamental.

### **Restrictions on human mobility and the emergence of the new coronavirus (2019ncov) in china:**

It has been assessed that the Wuhan closure, which restricted human portability, added to the spread of Coronavirus. The conclusion impact is recognized from the misgiving, viral, and Spring Festivity impacts utilizing DID estimations. The finish of Wuhan 76.98%ly impacted imports, a 56.31 percent effect on conveys, and a 55.91 percent effect on intra-city improvement. The quantity of Covid cases in 347 Chinese people group beyond the Hubei area would have been 105.27% higher in the event that Wuhan had not been closed down.

Because of limitations on human portability, remembering Wuhan's conclusion for January 23, 2020, we evaluate the control and postponement of the Clever Covid. 2019-nCoV). To detach the conclusion influence from other intruding effects, for instance, the fear influence, contamination impact, and Spring Festivity influence, we use difference-in-differences (DID) approximations. Since to the conclusion, brings into Wuhan fell by 76.98%, withdrawals from Wuhan fell by 56.31 percent, and moves inside Wuhan fell by 55.91 percent. Up to 22 deferred populace moves from Wuhan and different regions in Hubei, the focal point of the 2019-nCoV pandemic, are the subject of our examination, just like the impacts on new contaminations in the urban communities we are zeroing in on. We moreover show that general population moves from point of convergence towns in Hubei region upheld the spread of 2019-nCoV in target metropolitan networks in 98 Chinese metropolitan regions with further developed social division rules. The quantity of Coronavirus cases in the 347 Chinese towns beyond the Hubei region would have expanded by 105.27 percent in the event that Wuhan hadn't been closed down on January 23, 2020. Our discoveries may help out to worldwide endeavors to diminish scourges.

### **How many ways are there to characterise online education? An analysis of the definitions of online learning from 1988 to the present:**

The idea and adage "online advancing" have been the subject of instructive examination for over twenty years. Since the idea of web based learning has various implications, in spite of being habitually examined, we present the consequences of a thorough writing survey of definitions in this article. The term is utilized by scholastics and authors to portray thoughts that are entirely against, while possibly not even disconnected. To explore the by and large and explicit impacts of internet learning, we took a gander at the whole organization from

1988 to 2018. Following the assortment of 46 definitions from 37 sources, we led a substance investigation of these definition sets. The accumulated implications' substance examination uncovered the distinctive qualities of web based advancing as well as the word and elective vulnerability. The advancement of the significance of internet learning throughout the course of recent many years has been connected to the historical backdrop of innovation.

### **Investigating how multimedia might improve social presence in an asynchronous online course:**

Online education is filling in prominence, however worries about its quality are developing. One of the principal burdens of online guidance is social separation. A new report recommended focusing on well disposed presence methodologies in an electronic report corridor to lighten this feeling of gloom. Nonetheless, there are various snags to defeat while establishing an internet learning climate where understudies feel major areas of strength for an of social presence. This could be because of the way that most past exploration utilized self-report polls to quantify social presence, and the idea of social presence is vague. This study tried to examine how sight and sound upgrade social presence in a surprising electronic course and how friendly presence influences a multimodal conversation stage. Moreover, the goal was to learn, for different reasons, how teachers and understudies had an outlook on the utilization of video in a web-based course. This study assessed how sight and sound improved social presence in a web-based course utilizing a blended strategies, contextual investigation exploratory plan. The review used three obvious plans to comprehend the use of mixed media and the case of extending social presence: neighborhood demand, social constructivism, and casual association assessment. According to the investigation, video could help social presence in an online learning bundle in different ways. The discoveries demonstrate that specific sight and sound arrangements, as VoiceThread, expanded social presence however not contacts. Besides, the examination found that Rourke et al. 's (2001) social presence scoring technique was unequipped for perceiving explicit social presence markers during a sound and sight discussion stage. To get familiar with an understudy's popularity in a learning organization, this study checked out at various assumptions.

### **The potential immunological mechanisms behind the diverse immunopathogenesis of COVID-19 infections in young, old, and healthy adults:**

An overall scourge was started by Coronavirus, a fresh out of the box new Covid that was found in Wuhan, China, in December 2019. It is as of now obscure what causes this viral contamination in people or how the resistant arrangement of the host answers this clever danger. Covid sicknesses, of course, have a broad assortment of recurrence and passing rates, going from vague and delicate to deadly and outrageous. Babies are insusceptible to serious or lethal

contaminations, yet the older and delicate are generally defenseless to this infection. It is fundamental to decide the plausible contamination and host interchanges that make Covid patients experience such fluctuating unfavorable results.

### 3. Methodology

Colleges and other instructive foundations should have the option to meet their learning targets while at the same time bringing understudy fulfillment steps up in request to show the adequacy of their Good old frameworks SSL. How much understudies are satisfied with the instructive choices, offices, and administrations given by SSL is the way the school is assessed. SSL must be accomplished, as indicated by, on the off chance that the specialist co-op gives definitively what is expected of it. Low Web availability upsets understudy advancing all over the place, in spite of the requirement for strong computerized foundation and stages for online course conveyance and information assortment. New progression is essential for students and educators to participate in strong and autonomous progressing really. A reliable assessment part meant quite a bit to ensure the overall level of internet tutoring concerning learning results. Achievement of educational goals, as well as amicable and mental new development, are quality markers in the pandemic period. Hence, an instrument is significant in present high level learning settings on assess the complete developing experience as well as the commitments and correspondences of teachers, understudies, and instructive advances. The Student Satisfaction Level (SSL) and learning objectives that colleges and schools can accomplish are basic signs of the adequacy of their internet learning frameworks. All through the assessment period, SSL is a part of similar degree of occasions and recognizable triumphs related with instructive organizations. SSL is impacted by how understudies see their instructive encounters, offices, and administrations. As per this, SSL execution is just conceivable in the event that there is no inconsistency between what was expected and what the specialist organization gave.

#### Disadvantages

1. Since these pointers by suggestion show the capability of those establishments' Old fashioned systems.
2. Universities and institutions must achieve their educational goals in order to raise their Student Satisfaction Level (SSL).

#### Proposed System

An machine learning (ML) model was made in this article to ensure the most elevated level of precision during testing. The proposed model was developed with the assistance of the k-NN and SVM ML Classifiers. Before, the most broadly utilized order methods were k-NN and SVM. Furthermore, ABC, Whale

Streamlining, and Sailfish Improvement, as well as irregular timberland, choice tree, citizen classifier, and move learning with CNN and LSTM layers, were used in this review. K-NN and SVM have commonly been the most broadly perceived request computations. Likewise, 11 meta-heuristic strategies were used in this review to find important dataset attributes. An uncommon model that expects SSL with additional created accuracy and steadfast quality was built using 11 meta-heuristic philosophies.

### Advantages

1. The disclosures of the energy research clearly support the hypothesis that action accuracy is on the other hand comparative with the amount of characteristics present.
2. The utilization of FS for high-precision SSL projection is remarkable because of the way that the right blend of elements may essentially help with the production of effective educational exercises.
3. Our survey uncovers a component size lessening of up to 80% and portrayal accuracy of up to 100% on the picked consistent dataset.

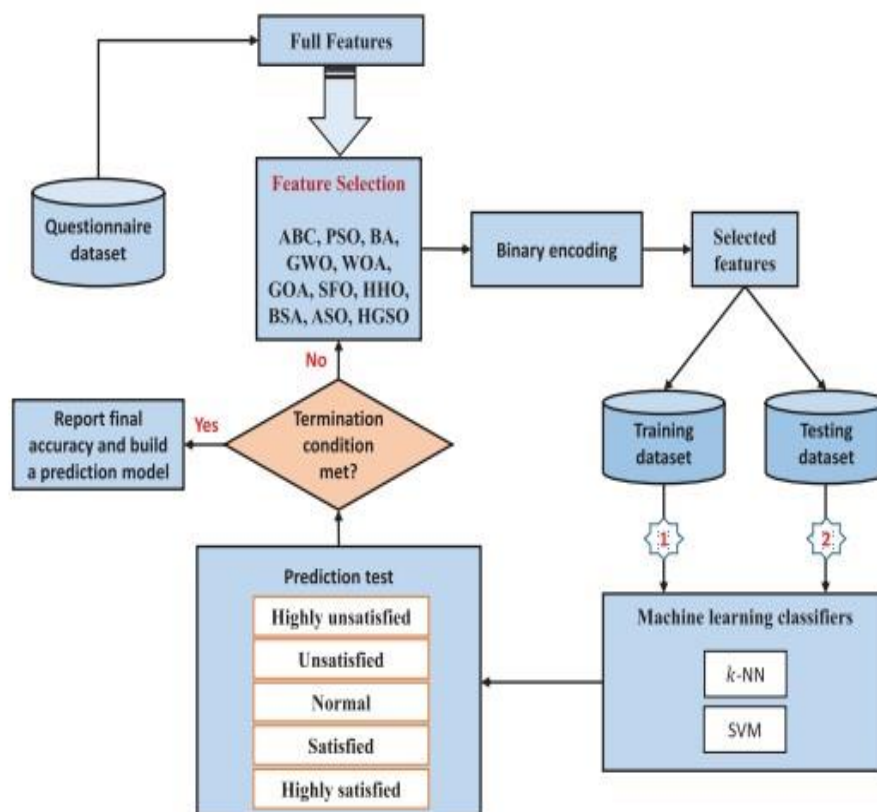


Fig.2: System architecture

**Modules**

We grown the modules filed beneath in consideration of carry out the earlier project.

- Exploration of dossier: we will place facts into the foundation handling this piece.
- Processing: Using this piece, we will state dossier for deal with.
- Utilizing this piece, facts will be private into train and test.
- Creation of models: Using CNN joined accompanying LSTM tiers ABC, porpoise growth, and sailfish addition, you can devise k-NN, SVM, haphazard thicket, conclusion sapling, and vote classifier. Algorithms' veracity is planned.
- Signing up and record in: Utilizing this piece will cause success admission and login.
- Client input: Predicted input will accompany utilizing this piece.
- Prediction: The presented anticipated ending profit

**Implementation**

**Algorithms:** K-NN: A nonparametric, directed education classifier, the k-nearest neighbours invention, or KNN or k-NN, uses closeness to categorize or forecast the arrangement of individual dossier points.

**SVM:** The aim of the SVM treasure search out find high-quality line or resolution barrier that can separate n-spatial scope into classes because new dossier points maybe surely filling a place the right classification from now on. The optimum choice bound is a hyperplane.

**RF:** The Supervised Machine Learning Algorithm Random Forest is commonly secondhand in Classification and Regression tasks. It uses the most select as representative categorization and the average for reversion from diversified samples to construct resolution saplings.

**DT:** A decision tree is a diagram that shows all likely result for a likely recommendation utilizing a separate machine. Decision trees maybe tense manually or established accompanying specific spreadsheet or a drawings program. Decision saplings grant permission assist in concentrating the consideration when a group is met accompanying a resolution.

**Voting Classifier:** A representative classifier is an ML judge that prepares miscellaneous base models or assessors and forecasts taking everything in mind the results of each base judge. For each estimator product, amassing tests maybe linked accompanying vote determinations.

Move grasping the swing of applying CNN accompanying LSTM layers: Transfer of education is the process of asking earlier well-informed facts to new



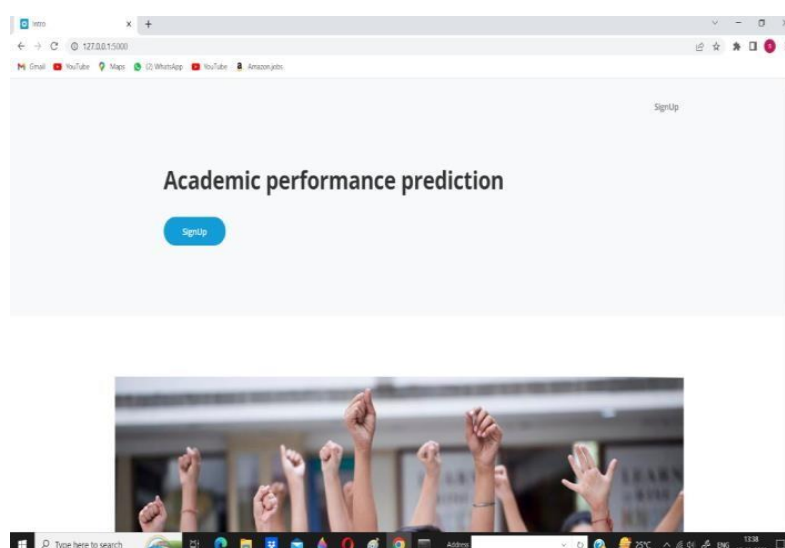
circumstances. Examples of movable knowledge: A graduate learns to answer polynomial equations in class and therefore applies that information to comparable exercise questions. A faculty member describes a difference of insane diseases in class. The fundamental premise of transfer education is honest: Apply the information of a model that was prepared on a abundant dataset to a tinier dataset. For object acknowledgment, we train only the last few levels that form prognoses and stop the network's early convolutional coatings.

LSTM Recurrent Neural Networks are a justifiable choice beyond any doubt order anticipation undertakings, but the arrangement is asserted on bearing adequate arrangement and experiment facts from a comparable transport.

**ABC:** ABC is a swarm intellect-located theory of probability search blueprint that mimics the endeavor of stinging insect swarms probing for meal. The ABC estimation divorces dear bees in a responsibility into three classifications: working bees, as known or named at another time or place thief bees, theatergoer bees, or scout bees, Each foodstuff beginning has singular contracted honey-making. This resources that skilled are as many meat beginnings as skilled are working bees.

**Whales Optimisation:** A novel addition game plan for handling growth issues is the Whale Optimisation Algorithm (WOA). This program involves three administrators that mirror the humpback mammal's seek to find something pillage, including pillage, and bubblenet scavenging conduct.

**Sailfish Optimisation:** The SFO is a sailfish-optimizing public-located metaheuristic invention. In this arrangement, the sailfish's site in the search room is individual of the issue variables, and they are destined to symbolize potential resolutions. Consequently, the answer scope community is produce every now and then.



**Fig.3: Home screen**

### 3. Experimental Results

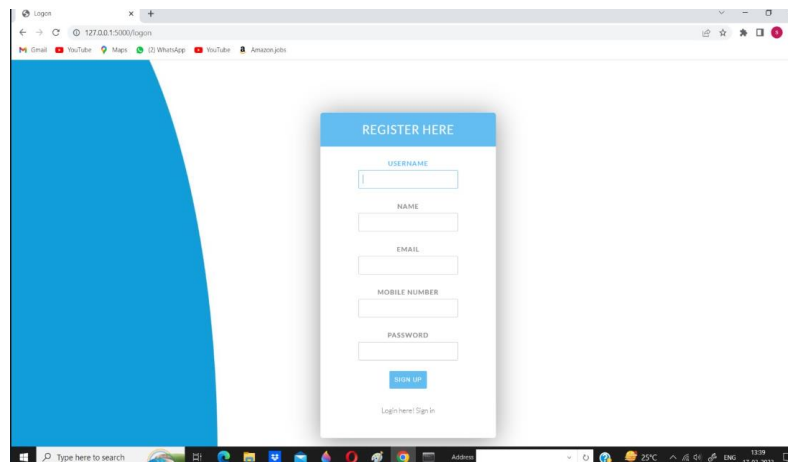


Fig.4: Registration page

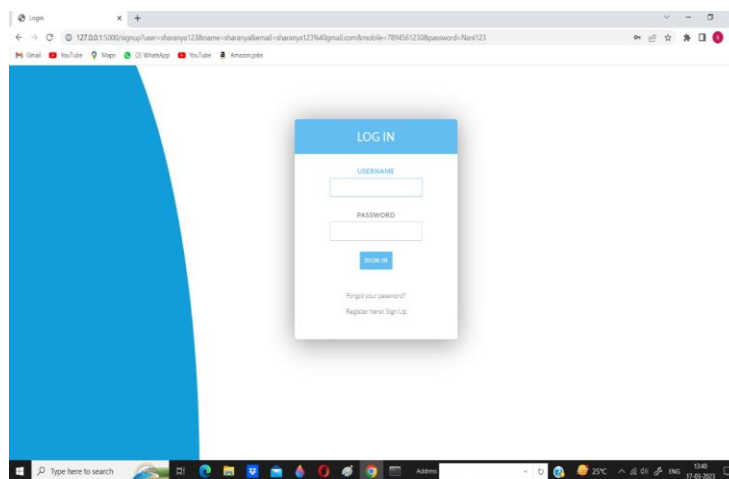


Fig.5: User Login page

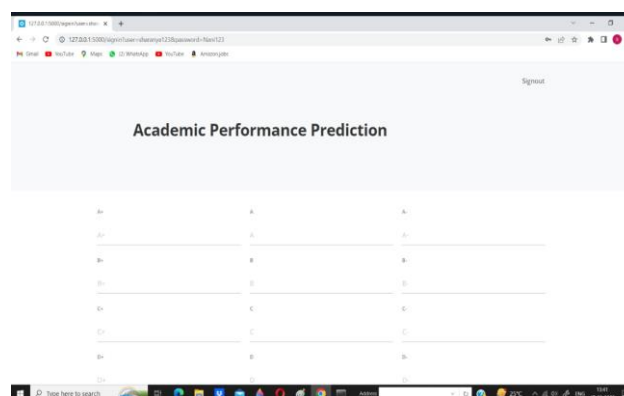
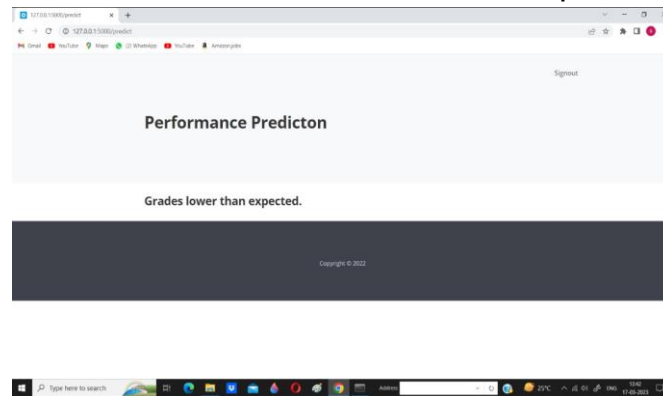


Fig.6: Academic performance prediction page



**Fig.7:Prediction page**

#### 4. Conclusion

In order to advance the educational process and address problems obstructing OL advancement, an SSL prediction model was put forth in this work. Data preparation, FS, ML classifiers, and ML model evaluation make up our model's four parts. It was a dataset.

Additionally, two ML classifiers, k-NN and SVM, were used to analyse data that was gathered through a questionnaire created specifically to elucidate how OL affects students. Some common SSL evaluation criteria were used in the current study, including faculty member obligations (online), online teaching and lectures, assessment methods, and E-Tests. 11 wrapper-based criteria were used to choose the best bundle of features. To distinguish between all traits and the chosen ones. The results showed that, when comparing k-NN and SVM, total accuracy based on the chosen features had increased by 2% and 8%, respectively. With FS algorithms, the total mean accuracy for k-NN and SVM was 100% when using all features. Fitness for exploration and exploitation is best achieved by the SFO algorithm with k-NN and SVM. Only four characteristics were determined. We come to the conclusion that only four factors (rather than the 20 features) had an impact on SSL and were therefore enough to accurately predict SSL with OL. "The lectures are delivered in an attractive format" and "The teaching method in this course motivates me to participate actively" are the two most basic yet important aspects. all features and the selected ones to sense the difference. The findings demonstrated that overall accuracy based on the selected features had been improved by 2% and 8% for k-NN and SVM, respectively, compared. With FS algorithms, the total mean accuracy for k-NN and SVM was 100% when using all features. Fitness for exploration and exploitation is best achieved by the SFO algorithm with k-NN and SVM. Only four characteristics were determined. We come to the conclusion that only four factors (rather than the 20 features) had an impact on SSL and were therefore enough to accurately predict SSL with OL. The fewest, yet most significant, features are: "The lectures are presented in an attractive style," "This course's teaching method encourages me to participate actively during the classes," "The form has been filled with degrees," and "Students trained on how

to solve exams online by designing an experimental quiz." This could assist HEIs in early SSL prediction.

### Future Scope

The Random Forest (RF) model might be assessed using 11 FS approaches on the predetermined constant dataset in the future since it can precisely coordinate the association among information and result with limitless intricacy.

## 5. References

- [1] G. Bonaccorsi, F. Pierri, M. Cinelli, A. Flori, A. Galeazzi, F. Porcelli, A. L. Schmidt, C. M. Valensise, A. Scala, W. Quattrocioni, and F. Pammolli, "Economic and social consequences of human mobility restrictions under COVID-19," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 27, pp. 15530-15535, Jul. 2020.
- [2] H. Fang, L. Wang, and Y. Yang, "Human mobility restrictions and the spread of the novel coronavirus (2019-nCoV) in China," *J. Public Econ.*, vol. 191, Nov. 2020, Art. no. 104272.
- [3] V. Singh and A. Thurman, "How many ways can we define online learning? A systematic literature review of definitions of online learning (1988-2018)," *Amer. J. Distance Educ.*, vol. 33, no. 4, pp. 289-306, Oct. 2019.
- [4] C. Khurana, "Exploring the role of multimedia in enhancing social presence in an asynchronous online course," Ph.D. dissertation, Rutgers Univ.-Graduate School, New Brunswick, NJ, USA, 2016.
- [5] M. A. Peters, H. Wang, M. O. Ogunniran, Y. Huang, B. Green, J. O. Chunga, E. A. Quainoo, Z. Ren, S. Hollings, C. Mou, S. W. Khomera, M. Zhang, S. Zhou, A. Laimeche,
- [6] W. Zheng, R. Xu, L. Jackson, and S. Hayes, "China's internationalized higher education during COVID-19: Collective student autoethnography," *Postdigital Sci. Educ.*, vol. 2, no. 3, pp. 968-988, Oct. 2020.
- [7] S. Raime, M. F. Shamsudin, R. A. Hashim, and N. A. Rahman, "Students' selfmotivation and online learning students' satisfaction among unitar college students," *Asian J. Res. Educ. Social Sci.*, vol. 2, no. 3, pp. 62-71, 2020.
- [8] S. Wilkins and M. S. Balakrishnan, "Assessing Student satisfaction in transnational higher education," *Int. J. Educ. Manage.*, vol. 27, no. 2, pp. 143-156, Feb. 2013.
- [9] A. Dutt, M. A. Ismail, and T. Herawan, "A systematic review on educational data mining," *IEEE Access*, vol. 5, pp. 15991- 16005, 2017.
- [10] N. Kapasia, P. Paul, A. Roy, J. Saha, A. Zaveri, R. Mallick, B. Barman, P. Das, and P.
- [11] Chouhan, "Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal, India," *Children Youth Services Rev.*, vol. 116, Sep. 2020, Art. no. 105194.

- [12] M. Canayaz, “MH-COVIDNet: Diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images,” *Biomed. Signal Process. Control*, vol. 64, Feb. 2021, Art. no. 102257.
- [13] Q. Al-Tashi, H. Rais, and S. Jadid, “Feature selection method based on grey wolf optimization for coronary artery disease classification,” in *Proc. Int. Conf. Reliable Inf. Commun. Technol. Springer*, 2018, pp. 257-266.
- [14] A. I. Hammouri, M. Mafarja, M. A. AlBetar, M. A. Awadallah, and I. Abu-Doush, “An improved dragonfly algorithm for feature selection,” *Knowl.-Based Syst.*, vol. 203, Sep. 2020, Art. no. 106131.
- [15] Q. Al-Tashi, H. M. Rais, S. J. Abdulkadir, S. Mirjalili, and H. Alhussian, “A review of grey wolf optimizer-based feature selection methods for classification,” in *Evolutionary Machine Learning Techniques*, 2020, pp. 273-286.