IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

STUDENTS PERFORMANCE ANALYSER USING MACHINE AND DEEP LEARNING

Vaibhav Patil
Computer Science & Engineering
G H Raisoni College of Engineering,Nagpur,India

Janhvi Jaiswal
Computer Science & Engineering
G H Raisoni College of Engineering,Nagpur,India

Harsh Pandey
Computer Science & Engineering
G H Raisoni College of Engineering,Nagpur,India

Aditi Gondane
Computer Science & Engineering
G H Raisoni College of Engineering,Nagpur,India

***Abstract*— Educational institutions generate humongous data globally based on student examinations and cumulative tests, which expand as education awareness concerns the insolvent and solvent nations. Maintaining such data is a big challenge but analysing and predicting a further growth rate of the particular student is essential for the average performing student in academic study. Machine learning algorithms aid the educational systems in reducing human efforts and focusing on the student more because man-made efforts will take time to analyse the student growth projection based on previous data. Making wields the diverse combinational type of coded algorithms, which can analyse the student performance by just giving the input dataset and some extrinsic and intrinsic family, friends, and social life relations will provide the actual situation of particular students who need urgent attention. This model uses different machine learning techniques such as support Gradient Boost, and Deep Learning to build the prediction Models.Root Mean Square Error (RMSE) method is used to measure the accuracy and effectiveness of the model to predict the results. The main frame of concern is to guide and identify the students for scoring better grades in the academic curriculum**.

Keywords*— *Prediction,Performance, Marks, Machine Learning, Deep Learning, Keras.

I. INTRODUCTION

One of the fastest-growing subfields of computer science, machine learning has a broad range of applications. The automatic discovery of significant patterns in data is known as machine learning. It is also commonly used in fields of study like astronomy, bioinformatics, and medicine. The performance of pupils may be predicted using machine learning, which can also be used to spot risks as early as possible. Making it feasible to take the necessary steps to improve their methods would aid students in raising their performance. Based on expected grades, that would make it possible for teaching staff to recognize those students who may require support with the coursework.

This research aims to facilitate students' preparations by forecasting their future academic outcomes based on past performance. With Machine Learning and Deep Learning approaches, students would perform better in class based on expected grades and other important variables, and teachers could identify students who might require more help.

Machine learning techniques can be employed to predict student performance and identify at-risk pupils as early as possible so that the proper steps can be taken to improve their performance. The purpose of utilizing Machine Learning to determine the listing's current situation and further forecast its future performance is to assist the student in avoiding the projected unsatisfactory result.

The environment plays a crucial role in developing one's future in the shape of intrinsic and extrinsic factors concerning social gatherings and daily basis work related pupils and family members.

Human work processing results is a bit tough; however, achievement of perfection comes with time, but humans can make mistakes in calculation unconsciously as a distraction with the respective event can cause accounting and analysis of the toe-toe report of hole stack of grade improvement in 12 dozen students takes plenty time

Following exam summing, result processing requires time because it's done by hand. Consequently, the system is a computerized one. The system is a developing field that is essential to schools and universities, supporting their professors and students to raise grades and performance while keeping in mind other psychological factors that impact their way of life. Using the data, we can evaluate the performance, which will be helpful for both mentors and pupils.

II. RELATED WORK

(Midhun Mohan M G et al. 2015) The typical forecast of the student's performance over a vast amount of data often uses two different approaches. These methods include predictive analytics and learning analytics. Learning Analytics focuses mainly on the information series and pre-processing information phase, where the necessary information is gathered from the schools to build the prediction model. They use a MySQL server to store an enormous amount of data. Data cleansing, data transformation, and other information pre-processing steps follow.[1]

(Madhav S. Vyas and Reshma Gulwaniet et al. 2017) They employ a selection tree model to forecast pupils' total academic performance. Data needed for creating the blueprint is gathered, and records pre-processing is done where continuous values are changed to discrete values, and null values are ignored. Afterward, a decision tree prediction model is built by applying the CART algorithm to the records. The students who perform poorly overall are dismissed. [2]

(Alana M. de Morais, Joseana M. F. R. Araújo, and Evandro B. Costa et al. 2014) The two methods that draw the most excellent attention are clustering and regression. Regression is used to make predictions, while clustering gives data pre-processing. Even if the data classification is done and the regression model is utilized to get the forecasts, comparable data are grouped in clustering.[3]

(PauziahMohdArsad, NorlidaBuniyamin, and Jamalullail Ab Manan et al. 2013) The concept of an ANN network is used to forecast pupils' academic performance. The resulting criterion is grade point average (CGPA). The electrical department of Malaysia's Technology MARA University provided the information needed for the endeavor. Pearson's correlation coefficient R and Mean Square Error were used to measure the models' outputs (MSE). The learning of the validated subjects has significantly impacted the CGPA.[4]

(B. K. Baradwaj, S. Pal, et al. 2011) Researchers have been attempting to solve the problem of student result analysis and prediction for the past few years. A classification method based on decision trees to forecast students' final exam scores is provided. The authors claimed that concealed data in educational databases might be essential for improving students' performance.[5]

(N. A. Gorikhan et al. 2016) The application of various classification techniques for vocational institutional analysis that aid teachers in working with struggling students to improve their performance and assert that decision trees are the most accurate prediction model for institutions' students' research.[6]

(Asaad Masood et al. 2019) Machine learning-based predictive models for assessing student performance (Asaad Masood - 2019), Decision Trees, Logistic Model Trees (LMT), and Association Rules Mining are some algorithms employed. 1021 Records from the test database were used in the data set. Benefits: Factors influencing student performance are analyzed, and timely, wise decisions are made. The given dataset is relationally old, which is a drawback.[7]

(Mehil B Shah et al. 2019) Machine Learning-based Student Performance Assessment and Prediction System [(Mehil Shah-2019)9] Support Vector Clustering (SVC), Decision Tree, and K Nearest Neighbor (KNN) Random Forest Model are the algorithms that were used. Portuguese data collection was used for binary/five-level classification and regression applications with 33 different attributes. Advantages: wide range of non-academic factors, such as commute time, the job of the father or mother, Etc. Disadvantages: reduced precision neural network, the accuracy rate of all employed algorithms is relatively low.[8]

On the platform, four methods are used to accurately forecast student performance and identify students at risk: (i) prediction of academic achievement; (ii) identification of students at risk; and (iii) assessment of learning challenges. (Alasdair Bruceet al. 2019) By creating and putting into place quick and regular intervention procedures, accurate early prediction of student dropout contributes to solving the underlying issue. This section reviews related research based on datasets, characteristics employed in ML approaches, and the results of the research in order to provide a detailed description of dropout prediction using machine learning techniques. In order to help educators and organizations that are interested in learning about ways to improve students' academic performance, it is necessary to examine their performance in the classroom. By combining individual variables including the student's age, demographic makeup, personal review ideology, and family background, their review performed the study of prior results using a variety of machine learning-based algorithms. They came to the conclusion that linear regression, supervised learning, and deep learning were the three most important models for the comparative performance study. Also highlighted was the potential of ML for early dropout prediction in students. After that, the dataset was split into training and testing sets, respectively. In situations when the Keras deep learning framework achieves high average scores or accuracy, Gradient Boosting Tree (GBT), SVM, and RF classifiers are used.[9]

(Diego Didonaet al. 2015) The correctness of these solutions ultimately depends on the accuracy of the Deep Learning technique used, as Machine Learning is simply used to define the workload. Conversely, the ensemble techniques we propose in our study rely on machine learning to correct an analytical model's errors, and as a result, their accuracy can be improved over time as new sample data is gathered from the system being modeled. The prediction curve of the system as a whole is then created by combining the sub-models in accordance with a formula. We identify applications for this technique in the domain of performance modeling for Deep Learning and Machine Learning applications.[10]

(Sher Afzal Khanet al. 2021) The authors suggested a machine learning technique while working on forecasting student performance by predicting student grades in terms of pass/fail. However, adopting automatic feature representation techniques seen in Deep Learning models can further boost the system's performance. The study aims to create a collaborative filtering method that uses academic data to forecast students' success. The experimental results demonstrate that the approach outperforms the standard support vector machine classifier. In their research on evaluating student performance, they developed a method based on low-range matrix factorization and a dispersed linear model that uses previous academic grade data from students as an input to estimate a student's performance.[11]

(Shah Hussein, Muhammad Qasim Khan, et al. 2021) In a study, the author used a novel technique to assess students' performance by using past academic data from their students as input. The study used a distributed linear model and low-range matrix factorization. About 12 years' worth of academic data from the University were included in the data set that was reviewed. The proposed system has increased the precision of grade prediction. The study's author proposed a novel technique for analyzing educational data that makes use of a recommendation system that is specifically designed to forecast performance and validated by comparison to existing cutting-edge regression models, such as logistic and linear regression. Utilizing a collaborative filtering process, to assess students at the beginning of a semester and forecast their academic success. The student's learning is reflected in the authorized courses that are selected. The information system gathered historical data to detect common traits among university students. According to the research analysis, this method works well as an SVM classifier.[12]

(Timothy Andersonet al. 2017) Family influences include the parents' educational backgrounds, their capacity to help their children with their education, and their capacity to foster an environment that is conducive to learning. The study's findings showed that while parent employment plays a significant effect in predicting grades, the type of school a student attends has little bearing on their performance. The authors discovered that a student's past academic success and parents' educational backgrounds are the two factors that matter most in forecasting that student's future academic success. Studies sometimes look at how parents' educational background and financial status affect their children's academic achievement.[13]

(HAYAT SAHLAOUI et al. 2021) It uses a number of machine-learning techniques to forecast students' grades across our data collection. We contrast the outcomes of the machine learning algorithms with the "simple average" method of prediction most frequently used by professors and students. The sections that follow provide a quick review of machine learning techniques, the outcomes of utilizing these techniques to forecast students' marks about a "basic average" prediction as a reference point, and a discussion of pedagogical applications of this method of grade prediction.[14]

III. METHODOLOGY

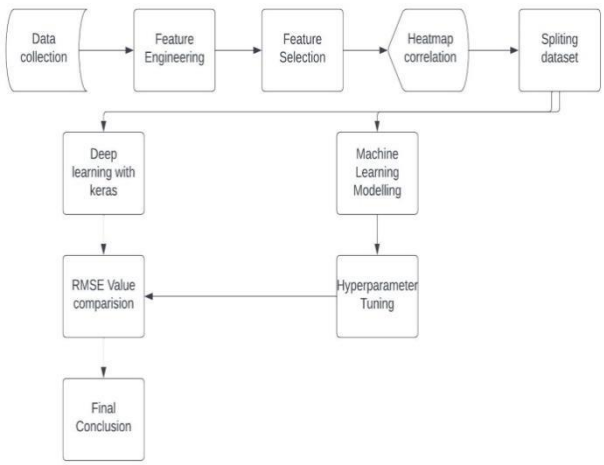

Fig 1. Proposed Methodology

The complete methodology consists of the following steps:

(A) *Dataset selection - Input data*

We have two schools in the dataset, including the student, parents, and personal information inside the file. Every critical information is intervened in the dataset so that model can take appropriate decisions for students despite teaching faculty.

(B)*Attribute Information School*

The dataset consists of 33 diverse columns and 395 rows of entries in terms of information concerning a particular student. School and educational institution is represented as the school with the option of Mounsinho Silveira(MP) and Gabriel Pereria(GP). Gender and sex are defined as the sex in the column, whether female (F) or male(M). The age column shows the numeric number range (15-23) of a person's age. The address represents whether a student lives in urban (U) or rural(R). Family size is classified based on less than 3(LT3) factors and greater than 3(GT3). Parental status signifies living together(T) or living apart (A). Mothers' education(Medu) is scaled on a 0 to 4 points basis, whereas 0-nope, 1-elementary education, 2- 5th to 9th grade, 3-secondary school, 4-Further education. Father's education(Fedu) is scaled on a 0 to 4 points basis, whereas 0-nope, 1-elementary education, 2- 5th to 9th grade, 3-secondary school, 4-Further education. Mother job(Mjob) can be categorized as a teacher, health care service, civil service (administrative or police), or at home(other). Father job(Fjob) can be categorized as a teacher, health care service, civil service (administrative or police), or at home(other). The guardian is also referred to as the student's mother, father, or further guardian justification for choosing this institution. Travel time is referred to as home-to-school travel time. Study time is a week of study time spent by a student in the research of a respected course or academic undergo specific curriculum books. Failure is mentioned as the number of times a student struggle to pass a specific grade in a particular year in sequence. The school's column shows the extra educational support for the future scope of the job or business purpose, which is shown in the dataset as yes and no. famsup column tells that family can assist the student's future struggle. The extracurricular column shows the different hobbies other than academic activity as yes or no in records. The paid column shows us that students take extra classes other than school or college. The higher column details students are willing to study in higher study after leaving school. Internet column has simple if students have access to them (yes or no). Romantic has the information as students are involved in relations with the opposite gender in their teenage years. famrel has the quality check of connection based on a link with family members on a Numeric scale: 1 for extremely poor to 5 for outstanding. Free time is free time for a student away from study or academics with numbers: extremely low = 1 to very high = 5. Walc represent the alcohol intake of a student on a scale of 0 to 5 as 0 -nope to 5- extreme drinker. The dalc daily alcohol intake during a day is under-noted in this column on a scale of 0 to 5 as 0 -nope to 5- extreme drinker. The absences column here refers to absences from school, skipping classes, or leaving school.

(C)*Exploratory Data Analysis*

We are importing specific libraries in the project, such as panda, NumPy,matlotlib, and seaborn, which aids the visual of the graph generation in the project. The dataset consists of 33 diverse columns and 395 rows of entries in terms of information concerning a particular student. Mean, median, and mode calculations to analyze the data. Comparison of attribute features with the grades concerning 30 attributes to check the result of the student in specific G1, G2, and G3

inint format numbers obtained by the student, such as school student performance versus Mounsinho Silveira(MP) and Gabriel Pereria(GP). Comparison of age and sex, mother and father education with G3, travel, and study with the G3 score of the student and other columns in the dataset. Probability distribution of dataset oblation marks to view the students' grades at a particular age in the three consecutive grade terms to understand the student's performance. Then, a comparison between G1 vs G3 and G2 vs G3 is taken to identify the improvement in the final term concerning the initial ones. The result from knowledge to interpret the linear graph shows the variation in performance.

(D)*Feature Engineering*

It plays a crucial role in the metamorphosis of data provided in the data set to operate the machine-learning language. The fundamental work of feature engineering is to convert the impeccable format which suits the model to work accurately.

(E)*Feature Selection*

Feature selection will shape the model in an accurate way which can make the model impeccably accurate so that negative terms can not lead to irrelevant errors and flaw developments of the system. Feature selection makes the system non-vulnerable to a flaw that will damage the model in the development phase.

(F)*Correlation Heat map*

Heat map correlation is used to see how every element is essential for the student to improve or impact his or her academic life in school. Heatmap correlation is generally availing when multiple variable columns are there in the dataset and to know how crucial column is to each other.

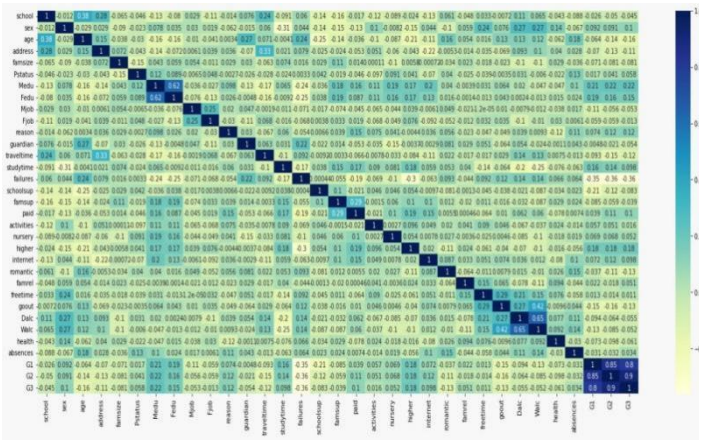

Fig 2 Heat correlation Map of dataset

(G)*Splitting data*

Splitting data into train and testing plays a vital role in predicting the accuracy of machine learning algorithms, which proceed in the model. Splitting data into train and test have a division of 3:7(test set: training set) as its fast and efficient procedure.

(H) *Calculation of Root Mean Square Error*

One of the methods most frequently used to assess the accuracy of forecasts is a root mean square error, also known as root indicate square deviation. It displays the Euclidean distance between measured actual values and expectations.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \|y(i) - \hat{y}(i)\|^2}{N}},$$

Formula of RMSE calculation

(I) *Machine Learning Algorithms used*

This model uses seven algorithms to predict the accomplishment of the best and most efficient accuracy result measured by a root mean square error minimum value.

a. Linear Regression-

The linear Regression method has two independent and dependent variables, which are used to show whether the relationship may or may not necessarily depend on each other. However, there must be a critical relationship between the two variables; then, a scatter plot is used to imply the strength between the variables, but if not, the graph will be stagnant.

b. LASSO (Least Absolute Shrinkage and Selection Operator)-

The regression coefficients are shrunk by the LASSO approach, which causes some of them to become zero. Every non-zero value is chosen to be utilized in the model during the feature selection phase, which follows shrinkage. This technique's reduction of common prediction errors in statistical models is significant.

c. Ridge algorithm--

Regularization methodology, or, to put it simply, a Linear Regression variant, is what the Ridge Regression is. One regularisation strategy where the data exhibits multicollinearity is this one. The variance is high in this multicollinearity, the least squares are unbiased, and the predicted value differs from the actual value. There is also an error term in this equation.

d. ElasticNet--

The two most widely used regularised linear regression methods, Lasso and Ridge, are combined to create an elastic net. L2 is the penalty used by Ridge, while L1 is the penalty used by Lasso. You can use either of these models with elastic net because it uses both the L2 and L1 penalties.

e. Gradient Boosting Regressor--

This estimator allows for the optimization of any differentiable loss function and constructs an additive model in a forward stage-wise manner. A regression tree is fitted on each level's negative Gradient of the provided loss function. Comparison of accuracy between the six algorithms Gradient Boost with a minimum value of 1.66 RMSE.

(J).HyperparameterterTuning

Hyperparameters in Machine learning are those parameters that are explicitly defined by the user to control the learning process. These hyperparameters are used to improve the learning of the model, and their values are set before starting the learning process of the model. Algorithms used in machine learning knowledge. These have been adjusted to improve the model's performance. The goal of hyperparameter tuning is to identify such parameters where the model performs best or where the model performs best and the error rate is lowest.

A mathematical model containing several parameters that must be learned from the data is referred to as a machine learning model. We can fit the model parameters by using existing data to train a model. Hyperparameters, on the other hand, are different kinds of parameters.

After Hypertuning the result, instead of getting a more efficient result, it got worse and turned out to be an RMSE value of 2.098.

(K)Deep Learning

Artificial intelligence's deep learning field focuses on simulating the functioning of the human brain to solve challenging issues. In deep learning, we employ neural networks that operate numerous operators placed in nodes to assist in decomposing the problem into small ones that are each tackled separately. However, it can be pretty challenging to implement Keras, a deep learning framework, that handles this issue. To implement neural networks, Google created the high-level Keras deep learning API. It is a modified implementation ion of neural networks and is written in Python. Additionally, multiple neural network backend computation is supported.

Deep Learning is used to make the result more accurate than the best optimism algorithm in the machine learning model. After using deep learning in this model RMSE value comes out to be 0.58, which is the lowest until the algorithm is used.

IV. RESULT

1. Compare Age and Sex with G3

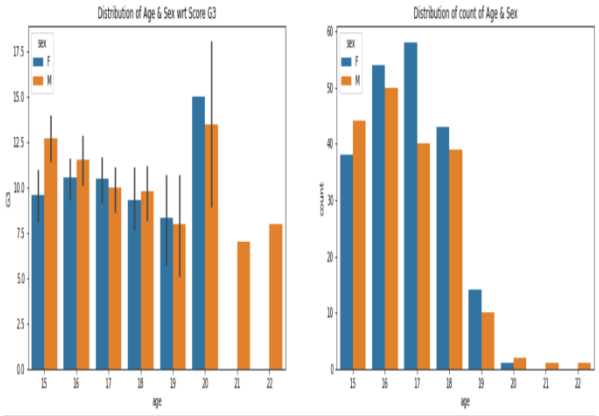

Fig 3 Graph demonstrating distribution in age, sex and G3

These two graphs illustrate the comparison between age & sex with G3 and age and sex with respective order in the graph. The first graph has the volatile scoring in the minuscule difference between opposite gender male and female in the age gamet 16-19, but 15 and 20 have measurable differences. As shown in the graph, the astonishing result as 21 to 22 age only males perform academic activities. The second graph shows that females are indulging in academics concerning age and sex but suddenly, a significant drop in education for both genders after 20.

2. Compare family (father and mother) education with G3

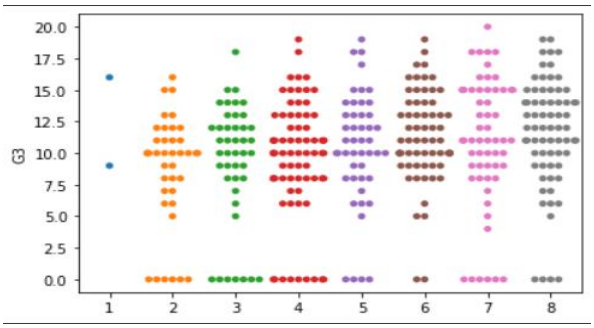

Fig 4 Comparison of educational background with grades

The above graph represents the student performance of students' grade basis on the attribute of educational background, which is essential for understanding the student's struggles in the academic life cycle. Performing exams and experiences have a valuable impact on their career later, so it turns out to be the most crucial parameter in taking grades up for the student's grade.

3. Comparing other features

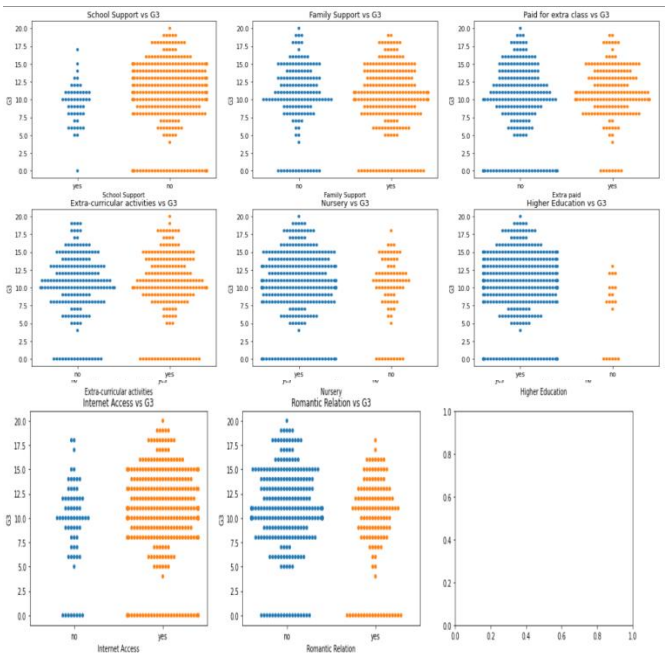

Fig 5 Comparison of other features with grades

Multiple graphs illustrate academic life's negative and positive impact on the student concerning parameters with exact values that impose change, whether an increase or decline in grade.

Factors such as not having school support or having a relationship in school have declined graph in period time involvement in such circumstances, but some factors like internet access, extra classes, and went to nursey have a positive effect on academic life to success in grade compared to other which have lacked in this situation.

4. RMSE value of applied Machine Learning Algorithms- Root mean square value (RMSE) is a particular parameter to measure the accuracy of the machine algorithms used in the marks prediction model to achieve an effective result than labor work by the human force. Lower the value of the RMSE for operation performed algorithms is directly proportional to the skyrocket accuracy of liability for the model to operate on correct values.

Algorithm Name	RMSE Value
Linear Regression	1.92
Ridge Regression	1.92
LASSO	1.91
ElasticNet	1.91
SVM	1.97
RandomForest	1.72
GradientBoost	1.66
Deep Learning	0.58

Numerous algorithms are applied to train the model and get impeccable accuracy levels, including the above mentioned list in the machine learning prediction model and Gradient Boost, which is the best RMSE value. Hypertuning of the best algorithm to achieve a greater accuracy level, but it got a worse accuracy score than a worse machine learning model, SVM(1.97). The accuracy level score drops by 0.03, and the results appear to be 2.08 after the hypertuning gradient boost. To boost the level of accuracy of the model or achieve the apex accuracy, Deep Learning is introduced to mitigate the problem statement. With the deep learning method, the accuracy level achieved was 0.58, which is an astonishing number in terms of accuracy.

5. Comparing G1, G2 with G3

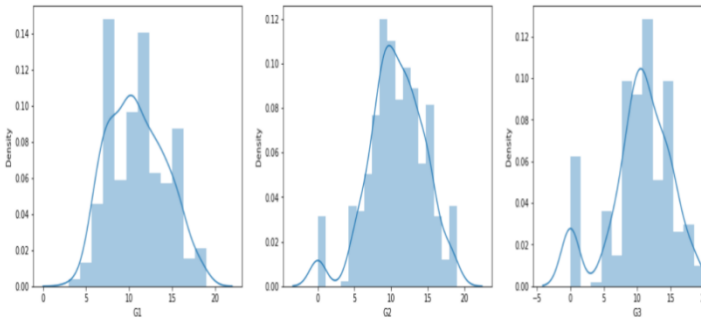

Fig 6 Probability distribution of G1,G2, G3

Three graphs on the line side by side are the probability distribution of marks over the particular grade set having grades with the highest score in the particular figure.

V.CONCLUSION

This machine learning model predicts student growth or plunge of academics with respective to life-leading factors, which can be intensive and extensive with basic catering. The increase or extension of more and more data in the dataset will bring the apex of accuracy level in time upgrade. Machine learning will eventually improve as the growth of constant data sources increases with diverse types of stream colleges.

Algorithms are designed so that excellent cut-off unwanting waste attributes will vanish, but the engagement of benevolent features will take the initiative and add value to improve the model. Insolvent or poor students can take advance of it and make a change to solvent or rich grades with an increase in certain essential factors, which can lead student future in bright life.

Gradient boost produces the most satisfying results. However, after hyper-tuning, the algorithm that was supposed to improve was worse. Deep learning then achieved accuracy levels that were skyrocketing compared to gradient boost.

VI. FUTURE SCOPE

Eventually, data gets updated on quick action across the globe, which helps our model increase the prediction results using algorithms. Deep learning is an exception to get expertise in analyzing results as algorithms used previously like gradient boost, linear regression, ridge regression, and SVR. Extrinsic and intrinsic factors also matter a lot in academic performance with family and surrounding supporting mates, which impact social life in particular

REFERENCES

[1] Midhun Mohan M G, Siju K Augustin and Dr. Kumari Roshni V S "A Big Data Approach for Classification and Prediction of Student Result Using Map Reduce", 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS) | 10-12 December 2015

[2] Madhav S. Vyas and Reshma Gulwani. "Predicting Student's Performance using CART approach in Data Science", International Conference on Electronics, Communication and Aerospace Technology ICECA 2017

[3] Alana M. de Morais, Joseana M. F. R. Araújo and Evandro B. Costa. "Monitoring Student Performance Using Data Clustering and Predictive Modelling", 978-1-4799-3922-0/14/$31.00 ©2014 IEEE

[4] PauziahMohdArsad, NorlidaBuniyamin and Jamalullail Ab Manan. "A Neural Network Students' Performance Prediction Model (NNSPPM)" IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), 26-27 November 2013

[5] B. K. Baradwaj, S. Pal, "Mining Educational Data to Analyze Students Performance", International Journal of Advanced Computer Science and Applications, 2(6), 2011. https://doi.org/doi:10.14569/IJACSA.2011.020609

[6] N. A. Gorikhan, "A study on Implementation of classification techniques to predict students' results for Institutional Analysis" MSc Thesis, Faculty of Engineering & IT, The British University in Dubai, 2016

[7] Asaad Masood-" Predictive model for predicting students' performance using Machine Learning", 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)

[8] Mehil B Shah- "Student Performance Assessment and Prediction System using Machine Learning", 2019 4th International Conference on Information Systems and Computer Networks (ISCON).

[9] Al-Samarraie H, Teng BK, Alzahrani BK, Alalwan N (2018) E-learning continuance satisfaction in higher education: a unified perspective from instructors and students. Stud High Educ 43(11):2003–2019

[10] Diego Didona1, Francesco Quaglia2, Paolo Romano1, Ennio Torre2 1 INESC-ID / Instituto Superior Técnico, Universidad de Lisboa 2 Sapienza, Università di Roma "Enhancing Performance Prediction Robustness by Combining Analytical Modeling and Machine Learning "2015

[11] Bashir Khan Yousafzai 1, Sher Afzal Khan 1, Taj Rahman 2, Inayat Khan 3, InamUllah 4,*, "Ateeq Ur Rehman 5, Mohammed Baz 6, Habib Hamam 7 and Omar Cheikhrouhou"2021

[12] Shah Hussein, Muhammad Qasim Khan" Student-Per formulator: Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning" 2021

[13] Timothy Anderson, Stanford University Randy Anderson, California State University, Fresno" APPLICATIONS OF MACHINE LEARNING TO STUDENT GRADE PREDICTION IN QUANTITATIVE BUSINESS COURSES"2017

[14] HAYAT SAHLAOUI1, EL ARBI ABDELLAOUI ALAOUI 2, ANAND NAYYAR 3,4, SAID AGOUJIL1, AND MUSTAFA MUSA JABER 5,6" Predicting and Interpreting Student Performance Using Ensemble Models and Shapley Additive Explanations"2021