On the Design of Student Assessment Model Based on Intelligence Quotient using Machine Learning

NIKHILA KATHIRISETTY1, RAJENDRASINH JADEJA 2, DEEPAK GARG 3, HIREN KUMAR THAKKAR 4

1Department of Information Technology, Vardhaman College of Engineering, Hyderabad, India. (e-mail: nikhila@vardhaman.org)
2Department of Electrical Engineering, Marwadi University, Rajkot, Gujarat, India. (e-mail: rajendrasinh.jadeja@marwadieducation.edu.in)
3Department of Computer Science and Engineering, Bennett University, Noida, India. (e-mail: deepak.garg@bennett.edu.in)
4Department of Computer Engineering, Marwadi University, Rajkot, Gujarat, India. (e-mail: hirenkumar.thakkar@marwadieducation.edu.in)

Corresponding author: Hiren Kumar Thakkar (e-mail: iamhiren@gmail.com)

ABSTRACT The goal of this research is to figure out how to calculate academic achievements and students’ cognitive quotients for placement. This study will attempt to forecast students’ intelligence quotients or academic grades to measure the IQ of a student in a holistic manner using all kinds of parameters, from students’ academic records to input from their professors and even their family background, thus creating a dataset of 9000 instances with all these data. We implemented and trained multiple machine learning algorithms on the data and collected the outcomes to select the best algorithm. Students’ quantitative reasoning ability was selected as a parameter that could be assessed by their performance on aptitude tests. Certifications of the student during their bachelor’s degree have been considered, which would also give us an idea about the student’s critical and logical thinking ability. All the parameters were rated on a scale of 1-10. The driving motivation behind this investigation was to discover what parameters force a student to be placed in a company then the final overall “student score” is calculated to determine a student’s intelligence quotient. The final IQ score of the student-generated was graded on a scale of 0-3 and a suitable salary package range for the student was estimated giving the company a good idea of the student’s capability.

INDEX TERMS Intelligence quotient (IQ), Machine Learning, Data Mining.

I. INTRODUCTION

General knowledge can be portrayed as a build comprising particular intellectual limits. This limits license people to acquire information and tackle issues, whereas the intelligence quotient can be a score determined with the help of a set of standard tests designed to study the insights of people. "IQ" was first given by a psychologist named for the German term William Stern. A concrete degree of insight cannot be finished because of the complex nature of these “insights”. Quotient scores are related to factors such as parental budgetary status, nourishment, parental social status, parental environment, horribleness, and mortality [1]. “The ability to manipulate information is defined as intelligence (tasks requiring a higher level of cognitive complexity). "While no single test can assess general brain function, study findings can be used to evaluate more specific cognitive abilities."

[21]. Intelligence quotient scores were utilized to assess mental failure and survey work candidates. In examination settings, they are considered markers of wage and work execution. In addition, they are utilized to consider the scattering of psychometric insights for the populace and their connections to further variables. Since the early twentieth century, unrefined marks on Intelligence quotient tests of diverse populations have risen at a rapid rate of three IQ centers every decade, a phenomenon known as the Flynn effect. Some factors influence the IQ score of Intelligence quotients. They are reliable and valid. Reliability is unwavering quality is the estimation test consistency. Like all factual amounts, any particular evaluation of IQ contains a related standard blunder that measures precariousness, approximately the evaluation. For show day tests, the conviction interim can be generally 10 focuses and a detailed standard error of estimation can be as moo as roughly three focuses; it does not account for all possible sources of inaccuracy; the detailed standard blunder
may be something to be wary about. External influences such as little motivation or high anxiety also affect a person’s IQ test score, thereby compromising the accuracy of a mental failure search.

The validity alludes to the need for a predisposition. While IQ tests are by and huge considered to a degree many shapes of insights, they might be short enough to aid an accurate degree of comprehensive definitions of human intellect that include creative capacity and social understanding. Numerous complaints expressed that IQ test scores disregarded a few other noteworthy subjects of mental capacity. Despite these disagreements, clinical clinicians for the foremost portion regard IQ scores as having satisfactory quantifiable authenticity for various experimental purposes. There are differences based on groups in IQ measurements such as gender and race.

**Gender:** Females performed more frequently on errands related to verbal capacity, and fellows performed much better on jobs associated with the transformation of objects, classified under spatial capacity. Females outperformed males in predictability, and the Gender-specific IQ-predictive models were well fitted to two distinct data sets, demonstrating significant gender-related variability in intelligence neurological [1]. With repeated firing, human pyramidal neurons with higher IQs sustain fast action potential dynamics. These studies are the first to demonstrate that neuronal complexity, action potential dynamics, and effective information transfer from inputs to outputs within cortical neurons are related to human intelligence [2].

**Race:** Impacts of generalization danger have been proposed as a clarification for the contrasts in Intelligence quotient test executions between racial bunches, as they have issues relating to social qualification and get to education logistics. [3].

Some of the intelligence quotient tests are [4] :
1. Stanford Binet (SB-V) calculates the capability of reasoning such as abstract, quantitative, verbal, and even working memory [5].
2. Wechsler Adult Intelligence Scale (WAIS-IV) Mainly comprises two essential parts. The verbal portion incorporates a few tests to assess verbal leaning, lexicon capacity, sound-related recognition, and outstanding memory. This insight test’s execution classification includes visual consideration, visual recognition, and 3D perception.
3. Wechsler Intelligence Scale for Children for ages 6 to16 (WISC-R) is close to the WAIS which comprises some questions for children to get easy.
4. Leiter International Performance Scale Assesses nonverbal cognitive capacities that make insight trials best for children with hearing and talking abilities.

The Performance Intelligence Quotient (PIQ) is a test score that examines your child’s mental aptitude in dealing with nonverbal abilities. An IQ test typically consists of two key components: a verbal exam and a performance test. The Performance IQ known as PIQ includes more subtests with which time is an important role in scoring than the Verbal IQ (VIQ) index (Wechsler 2008). As a result, those with slower processing speeds yet with Normative verbal functions may exhibit a larger than expected discrepancy between VIQ and PIQ [6].

The driving motivation and contribution behind this investigation were to discover what parameters force a student to be placed in a company, and then the final overall "student score" is calculated to determine a student’s intelligence quotient. The final IQ score of the student-generated was graded on a scale of 0–3, and a suitable salary package range for the student was estimated, giving the company a good idea about the student’s capability. Our study is motivated by a goal to identify individuals who are placed based on criteria other than grades or CGPA. As a result, we became intrigued by the prospect of working on them. We should be aware of the elements that influence grades and placements. We work in this article on the factors, their percentage of influence on a student, and their placement in the various types of firms based on salary package.

In the current world, the idea of student assessment has grown beyond quizzes and tests. Existing assessment models may provide current performance, but they are incapable of predicting future performance and success. Studies are being conducted globally (at the national and international levels). National and international levels to understand how various factors influence the students’ assessment [7], [8]. These studies help in understanding the students from a better perspective than just their grades.

**Objectives:**

a. To measure the IQ of a student in a holistic manner by considering all types of parameters and the percentage of influence to obtain an accurate estimation.

b. To measure the final score of students’ IQ and their placement along with salary packages to place in different types of firms.

c. To understand the importance of IQ in today’s highly competitive world and how it can help one survive in the industry.

The expansions used in research are represented in Table 1

II. RELATED WORK

“A Placement Prediction System Using K-Nearest Neighbours Classifier” [9] the authors’ proposal for predicting the likelihood of an undergraduate student placed in a software firm by applying K-Nearest Neighbour’s, and compared the results of both in comparison to the findings from added algorithms such as logistic regression and, SVM. “Student Placement Analyzer: A Recommendation System Using Machine Learning” [10] Paper grants a recommendation system for predicting the likelihood of students having one of the statuses during their on-campus placements. The status that the students could achieve was that of Core Company, not Eligible, Dream Company, Mass Recruiters, and not showing

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
Interest in Placements. This aids the placement cell in an organization to classify certain students, pay to respond to them, and recover their interpersonal and methodological skills.

In [11], a comparison of various machine learning algorithms for predicting student outcomes to exemplify the proposed work, five distinct machine learning algorithms were applied. They’re Naive Bayes, SVM, XG-boost, KNN, Multilayer Perceptron. Before fitting these classifiers, the data was preprocessed. Other classifiers produced satisfactory results as well, with all of them over 80% accuracy. MLP had the highest accuracy of the five classifiers, at 86.25%. “Ensemble method to predict the impact of student intelligence quotient and academic achievement on” [12] the paper attempts at predicting the likelihood that if a candidate’s Intelligence quotient or his/her scholastic records will be playing a vital role in their upcoming placements. The authors even worked on a dataset that had 193 candidates and then applied a machine learning algorithm was utilized to compare the influence of students’ intelligence, attitude, and academic achievement on placements. The voting classifier architecture is used to categorize and forecast the likelihood of a candidate being placed successfully.

The aim is to establish a prediction for students of computer and electronics departments in engineering using two classifier systems: a fuzzy genetic algorithm and a decision tree. KNN is a straightforward classification method; it supplies each accessible case and categorizes novel cases clustered into the comparability calculation distance function. The training samples were characterized by numerical qualities in multiple dimensions. Each sample represented a point in a space with multiple dimensions. Almost all exercise samples were kept in a multidimensional sequence at one point. When a new sample is discovered, k training samples are near the K-nearest neighbour classification for the model space.

A high-dimensional space, used for classification is generated in a hyperplane or classes known as hyperplanes [13]. Owing to the standardized quality of their applications and the number of applications, an SVM was chosen for testing.

Some common sets of techniques such as random forest, Adaboosting and bagging are used in academic achievement for students further precisely, as AdaBoosting offers the best accuracy in 80 percent of the artificial neural networks.

| S.No | Abbreviation | Expansion                  |
|------|--------------|----------------------------|
| 1.   | KNN          | K-Nearest Neighbor         |
| 2.   | SVM          | Support Vector Machine     |
| 3.   | DT           | Decision Tree              |
| 4.   | RF           | Random Forest              |
| 5.   | RMSE         | Root Mean Square Error     |
| 6.   | MSE          | Mean Square Error          |
| 7.   | ML           | Machine Learning           |
| 8.   | IQ           | Intelligence Quotient      |
| 9.   | EQ           | Emotional Quotient         |
| 10.  | EDM          | Educational Data Mining    |
| 11.  | EI           | Emotional Intelligence     |

TABLE 1. TABULAR FORM FOR EXPANSIONS
Students’ performance improved and their GPA was determined using a data mining technique in a case study by (GPA), within the WEKA software. For three years, they collected data from post-graduate engineering students. CART has the highest accuracy (80.35%), then comes the LAD tree (75.00%), and the Naive Bayes classifier (80.35%)

The goal of this review is to examine how academics have handled the previous and modern developments in data mining research in education. In addition to determining the possibility in the area of education through machine learning. Numerous shortcomings of existing research are reviewed, and recommendations for future research are suggested. Some of the article details and future scope have been explained briefly in Table 2 to help in the current proposal for implementation and methodology based on existing drawbacks. In [8], posits two strategies by which families impact children’s academic achievement. 1. Parents seek to offer their children high-quality educational options, and increased educational opportunities contribute to improved academic achievement. 2. Children’s learning habits and academic achievement may be influenced by their parents’ parenting style and educational support for their children. They also discover that, when compared to rural pupils, urban students’ academic performance is more substantially influenced by their families’ socioeconomic condition. Each household received three questionnaire surveys: family, adult (aged >=16), children’s questionnaires (aged <16), CFP 2010, all of which had survey data. In [18], promote the Using peer assessment as a formative method to improve academic performance. The findings show that peer assessment is more effective than no assessment when it comes to implementing peer assessment in the classroom. A meta-analysis of experimental and quasi-experimental research (54 papers, k = 141) examined the influence of formative assessments on academic performance (54 studies, k = 141). In [19], to research how mentors acquire ideas about student ability in practice and create summative assessment decisions.

In [20], higher learning institutions recognize the importance of industry skills and emotional intelligence in attracting new engineering graduates. Confirmatory Factor Analysis (CFA) assesses the constructs’ reliability and validity. In [7], A study was conducted on Emotional Intelligence (EI) and Intelligence Quotient (IQ) and their combined impact on Academic Performance (AP). This study used Structural equation modeling analysis using SPSS AMOS for evaluating the data collected and it highlights the comparative analysis on the impact of EI and IQ towards the private and public sector students’ AP. As per the study, 43.6% of the variance in the public sector and 56.3% variance in private sector students’ performance is observed. For this study, measurement of EI is based on Quantitative Reasoning (QR), Specific Knowledge (SK), Communication skills (CS), Visual-spatial Processing (VSP), Memory, Motor coordination, Perceptual skills, and measurement of IQ is based on Self-Motivation, Self-Awareness, Empathy, Emotional Management, Interpersonal skills. This study recommends adding logical, reasoning, application-based questions to bring out the managerial skills and personality qualities needed for professional life and also suggests conducting classes on Role modeling, case studies for developing sharing and supporting others, and promoting an environment of empathy. This study draws important conclusions for university students in terms of cognitive and emotional values development to improve their academic performance.

III. METHODOLOGY
We built a custom data set with parameters including inputs from a student’s homeroom teacher with the parameters of asking doubts, completion of assignments, and his attention during the class, three subject teachers having parameters of attention during the class and asking doubts. The homeroom teacher has an extra parameter as she knows her class students well. From Table 1 the students’ previous academics (academic history from kindergarten to their bachelor’s degree) were also considered. Parents’ literacy background and their living area (if the area was developed with proper resources such as electricity, schools, internet, etc which can prove invaluable in a student’s IQ development). We also reviewed a student by two of the toppers of the in-class and two of his/her best friends. We received feedback from the toppers because they are unbiased and can truly assess how a student is in class. We have even considered the best friends because they know the student closely, even outside the classroom. The students’ quantitative reasoning ability was selected as a parameter that can be assessed by their performance on aptitude tests. Certifications of the student during his bachelor’s have been considered which would also give us an idea about the student’s critical and logical thinking ability. All the parameters were rated on a scale of 1-10. Then the final overall “student score” is calculated to determine a student’s intelligence quotient. The Proposed workflow diagram is shown in Figure 2.

The collection of data is used in this article taken after an educational program designed to improve learning using digital technologies. The data were collected using a racking tool called API. To track their learning, students could use the API as if they were reading an article or viewing a teaching video. After compiling the dataset, it was pre-
| S. No. | Title | Approach | Future scope | Method / Technique | Year |
|-------|-------|----------|--------------|--------------------|------|
| 1.    | Factors Influencing [21] | Data Mining Algorithms | The state of the art has been compiled into a systematic approach that provides educators with step-by-step recommendations for making decisions and setting parameters. | P-Correlation | 2020. |
| 2.    | A PPS Using KNN Classifier [9]. | K-Nearest Neighbours, Logistic Regression, SVM, Prediction | Additional work can be done by employing various algorithms, which may result in improved outcomes. | i. Data mining ii. Instructional data iii. EDM | 2016 |
| 3.    | Student Placement Analyzer [10] | Decision Tree Classifier | Other algorithms can be used to improve the findings in future studies. | Machine Learning | 2017 |
| 4.    | Comparative study [11] | ML Techniques | Environmental considerations, as well as regular exam results, will be taken into account. With a huge dataset, several neural network topologies such as CNN, RNN, and others will be used. | i. Naive Bayes ii. K-Nearest Neighbours iii. SVM iv. XG-boost v. Multi-layer Perceptron | 2021 |
| 5.    | Student performulator [12] | Machine learning | The goal is to obtain the best weight for the characteristics to achieve the highest level of precision in the prediction framework. | i. EDM ii. Regression model | 2021 |
| 6.    | New DM model [13] | EDM algorithms | For a set of traditional learning models, hyperparameter tuning outperforms default settings. | Bayesian Optimization | 2019 |
| 7.    | Ensemble method to predict [14] | DB-SCAN algorithm, Random Forest. | Recommended to increase the number of correlative features with a size of the dataset, helps with performance, so it will be evaluated externally. | i. ML ii. Voting Classifier | 2021 |
| 8.    | Analysis of factors influencing [17] | Learning analytics. | improve the predictive power. | Learning analytics | 2020 |
| 9.    | How does family [8] | Multiple regression | This can be explained by the greater variation in family backgrounds and educational opportunities in urban areas compared to rural counterparts. | CFPS2010 | 2018 |
| 10.   | The impact of Peer [18] | Feedback Formative assessment | To identify the contextual and pedagogical aspects that influence the efficacy of peer assessment. | Meta-Analysis | 2019 |
| 11.   | Mentor Judgement [19] | PADs | Understand concerns better so that any measures have the greatest impact on areas of greatest risk, resulting in a student who is "fit for registration" at the end of their studies. | Mixed Methods | 2017 |
| 12.   | EI and relationship [20] | standardized survey questionnaire | t is suggested that other dimensions for employability skills and emotional intelligence be included, such as planning and organizing, problem-solving, self-management, learning, initiatives, stress management, and the ability to adapt to change. | CFA | 2019 |
| 13.   | A comparative analysis [7] | SEM | recognizing any indirect effect of EI and IQ in conjunction with other parameters as indicated in prior research. | Five-point scale Likert-type | 2019 |

**TABLE 2. LITERATURE SURVEY SUMMARY TABLE**
processed with attributes and features. After processing the data, machine learning models will be developed, and the results will be forecasted using advanced features in this analysis, as shown in Figure 1.

Some methods, such as distance-based, allow each item plotted in a certain class to be seen in the same way as other items in that class and discriminated from objects in other classes. There are two classification techniques based on distance such as:

- **Simple approach:** In this case, each class is presumed to be represented by its center. A novel thing can be added to a class with the highest possible resemblance charge.
- **KNN:** This is abbreviated as K-nearest neighbour. The parameters are not parametric approaches that rely on distance measurements. It can keep all accessible cases, and if a new instance arrives, it can be classified using the distance function.
- **DT:** This is abbreviated as a decision tree. These types of algorithms are necessary to represent the classification process for tree development. This categorization method requires two phases.
  a. Create a Decision Tree.
  b. Apply DT to a database EDM community.

Data mining is useful and productive decision-making. Classification is a widely used basic data-mining approach. Understanding classification requires knowledge of training data. The two phases of classification in data mining are:

- For training, development model.
- Using testing data, evaluate the model.

Different categorization methods emerge as a result of algorithms:

- One of the most important aspects of EDM is its use of the DM approach. They can be classified into two groups based on the following purposes:
  - Focusing on verification (hypothesis test, the goodness of fit, and analysis of variance are examples of traditional statistics)
  - Focuses on innovation (classification, clustering, prediction, relationship mining, neural networks, web mining, and other types of prediction and description).

Data mining in education involves the utilization of machine learning, data mining technologies, and various statistical approaches. The EDM is a method used to develop essential information from an educational setting. Student success is expected through the introduction of educational data mining for model development. This has led researchers to explore various mining techniques to develop current methods. Before a forecast of student success slowly began to dominate, EDM was the foremost research field in education. The most important goal of this project is to analyze and forecast students’ success in producing better outcomes. This study introduces ensemble methods to improve different classification algorithms by adding a new category called behavioral features to predict student results. An e-learning program provides an educational data kit. Data were collected using a tracking tool. The features obtained in these categories are academic history, demographic characteristics, and comparability.
Behavioral traits are a new classification correlated with student involvement in the education system cycle. This study used educational data to predict undergraduate success in academics. Therefore, this prototype examined the effects of students’ learning and behavioral characteristics on their academic success. This work was carried out through an application in data mining which is recognized as classification. The K-nearest neighbour, decision Tree, and support vector machine are three classifications used here. Some collaborative practices such as random forest, bagging and boosting improve the accuracy of the classification efficiency of a student’s performance model.

IV. EXPERIMENTAL RESULTS
For the experimental study, we have considered the customized dataset described as follows.

\[
\text{Precision} = \frac{TP}{TP + FP}, \quad (1)
\]

\[
\text{Recall} = \frac{TP}{TP + FN}, \quad (2)
\]

\[
\text{Accuracy} = \frac{TN + TP}{TN + FP + TP + FP}, \quad (3)
\]

Where
• TN is True Negative.
• TP is True Positive.
• FN is False Negative.
• FP is False Positive.

\[
F1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\quad (4)
\]

A. DATASET DESCRIPTION
Using all of the data, created a data set of 9000 instances, taking 80:20 percent of training and testing data into account. We used the data to develop and train different machine learning algorithms, then assessed the results to decide which one worked best. A student’s IQ can be measured holistically using a variety of characteristics ranging from academic records to feedback from academics and even family background. For the dataset, attributes are considered initially which are taken on a grade scale of 10. All attributes are –
• Parental Academic History [PAH].
• Previous Academics (Academic history from kindergarten to his Bachelor’s degree) [PA].
• Review of the student by his current 8th semester class teacher, which has three attributes of student listening to the class, asking doubts, and his performance in the assignments (Teacher 1) [RCT].
• Review of the student by his current 8th semester, three subject teachers, with two attributes of students listening to the class, asking doubts. (Teachers 2, 3, and 4) [RST].
• Certifications of the student from his 10th standard to his bachelor’s [RC].
• Residential area (if the area is developed with the Internet, electricity, and other essentials) [RA].
• A score is given for the student’s quantitative reasoning skills (based on tests conducted during his bachelor’s degree) [RQRS].
• Student score [SS].

The process used for the calculation of the student score,
• If the generated score is 0-6, then we assign a grade of 0 (not eligible).
• If the generated score is 7-8, then we give grade 1 (estimated salary package ranges from 3-5.5 LPA).
• If the generated score is 9, then we give a grade of 2 (estimated salary package ranges from 5.5-8 LPA).
• If it is 10, then we give a grade of 3 (estimated salary package ranges > 8 LPA).
• The data was divided into two parts: 80% training data and 20% test data.
• Review of students by two of his toppler friends in class.
• Review of the student by two of his best friends.

ALGORITHMS USED:

B. LOGISTIC REGRESSION
Linear regression auditing is a technique used to forecast one variable’s value depending on the value of another. The variable that needs to be expected is known as the dependent variable, which utilizes to predict the value of the other variable as the independent variable. Linear regression was performed using Microsoft Excel or by utilizing diverse factual computer program bundles that rearrange the method of straight relapse conditions, linear regression models, and linear regression equations. Using the logistic function, logistic regression helps us measure the relationship between categorical independent and dependent variables using probability estimation, which is the cumulative distribution function. This is a special case of the generalized linear model and is therefore comparable to linear regression. However, the model is based on assumptions.

\[
\log \left[ \frac{p}{1-p} \right] = (\beta_0 + \beta(age)) \quad (5)
\]

This is the formula used in LR. The odd ratio is \( \frac{p}{1-p} \) in this case. When the log of the odds ratio is positive, the likelihood of success is always greater than 50%.

The results of Table 4 are taken as a reference to build the confusion matrix in Figure 3, for better visualization of the results. From Table 4, it is clearly explained that accuracy is 76% and in existing research, accuracy is 75%. [11].

After applying the bagging technique, the results are,
Table 5 results are taken as a reference to build the confusion matrix in Figure 4, for better visualization of the results.
After applying boosting results are,
Table 6 results are taken as a reference to build the confusion matrix in Figure 5, for better visualization of the results.
### TABLE 3. DATA AND DATA FORM OF THE DATASET WITH THE ATTRIBUTES AND PERCENTAGES [21].

| Feature category | Features | Description | % |
|------------------|----------|-------------|---|
| Demo graphical   | a. Parental Academic History [PAH]. | a. Parents’ Education and Occupation. | 25% |
|                  | b. Review of Residential Area [RRA]. | b. Place of residences like Rural or Urban | |
| Academic Behavioral Background | a. Previous Academics [PA]. | a. Academic Records from kindergarten to Bachelor. | 44% |
|                  | b. Review with Quantitative Reasoning Skills [RQRS]. | b. A score is given for the students’ quantitative reasoning skills based on tests conducted during their bachelor’s Degree. | |
| UG Score         | • Student Score [SS]. | • Under Graduate Score. | 17% |
| Participation of Professors and Friends | a. Review with Toppers [RT]. | a. Review of the Student by two Topper Friends. | 11% |
|                  | b. Review with Friends [RF]. | b. Review of the Student by two Best Friends. | |
|                  | c. Review with Class Teacher [RCT]. | c. Review of the student by current 8th Semester class Teacher, which has three attributes of student listening to the class, asking doubts and performance in the assignments (Teacher 1). | |
|                  | d. Review with Subject Teachers [RST]. | d. Review of the student by current 8th Semester three Subject Teachers, with two attributes of student listening to the class, asking doubts and performance in the assignments (Teacher 2, 3 and 4). | |
| Behavioral       | • Review of Certifications [RC]. | • Certifications of the student during Bachelors. | 3% |

### TABLE 4. RESULTS FOR LOGISTIC REGRESSION

| Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|--------------|------------|----------|
| 0.76             | 0.76     | 0.42 | 0.17 | 0.82         | 0.82       | 0.81     |

### TABLE 5. RESULTS OF LOGISTIC REGRESSION AFTER BAGGING

| Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|--------------|------------|----------|
| 0.75             | 0.67     | 0.49 | 0.24 | 0.76         | 0.75       | 0.73     |

### TABLE 6. RESULTS OF LOGISTIC REGRESSION AFTER BOOSTING

| Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|--------------|------------|----------|
| 0.73             | 0.73     | 0.45 | 0.20 | 0.73         | 0.79       | 0.74     |

### C. DECISION TREE CLASSIFIER

As supervised learning, decision trees, or DTs, are utilised and non-parametric models for regression and classification. The aim is to develop a model for predicting target variable values based on data attributes using simple decision rules. The use of a decision tree has many advantages.
a. Are very easy to implement and are not complex and can be visualized.
b. Many techniques generally need data normalization, dummy variables, and no blank values.
c. The training cost of the decision tree is proportional to the number of data points used.
d. Decision trees can handle both categorical and numerical data.
e. These trees can be handled multi-output problems.
f. Decision trees use a white-box model. Suppose that any model can predict the situation; the required explanations are explained without any complexity by Boolean logic. In contrast, the results of the black box model are difficult to interpret.
g. The decision-tree model was validated using statistical tests. Therefore, making it possible for accounting for the reliability of the model.
h. The decision tree even if the assumptions provided by the true model from which the data were created are violated, the model performs well.

• **Entropy**  Entropy of the information under consideration is a measure of its randomness. The more entropy there is, the more difficult it is to draw any conclusions from the data. A coin flip is an example of an action that generates random data.

Mathematically, Entropy with one attribute is represented as:

\[
H(X) = - \sum_{i=1}^{c} p_i \log_2 p_i
\]  

(6)

where S \rightarrow current state, and \( p_i \) probability of an event \( i \) of state \( S \) or percentage of class \( i \) in a node of state \( S \).

Mathematically, Entropy with multiple attributes are represented as:

\[
E(T,X) = \sum_{c \in X} P(c) E(c)
\]  

(7)

where T \rightarrow current state and X \rightarrow Selected Attribute.

• **Information Gain**  A statistical parameter, often known as IG, determines how successfully a given attribute distinguishes training instances depending on their classification aim. The key to designing a decision tree is to choose a characteristic that provides the highest information gain while having the lowest entropy.

\[
IG = Ent(Before) - \sum_{j=1}^{K} \frac{\omega_j}{\sum_{j=1}^{K} \omega_j} \log_2 \omega_j
\]  

(8)

• **Gini Index**  The Gini index is a cost function used to evaluate dataset splits. The outcome is calculated by subtracting the total of the squared probabilities of each class from one. It favours larger partitions because they are simpler to build, whereas information gain prefers smaller partitions with varying values.

\[
Gini = 1 - \sum_{i=1}^{c} (p_i)^2
\]  

(9)

\[
GainRatio = \frac{IG}{SplitInfo}
\]  

(10)

\[
Ent(Before) - \sum_{j=1}^{K} \frac{\omega_j}{\sum_{j=1}^{K} \omega_j} \log_2 \omega_j
\]  

(11)

where "Before" refers to the dataset before the split, "K" refers to the number of subsets produced by the split, and (j,after) refers to subset \( j \) after the split.

The - decision tree classifier is capable of conducting multi-class classification on a dataset. The classifier receives two arrays as input: an array \( X \), sparse or dense, of shape (n samples, n features) containing the training data, and an array \( Y \) of integer values, shape (n samples, containing the class labels for training).

Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score
---|---|---|---|---|---|---
| 0.96 | 0.95 | 0.95 | 0.19 | 0.03 | 0.96 | 0.96

**TABLE 7. RESULTS FOR DECISION TREE CLASSIFIER**

Table 7 results are taken as a reference to build the confusion matrix below in Figure 6, for better visualization of the results.

After applying bagging results are,

Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score
---|---|---|---|---|---|---
| 0.95 | 0.94 | 0.20 | 0.04 | 0.95 | 0.95 | 0.95

**TABLE 8. RESULTS OF DECISION TREE CLASSIFIER AFTER BAGGING**
Table 8 results are taken as a reference to build the confusion matrix below in Fig. 7, for better visualization of the results.

![Confusion Matrix for Decision Tree Classifier after Bagging](image)

**FIGURE 7. CONFUSION MATRIX FOR DECISION TREE CLASSIFIER AFTER BAGGING**

After applying boosting results are,

| Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|---------------|------------|----------|
| 0.95             | 0.95     | 0.19 | 0.37| 0.96          | 0.96       | 0.96     |

**TABLE 9. RESULTS OF APPLYING DECISION TREE CLASSIFIER AFTER BOOSTING**

Table 8 results are taken as a reference to build the confusion matrix below in Figure 8, for better visualization of the results.

![Confusion Matrix for Decision Tree Classifier after Boosting](image)

**FIGURE 8. CONFUSION MATRIX FOR DECISION TREE CLASSIFIER AFTER BOOSTING**

**D. K-NEAREST NEIGHBOURS**

K-Nearest Neighbours (KNN) is a type of supervised ML algorithm, which is one of the simplest ML algorithms. KNN was used for regression and classification. However in most cases, they are used for classification. KNN does not make any assumptions regarding elemental data. Therefore, it is called a nonparametric algorithm. At the classification time, KNN stores the dataset but does not learn from the training, based on which it accomplishes actions within the dataset.

Therefore, it is also called a lazy learning algorithm. To select the correct K value, we would have to out the algorithm multiple intervals with various K values and choose K, which is the number of errors encountered, as well as maintain the to accurately make predictions with the data it has never encountered before. The predictions become less stable as the value of K decreases to one and as the value of K increases, the stability of the predictions also increases because of majority voting and averaging. Therefore, the prediction accuracy increased.

Advantages of using KNN are,

a. Easy to understand and implement.

b. Not necessary to make assumptions, tune multiple parameters, and build a model.

c. The versatility of the algorithm allows it to be used in regression, classification, and search problems with KNN neighbors for K values ranging between 1 to 19.

\[
d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}
\]  

(12)

To find the distance between any two points.

| K Value | Accuracy |
|---------|----------|
| K=1     | 0.93061  |
| K=2     | 0.93326  |
| K=3     | 0.93955  |
| K=4     | 0.94167  |
| K=5     | 0.94487  |
| K=6     | 0.94737  |
| K=7     | 0.94817  |
| K=8     | 0.94697  |
| K=9     | 0.94653  |
| K=10    | 0.94605  |
| K=11    | 0.94078  |
| K=12    | 0.94497  |
| K=13    | 0.94555  |
| K=14    | 0.94635  |
| K=15    | 0.94625  |
| K=16    | 0.94494  |
| K=17    | 0.94461  |
| K=18    | 0.94392  |
| K=19    | 0.94863  |
| K=20    | 0.94299  |

**TABLE 10. RESULTS FROM APPLYING VALUES OF K RANGING FROM 1 TO 20**

In Table 10 shows the results from applying the values of K ranging from 1 to 20.

KNN Classifier with 19 Neighbours i.e. K=19.

Table 11 results are taken as a reference to build the confusion matrix in Figure 9, for better visualization of the results.

KNN Classifier bagging with 3 Neighbours i.e. K=3.

Table 12 results are taken as a reference to build the confusion matrix in Figure 10, for better visualization of the results.
TABLE 11. RESULTS FOR KNN CLASSIFIER WITH K= 19

| Model Test Score | R² Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|---------------|------------|----------|
| 0.95             | 0.95     | 0.19 | 0.03| 0.96          | 0.96       | 0.96     |

FIGURE 9. CONFUSION MATRIX FOR KNN WITH K AS 19

1) Linear SVM
Find an ideal line splitting the data into classes. This ideal line is called the hyperplane. Primarily, from both classes, we obtain nearer to the line, and these points are called support vectors. Subsequently, we calculated the distance between the line and the support vectors. This distance is referred to as margin. A hyperplane with the maximum margin is referred to as the optimal hyperplane.

2) Nonlinear SVM
In linear SVM, data can be split by using a straight line, but for non-linear data, data cannot be split by a straight line. To split these data points, another dimension was added. Two dimensions are used in linear data that is, x and y and the third dimension is added to nonlinear data, referred to as z. This can be calculated as: 

\[ z = x^2 + y^2. \]

Equation of hyperplane:

\[ H : \omega^T(x) + b = 0 \]  

Hyperplane positive group:

\[ \omega^T(\phi(x)) + b > 0 \]  

Hyperplane negative group:

\[ \omega^T(\phi(x)) + b < 0 \]  

The product of a predicted and actual label would be greater than 0 (zero) on correct prediction, otherwise less than zero.

\[ y_n[\omega^T\phi(x) + b] = \begin{cases} 0 & \text{if correct} \\ < 0 & \text{if incorrect} \end{cases} \]  

TABLE 12. RESULTS FOR KNN CLASSIFIER WITH A K=3

| Model Test Score | R² Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|---------------|------------|----------|
| 0.93             | 0.93     | 0.21 | 0.04| 0.95          | 0.95       | 0.95     |

FIGURE 10. CONFUSION MATRIX FOR KNN AFTER BAGGING

In existing research [9], KNN accuracy is 78.57%, and in current research, KNN accuracy is improved to 95%, as clearly explained in Table 10.

E. SUPPORT VECTOR MACHINE CLASSIFIER
In classification, SVM is a linear model with regression problems that can find solutions to various mathematical problems, including linear and non-linear problems. The objective is to create a hyperplane that splits the classes from the data or the best line.

SVM is classified into two types:

1) Linear SVM
Find an ideal line splitting the data into classes. This ideal line is called the hyperplane. Primarily, from both classes, we obtain nearer to the line, and these points are called support vectors. Subsequently, we calculated the distance between the line and the support vectors. This distance is referred to as margin. A hyperplane with the maximum margin is referred to as the optimal hyperplane.

2) Nonlinear SVM
In linear SVM, data can be split by using a straight line, but for non-linear data, data cannot be split by a straight line. To split these data points, another dimension was added. Two dimensions are used in linear data that is, x and y and the third dimension is added to nonlinear data, referred to as z. This can be calculated as:

\[ z = x^2 + y^2. \]

Equation of hyperplane:

\[ H : \omega^T(x) + b = 0 \]  

Hyperplane positive group:

\[ \omega^T(\phi(x)) + b > 0 \]  

Hyperplane negative group:

\[ \omega^T(\phi(x)) + b < 0 \]  

The product of a predicted and actual label would be greater than 0 (zero) on correct prediction, otherwise less than zero.

\[ y_n[\omega^T\phi(x) + b] = \begin{cases} 0 & \text{if correct} \\ < 0 & \text{if incorrect} \end{cases} \]  

SVM Classifier results are,

| Model Test Score | R² Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|---------------|------------|----------|
| 0.99             | 0.98     | 0.09 | 0.008| 0.96          | 0.96       | 0.96     |

TABLE 13. RESULTS FOR SVM CLASSIFIER

Table 13 results are taken as a reference to build the confusion matrix shown in Figure 12, for better visualization of the results.

After applying bagging results are,

| Model Test Score | R² Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|----------|------|-----|---------------|------------|----------|
| 0.99             | 0.99     | 0.07 | 0.005| 0.99          | 0.99       | 0.99     |

TABLE 14. RESULTS FOR THE SVM CLASSIFIER AFTER BAGGING

Table 13 results are taken as a reference to build the confusion matrix in Figure 11, for better visualization of the results.
F. RANDOM FOREST CLASSIFIER

The RF algorithm can be applied to classification and regression issues in machine learning. Random forest is an assembly of various decision trees on different subsets of a dataset and finds the mean, which in turn leads to an increase in the accuracy of the dataset. There are two stages to RF. In the first stage, a random forest is generated by the combination of N decision trees. In contrast to other algorithms, Random Forest is less time consuming. Results from applying random forest are,

| Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|---------|------|-----|---------------|------------|----------|
| 0.95             | 0.95    | 0.18 | 0.03| 0.96          | 0.96       | 0.96     |

**TABLE 15. RESULTS OF RANDOM FOREST CLASSIFIER**

Table 14 results are plotted to build the confusion matrix in Figure 13, for better visualization of the results.

After applying bagging results are,

| Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|---------|------|-----|---------------|------------|----------|
| 0.95             | 0.94    | 0.20 | 0.04| 0.95          | 0.95       | 0.95     |

**TABLE 16. RESULTS OF RANDOM FOREST CLASSIFIER WITH BAGGING**

After applying boosting results are,

| Model Test Score | R2 Score | RMSE | MSE | Precision (P) | Recall (R) | F1 Score |
|------------------|---------|------|-----|---------------|------------|----------|
| 0.96             | 0.96    | 0.17 | 0.03| 0.96          | 0.96       | 0.96     |

**TABLE 17. RESULTS OF RANDOM FOREST CLASSIFIER AFTER BOOSTING**

G. GRADIENT BOOSTING CLASSIFIER

A group of ML algorithms combines various weak learning methods to create one robust model. Decision trees are commonly used when performing gradient boosting and are increasingly fetching the famous cause of how effective they are in handling complex data-sets. In hypothesis boosting, we see tall perceptions where the ML algorithm is trained on and take off only the perceptions that the ML strategy effectively classifies, stripping out the other unnecessary perceptions. One modern frail learner made and tried on the data-set, ineffectively categorized, and after the illustrations that were categorized effectively to be retained. In Ada Boost, forecasts
are made with the help of a majority vote; instances are classified as concurring to such a class receives foremost votes from slow learners.

The loss function is trusted by the gradient boosting classifier. Gradient boosting classifiers can use a custom loss function and many standardized loss functions are supported; however, the loss function must be differentiate. Gradient boosting models are prone to over fitting but can perform extremely well on complicated data sets, which can be resolved along with proper strategies.

Gradient boosting, like other boosting approaches, iterative merges weak learners into a single strong learner. It is easiest to understand in the context of least squares regression, where the goal is to “teach” a model $F$ to predict values of the type $\hat{y} = F(x)$ by reducing the mean squared error $\frac{1}{n}(\hat{y}_i - y_i)^2$, where $i$ indexes over some training set of size $n$ of actual values of the output variable $y$:
- $\hat{y}_i$ is the predicted value $F(x_i)$.
- $y_i$ is the observed value.
- $n =$ the number of samples in $y$.

### TABLE 18. RESULTS OF GRADIENT BOOSTING CLASSIFIER AFTER BOOSTING

| Algorithms | Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Fold 6 | Fold 7 | Fold 8 | Fold 9 | Fold 10 |
|------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| LR         | 0.71   | 0.72   | 0.72   | 0.68   | 0.71   | 0.75   | 0.7   | 0.71   | 0.73   |
| DT         | 0.93   | 0.96   | 0.96   | 0.92   | 0.91   | 0.92   | 0.92  | 0.93   | 0.96   |
| Bagging    | 0.93   | 0.96   | 0.95   | 0.92   | 0.91   | 0.95   | 0.92  | 0.95   | 0.96   |
| Boosting   | 0.95   | 0.96   | 0.95   | 0.92   | 0.91   | 0.95   | 0.93  | 0.92   | 0.95   |
| KNN        | 0.95   | 0.92   | 0.93   | 0.94   | 0.93   | 0.87   | 0.93  | 0.91   | 0.95   |
| Bagging    | 0.95   | 0.93   | 0.95   | 0.91   | 0.93   | 0.91   | 0.93  | 0.92   | 0.95   |
| SVM        | 0.98   | 0.98   | 0.98   | 0.97   | 0.98   | 0.96   | 0.97  | 0.98   | 0.98   |
| Bagging    | 0.98   | 0.98   | 0.98   | 0.97   | 0.98   | 0.96   | 0.98  | 0.99   | 0.98   |
| Bagging    | 0.95   | 0.96   | 0.97   | 0.95   | 0.93   | 0.96   | 0.95  | 0.95   | 0.96   |
| Boosting   | 0.93   | 0.94   | 0.96   | 0.95   | 0.95   | 0.91   | 0.95  | 0.92   | 0.93   |
| Bagging    | 0.93   | 0.94   | 0.96   | 0.95   | 0.95   | 0.91   | 0.95  | 0.92   | 0.93   |
| Bagging    | 0.95   | 0.96   | 0.97   | 0.96   | 0.96   | 0.94   | 0.96  | 0.95   | 0.95   |

Random forest”) algorithms to explain how behavioral functions affect them.

### TABLE 19. RESULTS OF ALL ALGORITHMS

#### Figure 15. Confusion Matrix for Random Forest Classifier with Boosting

#### Figure 16. Confusion Matrix for Gradient Boosting Classifier

#### Figure 17. Evaluation Measure

### H. EVALUATION RESULTS

Several features influence the model when forecasting student success. In this study, the characteristics of behavior are called important characteristics that can affect student success. We demonstrated the findings by Classification (“LR, Decision Tree, KNN, and Support Vector Machine, SVM”) algorithms to explain how behavioral functions affect them.

### I. THE RESULTS OF EVALUATION WITH COLLABORATIVE METHODS OF CONVENTIONAL TECHNIQUES OF DATA MINING

In this segment were used to increase the precision. The improved findings are shown in Table 20 through ensemble
methods with three traditional groups ("decision tree, K-nearest neighbour, and support vector machine"). The three classifiers are trained by ensemble trainees and integrate outcomes into a widely held voting process so that students perform best. In the “Decision Tree, K-Nearest Neighbour, and Support Vector Machine”, the boost strategies are better than others, but Decision Tree has the best efficiency, with greater accuracy of Decision Tree by the boost from 0.87 to 0.89 while the accuracy of the result is enhanced from 0.85 to 0.87 and reminiscence results from 0.86 to 0.87.

Table 20 works as a reference for the evaluation measure of the decision tree, KNN, and SVM, as plotted in Figure 17. Following the preparation of the classification model with a 10-fold cross-validation, the validation process begins. The validation method is a very important part of the predictive model construction, which determines the accuracy of the predictive model. The assessment results through the research and validation processes are described in the Table 20 with the usage of classification techniques like "Decision Tree, KNN, and SVM".

Table 21 can be used as a reference for evaluating the test results, validation results, and evaluation measures of the algorithms. As shown in Table 21, an 86% precision was attained through the validation process in the proposed model. Our model performed well compared to the others, with an 82.1% precision. Therefore the outcome of the validation process of the model demonstrated the reliability of the proposed model.

J. COMPARISON ANALYSIS

Table 22. COMPARISON OF PROPOSED WITH EXISTING

V. CONCLUSION AND FUTURE WORK

A person with a high intelligence quotient can basically, have high retention power which enables them to retain more information with them and instantly perform more mental operations in their minds. Great at recalling facts, academic content, key concepts, and other information in general. The rate of acquiring formal knowledge and solving abstract problems is faster. The ability to logically deduce conclusions and detect connections, analogies, and patterns that are more complex at a faster rate. Therefore, it translates into a series of edges regarding the performance of any intellectual task that requires require logical, mathematical, and linguistic thinking, which includes working with scientific technology, philosophy, and academia in general. Another main advantage is the scale of pay. If you can prove the high quality of your intelligence to your recruiters in any form and continue to utilize your skills in the workplace, you will be recognized and rewarded with increased promotion and bonuses.

However, surprisingly, IQ is not sufficient to measure the full extent of the human thinking ability tests alone does not necessarily account for the full range of thinking abilities. In recent years, researchers have added a wider set of skills to the definition of intelligence. In the last 15-20 years, the concept of Emotional Intelligence (EI) have paved the way for describing another set of skills necessary for problem solving. Emotional intelligence refers to one’s ability to recognize and regulate emotions and use them for social awareness and survival in hectic work schedules. We must continue to collect more attributes that have an impact on student intelligence in order to obtain more accurate results for better prediction. Giving students substantial implications in terms of growing intelligence and emotional values in order to increase their academic performance.

For future work it would be interesting to have combination of Intelligence Quotient and Emotional Intelligence together can yield better results and can get clear picture of human intelligence in predicting a student for placements. Combination of EI and IQ towards academic performance highlighted to assess a student performance for placements.

REFERENCES

[1] R. Jiang, V. D. Calhoun, L. Fan, N. Zuo, R. Jung, S. Qi, D. Lin, J. Li, C. Zhuo, M. Song, et al., “Gender differences in connectome-based pre-
dictions of individualized intelligence quotient and sub-domain scores,” Cerebral Cortex, vol. 30, no. 3, pp. 888–900, 2020.

[2] N. A. Goriouanova, D. B. Heyer, R. Wilbers, M. B. Verhoog, M. Giugliano, C. Verbeist, J. Obermayer, A. Kerkhofa, H. Smeekin, M. Verberne, et al., “Large and fast human pyramidal neurons associate with intelligence,” Elife, vol. 7, p. e41714, 2018.

[3] I. Feinkohl, P. Kozma, F. Borchers, S. J. van Montfort, J. Kruppa, G. Wünter, C. Spies, and T. Pischon, “Contribution of iq in young adulthood to the associations of education and occupation with cognitive ability in older age,” BMC geriatrics, vol. 21, no. 1, pp. 1–10, 2021.

[4] A. Gibbons and R. T. Warne, “First publication of subtests in the stanford-binet 5, wais-iv, wise-iv, and wppsi-iv,” Intelligence, vol. 75, pp. 9–18, 2019.

[5] E. T. Reinaldi and R. Hidayat, “Stanford-binet intelligence scale form lm predictive power on academic achievement,” JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia), vol. 10, no. 2, pp. 133–141, 2021.

[6] A. Abramovitch, G. Anholt, S. Raveh-Gottfried, N. Hamo, and J. S. Abramowitz, “Meta-analysis of intelligence quotient (iq) in obsessive-compulsive disorder,” Neuropsychology Review, vol. 28, no. 1, pp. 111–120, 2018.

[7] S. Khan, “A comparative analysis of emotional intelligence and intelligence quotient among saudi business students’ toward academic performance,” International Journal of Engineering Business Management, vol. 11, p. 1847979019880665, 2019.

[8] Z. Li and Z. Qiu, “How does family background affect children’s educational achievement? evidence from contemporary china,” The Journal of Chinese Sociology, vol. 5, no. 1, pp. 1–21, 2018.

[9] A. Giri, M. V. V. Bhagavath, B. Pruthvi, and N. Dubey, “A placement prediction system using k-nearest neighbors classifier,” in 2016 Second International Conference on Cognitive Computing and Information Processing (CCIP), pp. 1–4, IEEE, 2016.

[10] M. I. Al-Twijri and A. Y. Noaman, “A new data mining model adopted for higher institutions,” Procedia Computer Science, vol. 65, pp. 836–844, 2015.

[11] K. Thakur, K. Lal, and V. Kumar, “Ensemble method to predict impact of student intelligent quotient and academic achievement on placement,” in 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), pp. 249–253, IEEE, 2021.

[12] S. Hussain and M. Q. Khan, “Student-performulator: Predicting students’ academic performance at secondary and intermediate level using machine learning,” Annals of Data Science, pp. 1–19, 2021.

[13] P. M. Moreno-Marcos, T.-C. Pong, P. J. Munoz-Merino, and C. D. Kloos, “Analysis of the factors influencing learners’ performance prediction with learning analytics,” IEEE Access, vol. 8, pp. 5264–5282, 2020.

[14] K. S. Double, J. A. McGrane, and T. N. Hopfenbeck, “The impact of peer assessment on academic performance: A meta-analysis of control group studies,” Educational Psychology Review, vol. 32, no. 2, pp. 481–509, 2020.

[15] S. Burden, A. E. Topping, and C. O’Halloran, “Mentor judgements and decision-making in the assessment of student nurse competence in practice: A mixed-methods study,” Journal of Advanced Nursing, vol. 74, no. 5, pp. 1078–1089, 2018.

[16] P. K. Chand, A. S. Kumar, and A. Mittal, “Emotional intelligence and its relationship to employability skills and employer satisfaction with fresh engineering graduates,” International Journal for Quality Research, vol. 13, no. 3, p. 735, 2019.

[17] E. Alyahyan and D. Dişitşegör, “Predicting academic success in higher education: literature review and best practices,” International Journal of Educational Technology in Higher Education, vol. 17, no. 1, pp. 1–21, 2020.