INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

**ISSN: 2454-132X**
**Impact factor: 6.078**
***(Volume 6, Issue 3)***
Available online at: www.ijariit.com

# Student placement possibility prediction using Naive Bayes algorithm

*Brijmohan Lal Sahu*
*ipowerbrij007@gmail.com*
*Disha Institute of Management and Technology, Raipur, Chhattisgarh*

*Dr. Anil Tiwari*
*anil.tiwari@dishamail.com*
*Disha Institute of Management and Technology, Raipur, Chhattisgarh*

## ABSTRACT

***The Aim of my project is to analyze how we can use available student data to predict the standard of new students and the possibility of their placement. It's a student assessment on the basis of his academic skills, Innovation, Research and Development capabilities. Doing this can help students as well as academics to boost students' standards. Also it helps us to design our course and extra training on the basis of it for maximum benefits to students. For this I conducted a small survey using Google form survey and collected information about students of different colleges and universities around Chhattisgarh in the Computer science and Engineering Department and Information Technology Department. Creating a dataset for this project is not easy because motivating students to share their information is a challenging task. For predicting the result of student placement possibility, we use the dataset prepared from the survey and different types of Naïve Bayes Algorithm (Classifie), which is a Supervised Machine Learning Algorithm. Naïve Bayes Classifier is one of the most effective and simplest algorithms to implement for supervised labeled dataset.***

***Keywords***— *Student Performance, Machine Learning, Naive Bayes Classifier, Gaussian, Multinomial, Bernoulli.*

## 1. INTRODUCTION

Artificial Intelligence and Machine learning research works are running in every dimension possible. A lot of data and powerful processors are making it even easier to refine and prepare it for feeding to modern algorithms. The primary goal of education system is Skill, Knowledge and getting a good Job. To get good job you must be good in your academics and prepare for opportunity before it knocks your door. So, that you can prepare yourself and grab the opportunity. This project is focused on the same requirement of predicting the possibility of placement of a student, so that he or she can prepare himself for the upcoming opportunities. Everything we do, think or feel is data. That can be used to predict results, finding solutions, preventing something that might be harmful or saving a life. Predictive models are quasi experimental structures used to determine the future patterns in data. These meaningful data patterns form the building block of any decision support system. Researchers all over the world have built many prediction models for major industries [18][19]. Some of them are implemented using SVM, feature selection algorithms, Neural Network, Fuzzy System and Deep Learning [23].

Data act as a fuel to the powerful processors equipped with advanced AI algorithm that can extract remarkable conclusions from data. Prediction is a powerful data analysis task used to determine the future behavior of a user or a system [20]. The task of prediction has aided the decision makers to take the right decision at the right time. Predictive activities have gained wide popularity owing to the high availability of surplus data. Computers generate nearly 2.5 quintillion bytes of data every year.

Predicting student's performance is a significant research work in the educational sectors. Few Years back, students waited till the end of the semester to know their GPA now we can even predict that too [10]. With the advent of machine learning, we can predict all semester's grades easily and in my project work we are using model to predict the likelihood of a student being placed or not. Research works in the educational sector has increased sharply [25]. This sharp increase may be due to the high availability of data in the educational domain. In my work I am trying to comprehend a few literary works on academic performance prediction of engineering students with the focus on placement predictions. Meaningful interpretations have been made and inferences are presented at the end of this paper.

## 2. LITERATURE REVIEW

**Baradwaj and Pal et al.** [1] They use decision Tree algorithm to classify students to predict students' division on dataset obtained from VBS Purvanchal University, Jaunpur (Uttar Pradesh). Not only previous semester result has been chosen as the attributes but the lab work and the seminar performance also have contributed to the findings. Their research also able to identify those students which needed special attention in order to reduce fail ratio.

 

**Borkar and Rajeshwari et al.** [2] On 60 datasets of students from the Pimpri Chinchwad College of Engineering, Pune University, run a research using association method. Their research used the students' attendance, unit test, graduation percentage and the assignment as the attributes to predict students' performance in university end semester exam.

**Anal Acharya,Devadatta Sinha et al**. [3] A student data set with 309 records and 14 features collected by a survey from various graduation level students majoring in Computer Science under University of Calcutta. It shows how to select the best features to predict the results. The best results are obtained by Feature Selection algorithm with 8 features.

**E. Osmanbegovic, M. Suljic, and H. Agic et al.** [4] To predict the performance of student in secondary school at Tuzla was conducted. The study is performed on data set with 19 features using Gain Ratio (GR) feature selection algorithm. The results with Random Forest classification (RF) algorithm reveal best results in terms of prediction accuracy.

**Karthika & Sairam et al.** [5] Have done a research paper on method to categorize the educational qualification using the Naïve Bayes Classification algorithm. It is found that Naïve Bayes Classification algorithm performs well when the attributes are non-numerical.

**Strecht, Soares, et al.** [6] The goal of their research it to predicting students' final grade. They evaluated various popular regression algorithms, i.e., Ordinary Least Squares, SVM, CART, kNN, RF and AdaBoost R2. The experiments were carried out using administrate data from the university's Student Information System (SIS) of Porto, concerning approximately 700 courses. The algorithms with best results overall were SVM, RF and AdaBoost R2.

**Kaur et al.** [7], The prediction of whether a student should be considered as qualified or not. The authors experimented with five classification algorithms and four attribute evaluation methods, using a sample dataset of 152 regular high school students. The Multilayer Perceptron (MLP) was the best performing classifier among all other methods.

**M. Mayilvaganan, D. Kalpanadevi et al** [8] used C4.5 algorithm and Naive Bayes' Classifier for Predicting the Cognitive skill of students. The cognitive skill of a student is tested during the online test which is associated with the intelligent test score (IQ) and health, more specifically morbidity (physical and mental). The cognitive skills of the students are classified using naive Bayesian classifier and C4.5 algorithm.

The core idea, goal and summary of most of the papers covered for this research are listed above. A short explanation about the research papers taken as guiding tool for this project. Many papers covered from different researchers from across globe and different field. They have explored different aspects of student performance evaluation and predictive model development. Their research involved different algorithms, data sets and features. Work performed by them is truly remarkable and it works as a support for my project. The model proposed in this thesis is based on Naïve Bayes algorithm and three types of it.

## 3. RESEARCH QUESTION AND OBJECTIVES OF THE PROJECT

Significance of everything starts with just a very small word why?
Many researches are done on student, about GPA [11], result in exam, University exam, pass or fail status. They all uses the data, and only the current data. This is the beginning of the question.

- Is it possible to use past data to predict future?
- If we can predict his status in university exam then, can we predict his performance in interview or placement exams?

This project is the result of such questions. A quest to predict future with data and algorithms.

Every research work, be it descriptive or analytical, applied or fundamental, needs a set of research objective to work on. The objective of this project is:

- To identify the common methodology adopted in building a predictive model among some ML algorithms.
- To check how Naïve Bayes algorithm performs in predicting student performance or
- To find the important techniques used in predicting the performance of students.
- To find the commonly used attributes for building the prediction models.
- To perform the frequently used performance metrics.

## 4. IMPLEMENTATION

Implementation of Naïve Bayes is very simple using python Scikit-learn library. Before applying Naïve Bayes we must have prepared data set. Then we load the data set into data frames. Data frames are just a multidimensional mathematical array. Used for holding values of each attributes. As Naïve Bayes is supervised machine learning algorithm it requires labeled data, but processing labels is hard. To make it simple the labels in data frames were encoded using some other encoding algorithms. Like one hot encoding. Processing on encoded data makes it faster and simpler. Here in this project categorical to numerical encoding is used. After that we get a matrix called data matrix and it's ready to feed into ML models.

To make our task simpler the data set is divided into two parts training data set and testing data set known as train-test split. A general ratio for this is 70-30 means 70% of data for training the model and 30% for testing the trained model. To avoid the problem of ***over-fitting***, means testing on the trained data. It's important to keep training and testing data different otherwise it's just a memorizing system if the same data is used for training and testing. The trained model is then tested against the unseen data (Testing data set) known as model evaluation. The performance of model is calculated on the basis of its response. Means how

 

many times it is correct, how many times it failed and predicted false for true and true for false. Other parameters accuracy, precision and recalls were calculated on basis of it. The details of performance evaluation parameters are explained in next chapter.

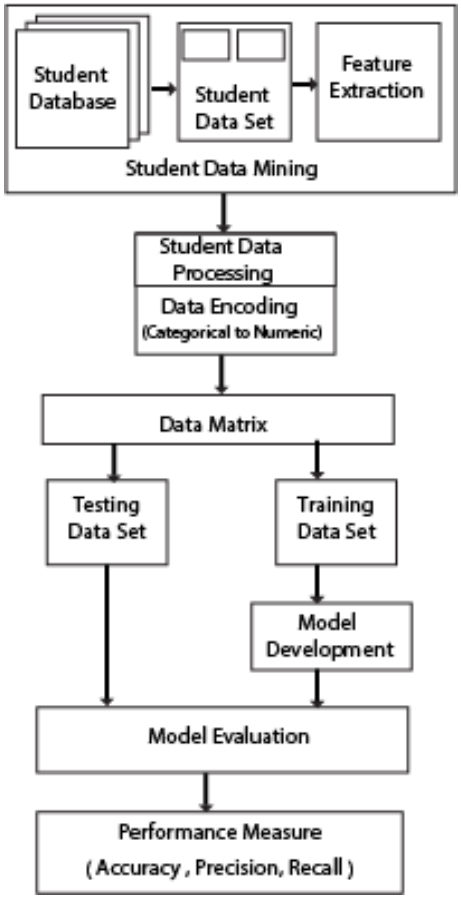


**Fig. 1: Machine Learning Model**

The above figure illustrates machine learning model and steps involved in it, along with data splitting, training, testing and model evaluation for performance measure. Naïve Bayes algorithm is Supervised learning algorithm that can only be used for classification of labeled data. That's why the given student data set is labeled dataset. The performance accuracy varies with amount of data along with computer resources. For predicting student placement possibilities 3 types of Naïve Bayes algorithm is used and collected 32 student (skill set/ result / features) property helps in it to make it even closer to the actual possibility. Some features it includes are belongs to major fields:

1. Past performance ($10^{th}$ and $12^{th}$)
2. Current Performance (engineering)
3. Skills (Programming Languages)
4. Subjective Knowledge
5. Projects and Innovation, Internship
6. Extra certification
7. Language and Communication.

The parameters used for measuring the performance are:

1. Confusion Matrix
2. Accuracy Score.
3. Precision.
4. Recall also known as Sensitivity.
5. F1-Score.
6. Support.

All the coding part is done in python using sklearn, numpy, pandas libraries and executed on spyder, jupyter notebook under anaconda package distribution environment. Many useful relations or we can say the interconnecting between a student past, present and future status is drawn. All these information and knowledge generated during this research project not only help school, college, industry but also help student and his family for proper planning of future. We can analyze the current scenario not in just our department but also around us. Can guide students about backlogs using statistics based on their $10^{th}$ class and $12^{th}$ percentage.

The figure 2 below shows relation between two features of student as x-axis and y-axis along with that it representing placement as the point hue. Green for not placed and Blue for placed.

 

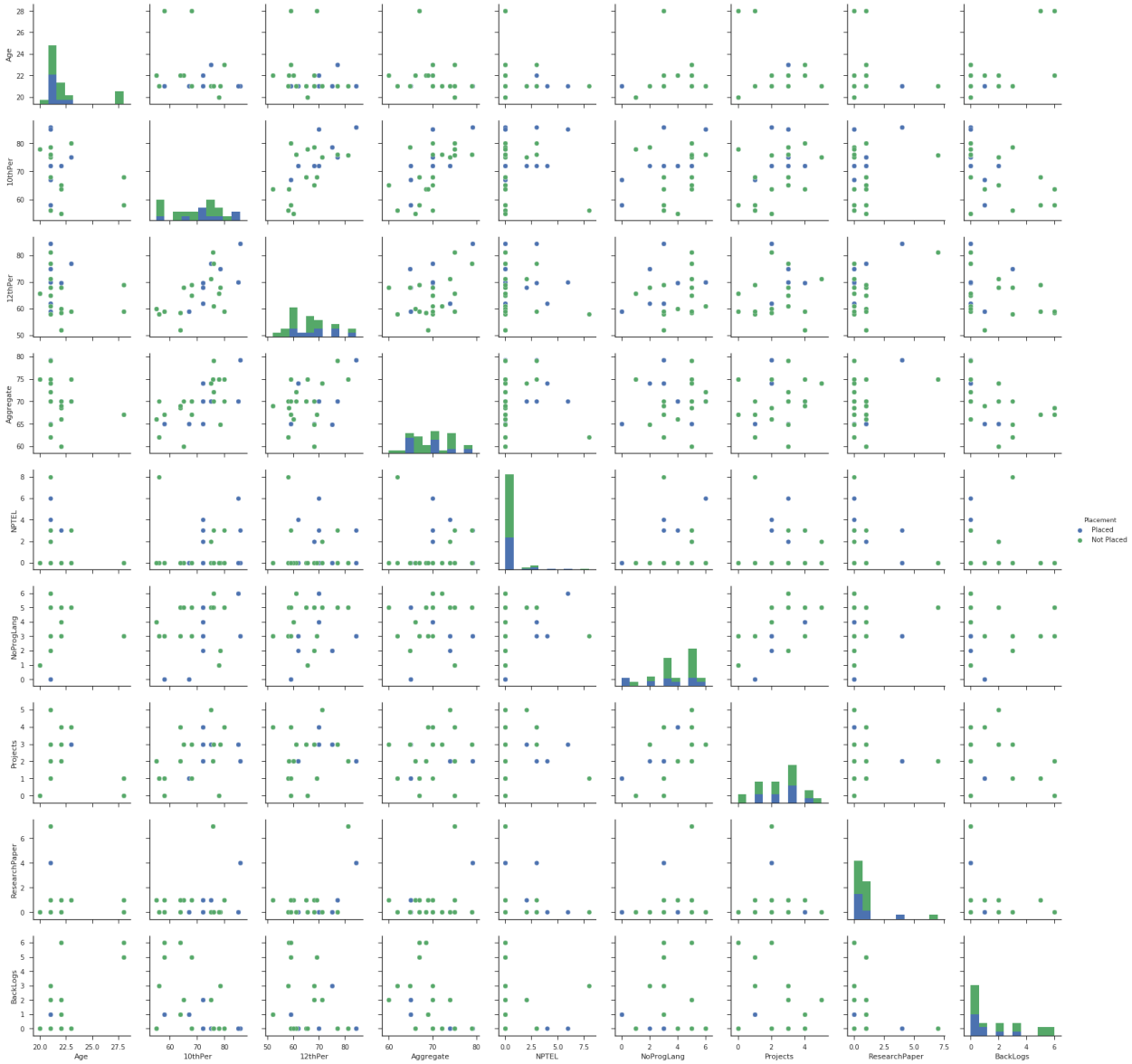


**Fig. 2: Relation with other features and placement**

When a student join in engineering, from Hindi medium school, the biggest challenge for hin is English, a language barrier.

## 5. RESULT

As it's a learning model and results varies. So, here result of all three Naïve Bayes Algorithm is generated by system for running multiple times. System generated accuracy, precision, recall, f1-score and support for every iteration for each type of Naïve Bayes model. Here, just for ease of representation only included 15 testing iterations.

======== Report Test Count :15
Confusion Matrix and Performance metrics
precision,recall, f1-score, support
==========
GaussianNB Accuracy : 0.633333333333

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.63 | 1.00 | 0.78 | 19 |
| Not Placed | 0.00 | 0.00 | 0.00 | 11 |
| avg / total | 0.40 | 0.63 | 0.49 | 30 |

Confusion Matrix GaussianNB:
[[19 0]
[11 0]]
---
True Positive: 19
True Negative: 0
False Positive: 11

 

False Negative: 0
==========
MultinomialNB Accuracy : 0.7

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.75 | 0.79 | 0.77 | 19 |
| Not Placed | 0.60 | 0.55 | 0.57 | 11 |
| avg / total | 0.70 | 0.70 | 0.70 | 30 |

Confusion Matrix MultinomialNB:
[[15 4]
[ 5 6]]
---
True Positive: 15
True Negative: 4
False Positive: 5
False Negative: 6
==========
BernoulliNB Accuracy : 0.733333333333

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.79 | 0.79 | 0.79 | 19 |
| Not Placed | 0.64 | 0.64 | 0.64 | 11 |
| avg / total | 0.73 | 0.73 | 0.73 | 30 |

Confusion Matrix BernoulliNB:
[[15 4]
[ 4 7]]
---
True Positive: 15
True Negative: 4
False Positive: 4
False Negative: 7
==========
GaussianNB Accuracy : 0.666666666667

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.68 | 0.89 | 0.77 | 19 |
| Not Placed | 0.60 | 0.27 | 0.37 | 11 |
| avg / total | 0.65 | 0.67 | 0.63 | 30 |

Confusion Matrix GaussianNB:
[[17 2]
[ 8 3]]
---
True Positive: 17
True Negative: 2
False Positive: 8
False Negative: 3
==========
MultinomialNB Accuracy : 0.7

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.78 | 0.74 | 0.76 | 19 |
| Not Placed | 0.58 | 0.64 | 0.61 | 11 |
| avg / total | 0.71 | 0.70 | 0.70 | 30 |

Confusion Matrix MultinomialNB:
[[14 5]
[ 4 7]]
---
True Positive: 14

 

True Negative: 5
False Positive: 4
False Negative: 7
==========
BernoulliNB Accuracy : 0.8

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.78 | 0.95 | 0.86 | 19 |
| Not Placed | 0.86 | 0.55 | 0.67 | 11 |
| avg / total | 0.81 | 0.80 | 0.79 | 30 |

Confusion Matrix BernoulliNB:
[[18 1]
[ 5 6]]
---
True Positive: 18
True Negative: 1
False Positive: 5
False Negative: 6
==========
GaussianNB Accuracy : 0.7

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.68 | 1.00 | 0.81 | 19 |
| Not Placed | 1.00 | 0.18 | 0.31 | 11 |
| avg / total | 0.80 | 0.70 | 0.62 | 30 |

Confusion Matrix GaussianNB:
[[19 0]
[ 9 2]]
---
True Positive: 19
True Negative: 0
False Positive: 9
False Negative: 2
==========
MultinomialNB Accuracy : 0.633333333333

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.75 | 0.63 | 0.69 | 19 |
| Not Placed | 0.50 | 0.64 | 0.56 | 11 |
| avg / total | 0.66 | 0.63 | 0.64 | 30 |

Confusion Matrix MultinomialNB:
[[12 7]
[ 4 7]]
---
True Positive: 12
True Negative: 7
False Positive: 4
False Negative: 7
==========
BernoulliNB Accuracy : 0.866666666667

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.83 | 1.00 | 0.90 | 19 |
| Not Placed | 1.00 | 0.64 | 0.78 | 11 |
| avg / total | 0.89 | 0.87 | 0.86 | 30 |

Confusion Matrix BernoulliNB:
[[19 0]
[ 4 7]]

 

---
True Positive: 19
True Negative: 0
False Positive: 4
False Negative: 7
==========
GaussianNB Accuracy : 0.666666666667

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.66 | 1.00 | 0.79 | 19 |
| Not Placed | 1.00 | 0.09 | 0.17 | 11 |
| avg / total | 0.78 | 0.67 | 0.56 | 30 |

Confusion Matrix GaussianNB:
[[19 0]
[10 1]]
---
True Positive: 19
True Negative: 0
False Positive: 10
False Negative: 1
==========
MultinomialNB Accuracy : 0.7

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.78 | 0.74 | 0.76 | 19 |
| Not Placed | 0.58 | 0.64 | 0.61 | 11 |
| avg / total | 0.71 | 0.70 | 0.70 | 30 |

Confusion Matrix MultinomialNB:
[[14 5]
[ 4 7]]
---
True Positive: 14
True Negative: 5
False Positive: 4
False Negative: 7
==========
BernoulliNB Accuracy: 0.7

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Placed | 0.69 | 0.95 | 0.80 | 19 |
| Not Placed | 0.75 | 0.27 | 0.40 | 11 |
| avg / total | 0.71 | 0.70 | 0.65 | 30 |

Confusion Matrix BernoulliNB:
[[18 1]
[ 8 3]]
---
True Positive: 18
True Negative: 1
False Positive: 8
False Negative: 3
==========
Above details are system generated precision, recall, f1-score and support for Gaussian Naïve Bayes, Multinomial Naïve Bayes, and Bernoulli Naïve Bayes Models.

**Table 1: Accuracy report (System Generated)**

| Test Count | GaussianNB Accuracy | MultinomialNB Accuracy | BernoulliNB Accuracy |
|---|---|---|---|
| 1 | 0.666666667 | 0.566666667 | 0.733333333 |
| 2 | 0.666666667 | 0.733333333 | 0.8 |
| 3 | 0.633333333 | 0.633333333 | 0.8 |
| 4 | 0.666666667 | 0.766666667 | 0.733333333 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |

 

| . | . | . | . |
|---|---|---|---|
| 66 | 0.7 | 0.766666667 | 0.8 |
| 67 | 0.733333333 | 0.633333333 | 0.566666667 |
| 68 | 0.666666667 | 0.666666667 | 0.766666667 |
| 69 | 0.666666667 | 0.633333333 | 0.733333333 |
| 70 | 0.766666667 | 0.6 | 0.833333333 |
| 71 | 0.766666667 | 0.766666667 | 0.833333333 |
| 72 | 0.6 | 0.7 | 0.6 |
| 73 | 0.733333333 | 0.7 | 0.766666667 |
| **Average Accuracy** | **0.693607306** | **0.684931507** | **0.750684932** |
| **Maximum Accuracy** | **0.8** | **0.866666667** | **0.866666667** |
| **Minimum Accuracy** | **0.6** | **0.466666667** | **0.566666667** |

The above table (Table 1) is Accuracy report of GaussianNB Accuracy, MultinomialNB Accuracy and BernoulliNB Accuracy. It's a system generated report for 73 iterations of all three models. The system is designed to consider most possible ways to generate data and result for simplifying the complexity of problem.

From above table it's hard to see any special changes as they all are just numbers. To express the above chart into a meaningful way we can plot a graph of above table data. It represents the accuracy of each Naïve Bayes Model at current iteration. This graph is based on Table 1 data, and now we can see the fluctuations in accuracy of each model. This graph shows that the accuracy of Bernoulli Naïve Bayes model is best over Gaussian and Multinomial. Gaussian model is consistent one with minimum deflection. Multinomial is the one which shows maximum change.

**Table 2: Max-Min Accuracy**

| | **GaussianNB Accuracy** | **MultinomialNB Accuracy** | **BernoulliNB Accuracy** |
|---|---|---|---|
| **Maximum Accuracy** | 0.8 | 0.866666667 | 0.866666667 |
| **Minimum Accuracy** | 0.6 | 0.466666667 | 0.566666667 |
| **Difference** | 0.2 | 0.4 | 0.3 |

After seeing the above table (Table 2) it's clear that the difference of Max-Min in Multinomial is maximum 0.4, and Gaussian is minimum 0.2.

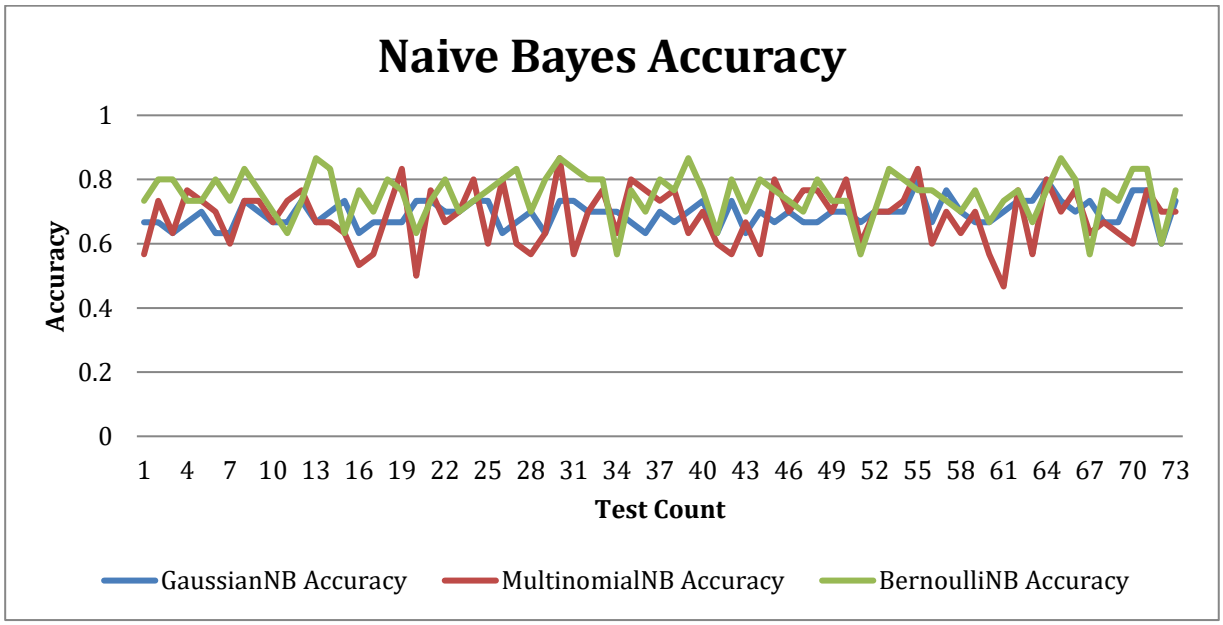


**Fig. 2: Accuracy-Times graph for Gaussian, Multinomial, and Bernoulli Naïve**

This graph is based on Table 1 data, and now we can see the fluctuations in accuracy of each model. This graph shows that the accuracy of Bernoulli Naïve Bayes model is best over Gaussian and Multinomial. Gaussian model is consistent one with minimum deflection. Multinomial is the one which shows maximum change.

**Table 2: Max-Min Accuracy**

| | **GaussianNB Accuracy** | **MultinomialNB Accuracy** | **BernoulliNB Accuracy** |
|---|---|---|---|
| **Maximum Accuracy** | 0.8 | 0.866666667 | 0.866666667 |
| **Minimum Accuracy** | 0.6 | 0.466666667 | 0.566666667 |
| **Difference** | 0.2 | 0.4 | 0.3 |

After seeing the above table (Table 2) it's clear that the difference of Max-Min in Multinomial is maximum 0.4, and Gaussian is minimum 0.2.

 

## 6. CONCLUSION

In future it can be extended to create job profile selector which help interviewers or reduce rush of students at campus drives and shortlist best candidates directly. That can save a lot more time and money for the company. It's the world of AI and ML. Every field, education, health, sports, or industry all of them are fusing it to their technology and results are amazing. At current COVID-19 situation the important role that AI and ML is playing is just beyond imagination. We never thought about occurrence of such situation when whole world will be in lockdown state. Even the biggest researchers are unable to find the actual break through for this Corona Virus. It causes a major problem for school, colleges, training and education centers. As it's nearly impossible to gather students and teach or train them. As well as problem for recruiters because it's not possible to conduct interview or campus recruitment at colleges. In such situation my project can be a handy option to shortlist good quality students from pool of student data. It will also encourage students to improve their skills as they see they need a lot more subjective as well as practical knowledge.

The collected data from 100 students, of Computer Science and Engineering, and Information Technology department is feeded to Naïve Bayes Algorithm based 3 models. Gaussian Naïve Bayes, Multinomial Naïve Bayes, and Bernoulli Naïve Bayes for predicting student placement possibility. The data set has 32 different labeled features related to students present and past education qualification, as well as confidence and strength in industry demanded subjects. Availability of more data can make it even more accurate, we achieved about 0.693 average accuracy score for Gaussian Naïve Bayes model, 0.684 average for Multinomial NB, and highest accuracy rate of 0.75 for Bernoulli NB. Accuracy count 0.75 means Bernoulli Naïve Bayes can predict with 75% correctness.

**Table 3: Accuracy Table of Gaussian, Multinomial and Bernoulli Naïve Bayes Model.**

| | GaussianNB Accuracy | MultinomialNB Accuracy | BernoulliNB Accuracy |
|---|---|---|---|
| **Average Accuracy** | 0.693607306 | 0.684931507 | 0.750684932 |
| **Maximum Accuracy** | 0.8 | 0.866666667 | 0.866666667 |
| **Minimum Accuracy** | 0.6 | 0.466666667 | 0.566666667 |

With the help of AI and ML we can automate almost every kind of analytical analysis and reach to results. Predicting student's performance and other qualities on basis of student skills and previous performance will really help many students, institutes and industry. AI and ML, are booming concepts of today's generation. Application of these technologies will change the world. This project will contribute to the education system by helping students, their parents, college and industry. This project is just another application of technology in field of education.

## 7. REFERENCES


[1] B. K. Baradwaj and S. Pal, "Mining educational data to analyze students' performance," Int. J. Adv. Comput. Sci. Appl., vol. 2, no. 6, pp. 63–69, 2011.

[2] S. Borkar and K. Rajeswari, "Predicting students academic performance using education data mining," Int. J. Comput. Sci. Mob. Comput., vol. 2, no. 7, pp. 273–279, 2013.

[3] Anal Acharya,Devadatta Sinha, Application of Feature Selection Methods in Educational Data Mining International Journal of Computer Applications (0975 – 8887) Volume 103 – No.2, October 2014.

[4] E. Osmanbegovic, M. Suljic, and H. Agic, "DETERMINING DOMINANT FACTOR FOR STUDENTS PERFORMANCE PREDICTION BY USING DATA MINING CLASSIFICATION ALGORITHMS," Tranzicija, vol. 16, pp. 147-158, 2015.

[5] S. Karthika and N. Sairam, ?A Naïve Bayesian Classifier for Educational Qualification,? vol. 8, no. July, 2015.

[6] Strecht, P.; Cruz, L.; Soares, C.; Mendes-Moreira, J.; Abreu, R. A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance. In Proceedings of the 8th International Conference on Educational Data Mining, Madrid, Spain, 26–29 June 2015.

[7] Kaur, P.; Singh, M.; Josan, G.S. Classification and prediction based data mining algorithms to predict slow learners in education sector. Procedia Comput. Sci. 2015, 57, 500–508.

[8] M.Mayilvaganan, D.Kalpanadevi,"Comparison of Classification techniques for Predicting the Cognitive skill of students in education environment", IEEE, September 2015.

[9] C.Anuradha, T.Velmurugan, "Feature selection techniques to analyse student acadamic performance using naïve bayes classifier", The 3rd International Conference on Small & Medium Business 2016 January 19 - 21, 2016, Nikko Saigon Hotel, Hochiminh, Vietnam 345

[10] Mashael A. Al-Barrak, Muna Al-Razgan, ?Predicting Students Final GPA Using Decision Trees: A Case Study?, Int. Journal of Information and Education Technology, Vol.6, No.7, 2016, pp.528-533.

[11] Mueen, A.; Zafar, B.; Manzoor, U. Modeling and predicting students' academic performance using data mining techniques. Int. J. Mod. Educ. Comput. Sci. 2016, 8, 36–42.

[12] Meier, Y.; Xu, J.; Atan, O.; van der Schaar, M. Personalized grade prediction: A data mining approach. In Proceedings of the 2015 IEEE International Conference on Data Mining, Atlantic City, NJ, USA, 14–17 November 2015.

[13] Mokhairi makhtar, hasnah nawang, syadiah nor wan shamsuddin, "Analysis on students performance using naïve bayes classifier", Journal of Theoretical and Applied Information Technology 31st August 2017. Vol.95. No.16

[14] Xu, J.; Moon, K.H.; van der Schaar, M. A machine learning approach for tracking and predicting student performance in degree programs. IEEE J. Sel. Top. Signal Process. 2017, 11, 742–753.

[15] Astha Soni, Vivek Kumar, Rajwant Kaur, D. Hemavathi, International Journal of Pure and Applied Mathematics, Volume 119.

[16] Kaplan, R.M., & Saccuzzo, D.P. (2010). Psychological Testing: Principles, Applications, and Issues. (8th ed.). Belmont, CA: Wadsworth, Cengage Learning.


 

[17] Maryam Zaffar, Manzoor Ahmed Hashmani, K.S. Savita, Syed Sajjad Hussain Rizvi(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 5, 2018, Page 541.
[18] Reeta R, Pavithra G, Priyanka V and Raghul J S., Predicting Autism using Naive Bayesian Classification Approach, International Conference on Communication and Signal Processing, April 3-5, 2018, India
[19] Sathya Durga, J. Thangakumar, "A complete survey on predicting performance of engineering students", International Journal of Civil Engineering and Technology (IJCIET) Volume 10, Issue 02, February 2019, pp. 48?56, Article ID: IJCIET_10_02_007
[20] Maria Tsiakmaki, Georgios Kostopoulos, Sotiris Kotsiantis and Omiros Ragos, Implementing AutoML in Educational Data Mining for Prediction Tasks, Appl. Sci. 2020, 10, 90; doi:10.3390/app10010090
[21] Fahad Razaque., Nareena Soomro., Shoaib Ahmed Shaikh., Safeeullah Soomro., Javed Ahmed Samo., Natesh Kumar. and Huma Dharejo. Using Naïve Bayes Algorithm to Sudents' Bachelor Academic Performances Analysis. Proceedings 4th IEEE International Conference on Engineering Technologies and Applied Sciences, Salmabad, 2017, pp. 1-5.
[22] Srinivas, A and Ramaraju, V. Detection of Failures by Analysis of Student Academic Performance using Naïve Bayes Classifier. International Journal of Computer & Mathematical Sciences, 7(3), 2017, pp. 277-283.
[23] Bendangnuksung and Prabu, P. Students' Performance Prediction Using Deep Neural Network. International Journal of Applied Engineering Research, 13(2), 2018, pp. 1171-1176.
[24] V. Sathya Durga, J. Thangakumar, "A complete survey on predicting performance of engineering students" International Journal of Civil Engineering and Technology (IJCIET) Volume 10, Issue 02, February 2019, pp. 48–56, Article ID: IJCIET_10_02_007.
[25] Brijmohan Lal Sahu, Shahnaz Parveen, Yasmeen Dani, "Glimpse of future Artificial intelligence and its impact on human lifestyle", Published in international journal of advance research, ideas and innovations in technology, KITE/NCISRDC/IJARIIT/2018/IT/101.
[26] https://www.ijariit.com/conference-proceedings/41%20IT%20101.pdf

 