A NOVEL STRATEGY TO ASSESS AND EVALUATE THE PERFORMANCE OF CANDIDATES IN ONLINE TEACHING LEARNING PLATFORMS USING MACHINE LEARNING MODELS

UGC CARE APPROVED

ABSTRACT

The Online Teaching Learning (OTL) platform is an emerging strategy for the delivery and study of technical content in the fields of Engineering, Science, Mathematics, Arts, Social Science, and many other programs. This new OTL platform allows candidates to share posts, raise queries, answer others' queries, and show their responses to the queries using emotion-based reactions. Because of this new behavior of the candidates, the traditional assessment and evaluation system needs to be updated to suit the current scenario. To augment the above-said need for new evaluation techniques in the OTL platform, this work investigates a multi-dimensional OTL platform dataset for understanding candidates' behavior, interactions, and skills, projected by them in the OTL platform using machine learning techniques. For this analysis, the OTL dataset from a Brazilian University is taken for a model. This work focuses on the skill sets and their participation in the online discussion forum throughout the course, using machine learning models like Random Forest, Gaussian Naïve Bayes, XGB Regressors, SVC, and Logistic Regression. First, the Prediction of the successful completion of the online course by the candidate using their skills and their interactions of their participation in an online discussion forum is performed. The accuracy obtained from the Machine learning models increased by 7% for the Prediction of completion using soft skills than the Prediction using their online peer group interaction. Secondly, the study on the required skillsets of the candidate is carried out based on their continuous online participation throughout the course. The machine learning model shows that candidates who have creative, innovative skills and are ready to learn by themselves show more interaction using the post than other candidates. Thus, a new strategy to assess and evaluate the performance of candidates in the OTL platform using machine learning models is proposed using their online behavior from the beginning of the course to completion.

***Keywords**: Online classroom, self-learning skills, machine learning, MOOC.*

B. LAVANYA
Department of Computer Science
University of Madras
Guindy, Chennai

U. VAGEESWARI
Department of Computer Science
University of Madras
Guindy, Chennai

G. SASIPRIYA
Department of Computer Science
University of Madras
Guindy, Chennai

I. Introduction

The Online Teaching Learning (OTL) platform is an emerging strategy for the delivery and study of technical content in the fields of Engineering, Science, Mathematics, Arts, Social Science, and many other programs. This new OTL platform allows candidates to share posts, raise queries, answer others' queries, and show their responses to the queries using emotion-based reactions.When it comes to education, the term "pedagogy" refers to an educator's comprehension of how their students study. The teachers concentrate on delivering the curriculum to the pupils in a fashion that is tailored to meet the specific requirements of the pupils. Pedagogy requires exchanges in the classrooms between the instructor and the students, which have the potential to make a profound impression on the mind of the learner. A teaching and learning process can be carried out using any of the various pedogeological methodologies [1]. The following is a list of the top five most important pedagogical approaches:

Constructivist: Students actively participate in learning in this method.

Collaborative: Students form groups to solve problems and generate strategies, ideas, and products.

Integrative: Students are given an integrative learning environment to connect their learning across the curriculum.

Reflective: Students self-evaluate under the reflective technique.

Inquiry-Based Learning: Teachers must answer students' questions and create an environment where their ideas are challenged, enhanced, and refined.

Modes of Learning Platform:

The current state of information and communications technology has made it possible to create many kinds of educational platforms, which is a positive development for the field. Every single educational medium can be placed into one of these categories, depending on the specifics of the context in which it is used as In-Person Learning, In-Person Technology -Supported Learning, Online Learning, Offline Distance Learning, and Hybrid Learning [2].

In-Person Learning: This is a more conventional approach to education in which the Candidate or Student, as well as the Instructor or Teacher, are required to be physically present on campus. Learning takes place in a face-to-face setting using this approach. This approach does not make use of any digital resources at any point.

In-Person Technology -Supported Learning: It is also a form of learning that is done in person. Along with traditional materials, this approach makes use of digital resources as teaching assistants. Learning takes place in a face-to-face setting using this approach.

Online Learning: Under this strategy, each class is conducted entirely online, including the discussions and the lessons. The delivery of online education can either be done synchronously or asynchronously.

Offline -Distance Learning: Students who choose to get their education in this manner will have the option to have printed materials sent to their homes via the postal service by the educational institution in which they are enrolled. This will be done by the institution where they are enrolled.

Hybrid Learning: It is a combination of the education that one receives in a standard classroom environment and the knowledge that one can acquire at one's own pace through the means of the internet. In other words, it is a hybrid kind of education. It is possible to transmit educational content through the medium of the internet in either a synchronous or an asynchronous manner using either live or prerecorded sessions. Both approaches have potential advantages, each of their kind.

A massive open online course, commonly known as a MOOC, is a type of web-based online education program that is offered for free and is designed for a huge number of learners who are in a range of locations all over the world. MOOCs are becoming increasingly popular. Following the epidemic, it will be hard to avoid using online learning platforms. The candidate's behavior, approach, and engagement with the pears will be substantially different depending on which platform they choose to participate on. As a result of this shift in behavior on the part of the candidates, the traditional assessment and evaluation system needs to be updated to suit the current scenario.

The working procedure of both the candidate and the instructor in an OTL platform is illustrated in Figure 1. For the synchronous mode of learning, the instructor connects to the online platform remotely. The candidate also connects online at the same time. In the asynchronous method of learning, the instructor posts a recorded video, and the candidate can view it at their convenient time. The learning platform includes a discussion forum, which is an essential part of the whole system. A candidate can pose questions in the forum's discussion section, where both other candidates and the instructor can observe and respond to them.

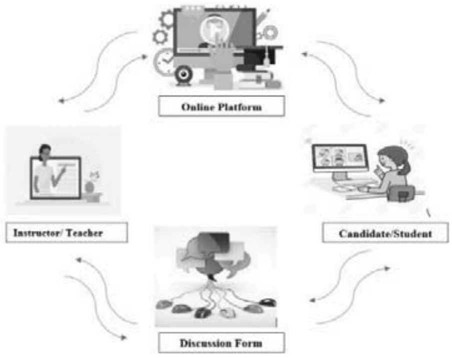

Fig 1. Online Teaching Learning (OTL) Platform working Process

To augment the above-said need for new evaluation techniques in the OTL platform, this work investigates a multi-dimensional OTL platform dataset for understanding candidates' behavior, interactions, and skills, projected by them in the OTL platform using machine learning techniques. For this analysis, the OTL dataset from a Brazilian University is taken for a model. This work focuses on the skill sets and their participation in the online discussion forum throughout the course, using machine learning models like Random Forest, Gaussian Naïve Bayes, XGB Regressors, SVC, and Logistic Regression.

First, the Prediction of the successful completion of the online course by the candidate using their skills and their interactions of their participation in an online discussion forum is performed. Secondly, the study on the required skillsets of the candidate is carried out based on their continuous online participation throughout the course. A new strategy to assess and evaluate the performance of candidates in the OTL platform using machine learning models is proposed using their online behavior from the beginning of the course to completion.

II. Related Work

The researchers investigated [3] [4] how award-winning online educators create, evaluate, and teach. To achieve this purpose, we developed a conceptual framework for online course design, assessment and evaluation, and facilitation and reviewed the relevant literature. Eight online American professors were surveyed. The Online Learning Consortium (OLC), the Association for Educational Communications and Technology (AECT), and the USDLA have recognized these online educators. The interviews showed that online educators use reverse design, consider student needs, and plan for their engagement. Teachers supported traditional and authentic assessments, rubrics for grading student work, course designs, quality assurance methods, surveys, learning analytics, and peer reviews. Timely feedback, presence, and communication were used by award-winning teachers.

To better audit the quality of English instruction delivered via the internet, this research proposes incorporating remote supervision with machine learning algorithms (IRS-MLA) [5]. Here, IRS-MLA models the integration of supervisory approaches into the classroom in response to the actual demands of English language instruction delivered online. Teachers' efficacy can also be evaluated by having both the teachers and their students seek and report on student performance and the learning process. This study presents research that evaluates traditional English-language online supervision, and it investigates the practical impact of this approach. The results of this investigation validate and demonstrate the efficacy of the model created in this paper [6]. In comparison to other models, the suggested IRS-MLA achieves the highest performance ratio (97.8%), accuracy (96%), efficiency (99.3%), and success rate (98%).

Within each procedure, we located and labeled the unsupervised or supervised Machine learning algorithms at work. Important facets of exams are explored from a Machine learning point of view, including authentication, scheduling, proctoring, and cheat/fraud detection. To better comprehend the function of machine learning in test preparation and its management of the post-exam process, the main qualities, such as Prediction of at-risk students, adaptive learning, and monitoring of students, are merged [7] [8]. While the ultimate purpose of such models is to forecast and intervene in real-world student achievement, this assessment highlights numerous important methodological inadequacies, including substantial filtering of experimental subpopulations, inefficient student model evaluation, and the use of experimental data [9].

Micro-incentives increased SET completion rates, according to student data. Psychological cues, micro-incentives plus prompts, or a no-information standard-practice control condition were randomly assigned to 36 instructors in Study two. The control condition (53.9%) had the lowest SET completion percentage, followed by prompts (64.5%) and micro-incentives with prompts (79.7%). Only micro-incentives differed from the control. These findings suggest that prompts and incentives can boost SET completion [10].

In this study, we build two different models, one to predict students' success on future exams and another to predict their overall performance. The models can be used to identify the determinants of learning success in massive open online courses (MOOCs). The findings demonstrate that both models can produce workable and reliable outcomes. For the student evaluation grades model, RF's RSME gain of 8.131 is the lowest, while GBM's RSME gain of 0.086 on average represents the highest accuracy [11].

III. Types of assessments

Normalized tests for determining grades and promotions are often misunderstood as assessments. Assessment is an umbrella term that covers not only standardized tests but a wide variety of other forms of evaluation as well. A wide range of methods is employed by teachers to evaluate their student’s progress at different points in the classroom. The purpose of assessment is to provide a clear picture of the learner's knowledge target, learning progress, and skill acquisition through the systematic collection of information that is evaluated, measured, and documented. The benefits of testing students go beyond simply revealing their abilities and weaknesses. Therefore, educators are better able to tailor their lessons to the individual learning styles of their students.

In addition to serving as a diagnostic tool, assessments can serve as a source of motivation for both students and educators. Both students and teachers benefit from the assessment process; students can showcase their abilities, and teachers can receive feedback to enhance their teaching. Beginning with a strategy for instructing students, the assessment cycle then moves on to the actual assessment. To build on that, step two is to implement the plan in the classroom. Third, evaluate the work's outcomes to see if they have been achieved as planned. Finally, it demonstrates how to put that knowledge to use. Assessment is critical for determining not only how much students have learned but also how well they can apply what they have learned to new and unfamiliar scenarios. That is why it's so useful for a wide range of pedagogical functions. Important forms of evaluation include.

A. Diagnostic Assessment or Pre-Assessment

Before beginning a new lesson, a new unit, a new course, or any other academic program, students go through either diagnostic testing or pre-assessment testing. This occurs in the time leading up to the beginning of the new instructional component. They are helpful tools for gathering information about the talents, knowledge, strengths, and shortcomings possessed by the individuals who are being evaluated in a certain context. After that, the material that is used in education is prepared in a way that is adapted to the specific requirements of the students.

B. Formative Assessment

These are the evaluations that are carried out concurrently with the learning that is taking place. Multiple instances of formative assessment will be carried out by the instructor through a topic, lesson, or course as the teacher moves through its entirety. They have been formatted in a way that makes them particularly useful for instructional purposes. Additionally, the two most important functions of formative assessment are the monitoring of learning and the provision of feedback to facilitate the modification of instructions. This is because formative assessment is used to help students learn. It is an approach to continuous learning that breaks down the material into manageable chunks and monitors how far along the student is in the process of acquiring new knowledge. Nevertheless, this kind of evaluation does not place the participants in any order or assign grades to them.

C. Summative Assessment

After a given period of instruction, this kind of evaluation seeks to determine the levels of learning that have taken place. It makes an effort to measure how effective learning takes place, as well as the student's level of proficiency and their level of accomplishment. This method accomplishes this goal by using assessments, assignments, and projects as the basis for specific grading and student ranking.

In addition, it assesses the learners' knowledge, skills, and attitude over the course of the entire program. As a result, it provides a glimpse into the performance of the student as well as the efficiency with which the instructor conveyed information. It reveals the extent to which students can meet the requirements of the assessment. In addition, the mark it receives is factored into the overall score for the unit. These are done after each unit has been completed. In addition to this, it offers data that can be used in the selection process for the subsequent level.

The ability of the summative assessment to adhere to logic and be relied upon is of the utmost significance. Additionally, it has the potential to supply information that has a formative value.

D. Confirmative Assessment

Confirmative assessment is a procedure that is performed one year after the finish of a certain educational session to determine whether or not the instructions were successful. The purpose of this evaluation is to determine whether or not the lessons were beneficial. The final aim here is to figure out whether or not the teaching is correct and whether or not the instructional tactics that are currently being employed are still successful. As a result of this, they provide a comprehensive type of summative evaluation, which is the fundamental objective of these exercises.

E. Norm-Referenced Assessment

In this section of the evaluation, teachers will measure a student's performance in comparison to a predetermined set of national norms that will represent the average. The norms will serve as an accurate representation of the mean. Take, for example, the typical grade that students in all the schools across the state receive in their English classes.

In addition to this, it assesses the student's performance in a competitive setting about the performance of their classmates. In this circumstance, the comparison of average grades is carried out with the other students that are enrolled at the institution. Because of this fact, the question test is also known as the Group Assessment or the Demographic Assessment.

F. Criterion-Referenced Assessment

Criterion-referenced examinations compare a candidate's current level of knowledge or skills to a set of previously established standards for academic achievement. It evaluates the academic needs and capabilities of the pupils at a certain stage of their educational development. The pupils' performance on these examinations is judged according to predetermined targets, requirements, or criteria. In other words, it analyzes the entirety of the instructional material for the course.

G. Ipsative Assessment

The learners' progress is tracked using their previous performance as a benchmark for the ipsative assessment tests. The learners evaluate their progress by comparing their current results to their prior ones. This kind of evaluation takes into account the fact that comparing students to their contemporaries isn't always a smart idea because it can affect the student sense of self-confidence. On the other hand, comparing current results with those of the learner's earlier endeavors helps develop the learner's total knowledge and personality.

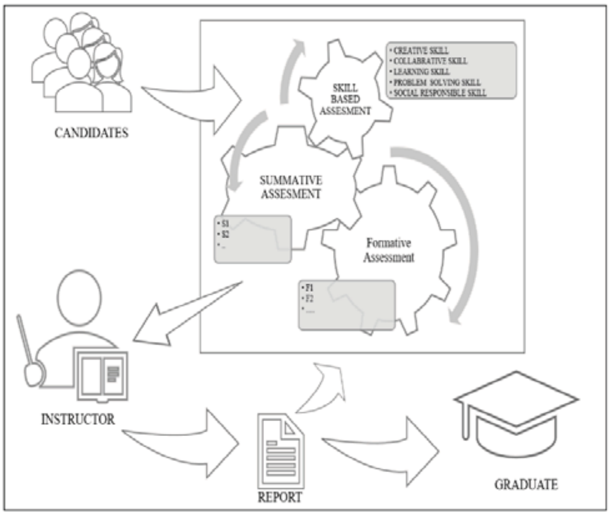

Fig 2: Proposed Skill-based Assessment framework

H. Work Integrated Assessment

This type of assessment is where the tasks and situations are closely associated with what you experience. It helps to develop students' skills and capabilities along with educational development. Also, this helps and supports educational staff in designing genuine assessments. It designs according to the need of the teacher.

IV. Proposed Method

The machine learning models have been broken into seven important steps collecting data, Preparing data, choosing a model, training a model, evaluating the model, parameter tuning, and making predictions. In this paper, data collected is from Brazilian University. By using this data and a few machine learning models, we have predicted the student's skills and posts.Figure2 represents the skill-based assessments and their importance.

Using this proposed skill-based assessment framework the candidates are assessed in three important dimensions. Formative assessment helps to evaluate the progress in each stage. Summative assessment helps to evaluate the overall understanding of the course content. Skill-based assessment helps to evaluate the skills projected by the candidate in the online platform. All these three are necessary to evaluate the candidate and grade them in an online platform. The Proposed framework incorporates all these assessment techniques and at the end gives the instructor a report about each candidate. The candidate either graduated successfully or was asked to retake some assessments based on the score in the report. Thus, this newly proposed framework ensures the efficiency of successfully completed candidates.

A. Random Forest

There are three primary hyperparameters for random forest algorithms that must be adjusted before beginning training. The node size, the number of features sampled, and the number of trees used are all examples. After that, you can use the random forest classifier to address your regression or classification issues.

Each decision tree in the random forest algorithm's ensemble is built using a sample of data selected from the training set via replacement, also known as a bootstrap sample. We'll return to this one-third of the training sample, which is known as the out-of-bag (OOB) sample, for testing purposes. Afterward, feature bagging is used to introduce yet another random event into the dataset and lower the correlation between the several decision trees. How the Prediction is arrived at varies with the nature of the underlying problem. Individual decision trees will be averaged for regression work, while the most common categorical variable will win out in a classification task and be used to determine the projected class. The final step is cross-validation, which uses the OOB sample [12] [13].

- The steps and pictures below show how the procedure works:
- First, take K sample points at random from the information in the training set.
- Second, create the trees of decision for the chosen data points (Subsets).
- Determine the size of the decision forest you wish to construct (N).
- Four, do the first two steps again.
- Determine the predictions of every decision tree for the newly added data points, and place the data into the category with the most votes.

B. Gaussian Naïve Bayes

A machine learning model that is used to differentiate between distinct types of objects based on particular characteristics is called a classifier. A probabilistic machine learning model that is used for classification tasks is called a Naive Bayes classifier. The Bayes theorem serves as the foundation for the majority of the classifier [14] [15].

Given that event B has taken place, we may use Bayes' theorem to calculate the likelihood that event A will also take place. In this case, Hypothesis A is being tested against Evidence B. In this case, it is assumed that the predictors and features can be considered separate entities. That is to say, the presence of one characteristic does not affect the other. Because of this, we refer to it as naïve.

C. SVM

An SVM classifier, also known as a support vector machine classifier, is a specific kind ofalgorithm that belongs to the field of machine learning. Its purpose is to assess and categorize data. The tasks of classification and regression are both within the purview of the support vector machine, which is an example of an algorithm for supervised machine learning. Finding the hyperplane that maximizes the difference between the two classes is how the Support Vector Machine classifier gets its job done. The support vector machine has numerous advantages over other machine learning algorithms, such as its resistance to noise and its capacity to manage big datasets. By utilizing kernel functions, SVM can be utilized to successfully tackle non-linear problems. After the data points have been mapped, the support vector machine will search for the hyperplane in this new space that is most suitable for dividing the data points into the two classes that are desired [16] [17].

D. XG Boost Regression

When it comes to developing supervised regression models, XGBoost is an effective method to use. If one is familiar with the objective function and base learners of XGBoost, then one maydeduce that this statement has some degree of truth to it. Both a loss function and a regularisation term are contained within the objective function [18] [19]. It provides information regarding the disparity between the observed values and the expected values or the degree to which the model findings deviate from the observed data. For regression problems, the reg:linear loss function is the most popular one used in XGBoost, whereas the reg:logistics loss function is the most common one used for binary classification. XGBoost is one of the methods that fall under the category of ensemble learning. Ensemble learning entails training and integrating individual models, sometimes referred to as base learners, to obtain a single prediction. When all of the predictions are combined, XGBoost anticipates having the base learners who are uniformly awful at the remainder. This ensures that when all of the bad predictions are canceled out and the better ones are added up, the final forecasts will be accurate.

E. Logistics Regression

The dependent or response variable in this logistic regression equation is denoted by logit(pi), whereas the independent variable is denoted by x. Maximum probability estimation is a method that is frequently utilized in the process of estimating the beta parameter, also known as the coefficient (MLE). This approach uses several iterations to examine a variety of alternative beta values to find the one that provides the greatest fit to the log odds [20] [21]. The log-likelihood function is the product of all of these iterations, and the goal of logistic regression is to maximize this function to obtain the most accurate parameter estimate. Once the optimal coefficient (or coefficients, if there is more than one independent variable) has been located, the conditional probabilities for each observation can be calculated, logged, and added together to produce a predicted probability. This can be done once the optimal coefficient (or coefficients, if there is more than one independent variable) has been located in the context of binary classification, a probability that is less than. Five will predict the value 0, whereas a probability that is larger than 0 will predict the value 1. After the model has been computed, it is best practice to evaluate the model's "goodness of fit," which refers to how accurately the model predicts the dependent variable. This evaluation should be done as soon as possible.

F. KNN

If you're looking for a simple Machine Learning method that uses the Supervised Learning approach, go no further than K-Nearest Neighbor. Assuming we have two groups, let's call them A and B, and a new data item x1, we want to know where it belongs. As such, a K-NN method is required to find a solution. K-NN is a useful tool for determining the general type of a dataset [22] [23]. Take into account the picture below: The K-NN working can be explained based on the below algorithm:

- Choose a neighborhood size, K.
- Find each neighbor's Euclidean distance up to K neighbors.
- Apply the Euclidean distance measure to find the K closest neighbors.
- The number of instances in each class can then be tallied from among these k nearest neighbors.

- Place the new information into the group with the highest average neighbor count.
- The concept is complete.

V. Results and Discussion

Candidates are graded according to their performance based on one of three separate criteria utilizing the recently suggested assessment and evaluation method. There are three different methods that can be used to evaluate the student’s knowledge and abilities: formative assessments, summative assessments, and skill-based assessments. The performance of candidates who are taking part in a skill-based examination is graded according to the skills that they demonstrate through online discussion forums. The teacher gives each candidate a score that is determined by how well they exhibit competency in five different categories of skills.

Table 1. Prediction Results by Posts

Model	Precision	Recall	F-1 Score	Accuracy
XGB	0.98	0.77	0.87	0.8
RF	0.94	0.92	0.96	0.93
LR	0.93	0.85	0.92	0.86
KNC	0.92	0.77	0.87	0.8
SVC	0.95	0.77	0.87	0.8
GNB	0.96	0.61	0.76	0.66

Table 2. Prediction Results By Skills

Model	Precision	Recall	F-1 Score	Accuracy
XGB	0.99	0.97	0.99	0.99
RF	0.98	0.96	0.97	0.98
LR	0.96	0.95	0.96	0.97
KNC	0.95	0.96	0.96	0.99
SVC	0.97	0.99	0.97	0.99
GNB	0.98	0.92	0.93	1

Skill-based assessment:

The originality and inventiveness of the candidate's posts, as well as the candidate's ability to provide novel solutions to both technical and non-technical problems, are both taken into consideration when determining the candidate's score for creative skill. The contributions that candidates make to team projects are evaluated to determine their scores in the Collaborative skills category. During this competency test, the candidate will be evaluated on their ability to work together with any member of the team, regardless of the similarities or differences that may exist between the members. One of the most important aspects taken into consideration when assigning a candidate, a score in the learning skills category is the extent to which they display an interest in expanding their knowledge base.Observing the candidate's ability to deal with ambiguity and inaccuracies in the post is a good way to judge whether or not they have the ability to come up with creative answers to challenging problems. In this day and age, it is more important than ever to be socially and culturally responsible. Because there is no monitoring system for social forms, the candidate should exercise responsibility when asking questions in the forum, and they should also exercise responsibility while answering, providing only responses that may be considered authoritative. The candidate's performance in this area will be graded by the teacher.

Accuracy, Precision, Recall, and F1 Score are the four metrics that are used to evaluate a model's Prediction.

Formula 1 presents the equation that must be used to calculate the precision metric.

$$\text{Precision} = \frac{TP}{TP+FN} \quad (1)$$

The equation needed to compute the recall metric is shown in Formula 2.

$$\text{Recall} = \frac{TP}{TP+TN} \quad (2)$$

Calculating the f1-score metric requires Formula 3.

$$\text{F1- Score} = 2 * \frac{precision*recall}{precision+recall} \quad (3)$$

The formula for determining the accuracy metric is given in formula 4.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

TP: The presence of positive cases was accurately predicted

TN: The absence of positive cases was accurately predicted.

FP: There is a misclassification of negative cases (wrong positive predictions)

FN: Cases that are positive are incorrectly categorized (wrong negative predictions)

Table 1 shows the Prediction Results from Various Machine Learning Methods Using Post. Table 2 shows the Prediction Results from Various Machine Learning Methods Using Skills. Figure 3 shows the Precision values by various machine learning methods using post and skill. It is evident from the figure that making a prediction based on one's skills yields better results than making a prediction based on performance by post.

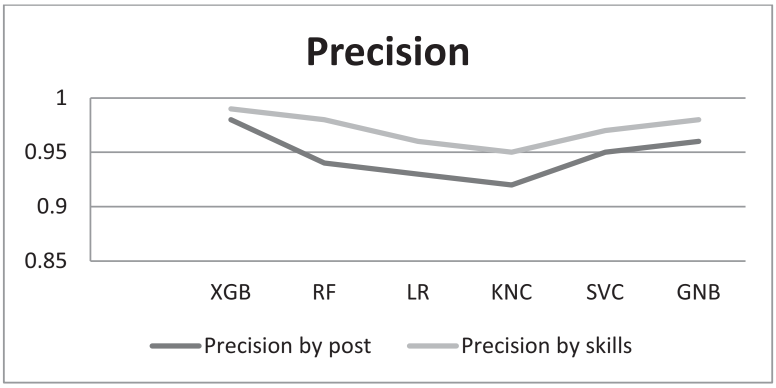

Fig 3. Precision by post and skills

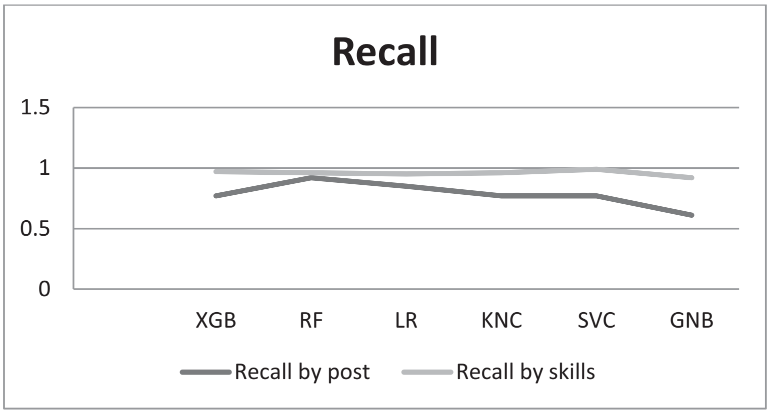

Fig 4. Recall by post and skills

Figure 4 displays the recall values obtained using a variety of machine-learning approaches that make use of post and skill data. A prediction that is made based on an individual's skills is shown to have a higher level of accuracy when compared to a prediction that is made based on performance by post.

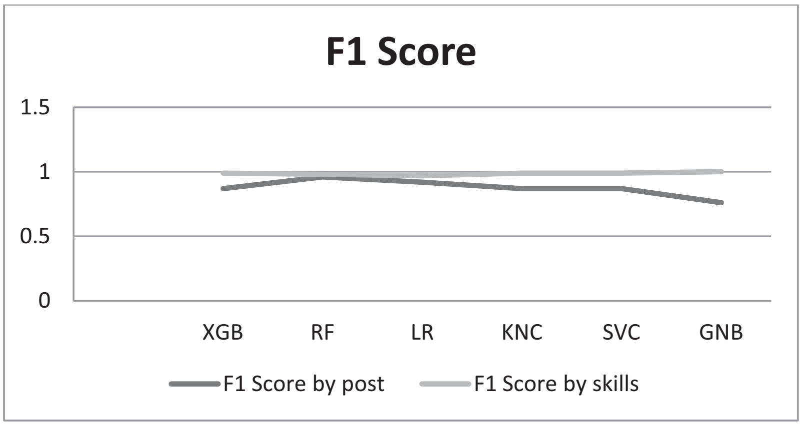

Fig 5. Precision by post and skills

Figure 5 displays the F1-score values derived from a number of different machine-learning approaches that make use of post and skill. The figure demonstrated that making predictions based on one's talents rather than making predictions based on performance by post results in a higher degree of accuracy.

Figure 6 displays the Accuracy values obtained using a variety of machine-learning approaches that make use of post and skill data. Figure 5 displays the levels of precision that can be achieved with the application of machine learning techniques that take into account both abilities and posts.

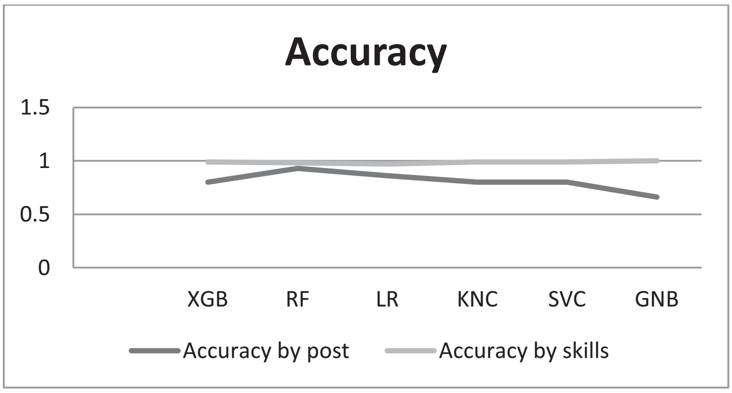

Fig 6. Accuracy by post and skills

Table 3. t- value and p-value by machine learning models

Machine Learning model	t-value	p-value
XGBRegressor	-3.85904	0.002408
Random Forest	-2.26579	0.02662
Logistic Regression	-2.28192	0.025958
KNC	-5.10159	0.000464
SVC	-4.4106	0.001127
Gaussian Naive Bayes	-5.8387	0.000194

To determine the statistical significance of factors in determining the candidate's course completion ratio, the statistical "t-test" is performed at a 5% significance level.

Hypothesis H0: Predicting the Candidate’s course completion using posts by them gives more accuracy than using skills assessed by the instructor.

Alternate Hypothesis H1: Predicting the Candidate’s course completion using posts by them does not give more accuracy than using skills assessed by the instructor.

For this different train, test sets of the same dataset are utilized. Table 3 is a tabular representation of the t values and p values obtained from several machine learning models.

By the F1 Score, the p values come in lower than 0.05. The lower the p-value, the greater the amount of support there is for hypothesis H1. The findings of the test provide support for the alternative hypothesis, indicating that skill-based assessment is statistically significant and that forecasting student performance using skills delivers superior accuracy is not something that happened by chance.

The Different machine learning models are used to check the required skillset for the candidate for better interaction with peers. Based on the total post count, the models show that candidate who has creative, innovative skills and is ready to learn by themselves shows more interaction using the post than other candidates. These skills are essential for performing better in discussion forums. Candidates with these skills project their knowledge in a better way, and at the same time, they will influence others to the right path.Because of this, the evaluation for the OTL platform ought to include both formative and summative assessments in addition to a skill-based evaluation.

VI. Conclusion

The aim of this research is to propose a strategy to assess and evaluate the performance of candidates in an OTL platform. For this, it investigates a multi-dimensional OTL platform dataset for understanding candidates' behavior, interactions, and skills, projected by them in the OTL platform using machine learning techniques. The prediction of the successful completion, of the online course by the candidate, in the OTL platform is performed. The accuracy obtained from the machine learning models increased by 7% for the prediction of completion using soft skills than the prediction using their online peer group interaction. Thus, the proposed strategy to assess and evaluate the performance of candidates, from the beginning of the course to completion in the OTL plat formought to include skill-based assessment as one of the essential components in addition to formative and summative assessments. This research work also investigates the required skill sets of the candidate to perform better in an online discussion forum. The candidate who has creative, innovative skills and self-learning skills shows more interaction using the post than other candidates. In order to effectively complete the course, the candidate must therefore enhance these competing skills. The outcomes of this research work are also checked for accuracy using the statistical t-test.

References

1. *C. Herodotou, M. Sharples, M. Gaved, A. Kukulska-Hulme, B. Rienties,E. Scanlon, and D. Whitelock, "Innovative pedagogies of the future:An evidence-based selection," in Frontiers in Education, vol. 4, p. 113, Frontiers Media SA, 2019.*
2. *R. Alcarria, B. Bordel, D. M. de AndrA˜ , et al., "Enhanced peerassessment in mooc evaluation through assignment and review analysis,"International Journal of merging Technologies in Learning (iJET),vol. 13, no. 1, pp. 206–219, 2018.*
3. *F. Martin, A. Ritzhaupt, S. Kumar, and K. Budhrani, "Award-winningfaculty online teaching practices: Course design, assessment and evaluation,and facilitation," The Internet and Higher Education, vol. 42,pp. 34–43, 2019.*
4. *P. M. Moreno-Marcos, C. Alario-Hoyos, P. J. Mu˜noz-Merino, and C. D.Kloos, "Prediction in moocs: A review and future research directions,"IEEE transactions on Learning Technologies, vol. 12, no. 3, pp. 384–401, 2018.*

5. S. Kaddoura, D. E. Popescu, and J. D. Hemanth, "A systematic review on machine learning models for online learning and examination systems," PeerJ Computer Science, vol. 8, p. e986, 2022.
6. W. Lu, G. Vivekananda, and A. Shanthini, "Supervision system of english online teaching based on machine learning," Progress in artificial intelligence, pp. 1–12, 2022.
7. X. Zhai, Y. Yin, J. W. Pellegrino, K. C. Haudek, and L. Shi, "Applyingmachine learning in science assessment: a systematic review," Studiesin Science Education, vol. 56, no. 1, pp. 111–151, 2020.
8. H. Luan and C.-C. Tsai, "A review of using machine learning approaches for precision education," Educational Technology & Society, vol. 24, no. 1, pp. 250–266, 2021.
9. J. Gardner and C. Brooks, "Student success prediction in moocs," UserModeling and User-Adapted Interaction, vol. 28, no. 2, pp. 127–203,2018.
10. N. Lipsey and J. Shepperd, "Examining strategies to increase studentevaluation of teaching completion rates," Assessment & Evaluation inHigher Education, vol. 46, no. 3, pp. 424–437, 2021.
11. R. Alshabandar, A. Hussain, R. Keight, and W. Khan, "Students performance prediction in online courses using machine learning algorithms," in 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1–7, IEEE, 2020.
12. M. Belgiu and L. Dr˘agut¸, "Random Forest in remote sensing: A review of applications and future directions," ISPRS journal of photogrammetry and remote sensing, vol. 114, pp. 24–31, 2016.
13. A.T. Azar, H. I. Elshazly, A. E. Hassanien, and A. M. Elkorany, "A random forest classifier for lymph diseases," Computer methods and programs in biomedicine, vol. 113, no. 2, pp. 465–473, 2014.
14. S. Xu, "Bayesian na¨ıveBayes classifiers to text classification," Journal of Information Science, vol. 44, no. 1, pp. 48–59, 2018.
15. W. Lou, X. Wang, F. Chen, Y. Chen, B. Jiang, and H. Zhang, "Sequence-based prediction of DNA-binding proteins based on hybrid feature selection using random forest and gaussian naive Bayes," PloS one, vol. 9, no. 1, p. e86703, 2014.
16. Y. Zhang, "Support vector machine classification algorithm and its application," in International conference on information computing and applications, pp. 179–186, Springer, 2012.

17. B. Gupta, S. Dhawan, and G. M. Mamdapur, "Support vector machine (SVM) research in India: A scientometric evaluation of India's publications output during 2002-19," Journal of Indian Library Association, vol. 57, no. 3, pp. 12–25, 2022.
18. X. Zhang, C. Yan, C. Gao, B. A. Malin, and Y. Chen, "Predicting missing values in medical data via xgboost regression," Journal of healthcare informatics research, vol. 4, no. 4, pp. 383–394, 2020.
19. K. D. Kankanamge, Y. R. Witharanage, C. S. Withanage, M. Hansini, D. Lakmal, and U. Thayasivam, "Taxi trip travel time prediction with isolated xgboost regression," in 2019 Moratuwa Engineering Research Conference (MERCon), pp. 54–59, IEEE, 2019.
20. M. Villalobos-Arias, "Using generalized logistics regression to forecast population infected by covid-19," arXiv preprint arXiv:2004.02406, 2020.
21. S. Sperandei, "Understanding logistic regression analysis," Biochemia Medica, vol. 24, no. 1, pp. 12–18, 2014.
22. Y. Wang and Z.-O. Wang, "A fast knn algorithm for text categorization," in 2007 international conference on machine learning and cybernetics, vol. 6, pp. 3436–3441, IEEE, 2007.
23. K. Alkhatib, H. Najadat, I. Hmeidi, and M. K. A. Shatnawi, "Stock price prediction using k-nearest neighbor (knn) algorithm," International Journal of Business, Humanities and Technology, vol. 3, no. 3, pp. 32– 44, 2013.