IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 11, Iss 08, 2022

FORECASTING THE STUDENT'S PLACEMENT HARNESSING THE POWER OF MACHINE LEARNING

#1A.Srinivasa Rao, Associate Professor,
#2S.Sunil Kumar, Assistant Professor,
#3B.Srinivasa Rao, Assistant Professor,
Department of CSM,
SAI SPURTHI INSTITUTE OF TECHNOLOGY, SATHUPALLY, KHAMMAM.

Abstract – Higher education institutions view the placement procedure as a major concern. The placement variables have a tremendous influence on both the names of the schools and the people who choose to attend them. As a result, educational institutions such as colleges and universities work very hard to improve the programs that they offer to students in order to help them find work. The main goal of this article is to assess the previous academic year's student placement statistics, speculate on what the placement outcomes would be like for current students, and offer some suggestions on how the educational system might improve its placement rate. This study provides a strategy to help educational institutions choose applicants for admission. You can predict how well a company will perform in the future if you successfully place a student and then use data from previous students who have been put in the same firm. Machine learning employs a number of classification algorithms, including the Naive Bayes Classifier and the K-Nearest Neighbors (KNN) approach. Both of these algorithms categorize data. The algorithms create their own predictions about the outcomes, and the dataset is used to assess how well the algorithms did their jobs. A corporation's positioning division may use the aforementioned framework to find applicants for employment and aid those individuals in focusing their education on the development of their technical and social talents.
***Key Words*:** Machine Learning, Naive Bayes, K- Nearest Neighbors (KNN), Database.

1. INTRODUCTION

There has been a noticeable increase in the number of newly founded educational institutions in recent years. In most circumstances, every school will have a placement office, and the major aim of these offices is to help people find decent jobs that pay well. Colleges and universities are under intense pressure to increase the employment rates of their graduates. There is a one-to-one relationship between the quality of available educational options and the complexity of making a placement decision. Keeping up to date on new educational practices and institutions is a critical component of any complete quality management system. Machine learning methods are frequently used to extract information from a wide range of sources, including research databases, archives, and expert knowledge bases. A dataset containing historical information about students is utilized to set the system in motion. These discoveries are used to assess the model's capacity to categorize data and make it easier to train the model to spot patterns and generalizations. Students are greatly motivated when they realize that there are several career opportunities in many different fields related to technology, and as a result, they work harder in school. There are extra benefits for the university's placement office and teachers in the classroom who use this technology to help their students perform well academically. The percentage of a school's graduates who are able to find work after graduation is an important credibility element. As a result, it is critical that this methodology be used in all of the university's classes. To achieve this goal, we employ various machine learning algorithms such as Naive Bayes and K-Nearest Neighbors (KNN).

Prediction System

The goal of this study is to use machine learning approaches to make informed placement decisions

based on information provided in written form. To detect the location of the object, machine learning methods such as Naive Bayes and the K-nearest neighbor (KNN) algorithm are used. This approach takes into account more than just a candidate's technical and aptitude qualities. A person's USN ranking, achievement in the tenth grade, completion of the Pre-University Course (PUC) or diploma, and Cumulative Grade Point Average (CGPA) scores are all considerations considered.

Naive Bayes Classifier

A wide range of data from the real world has been used to demonstrate the efficiency of the Naive Bayes Classifier. When the univariate conditional probabilities are precisely predicted and the factors are carefully selected, the Naive Bayes strategy typically outperforms other methods.

K-nearest neighbor (KNN)

The K-nearest neighbors (KNN) approach uses the full training dataset to make educated assumptions about the features of incoming data. Following a thorough assessment of the material in the training set, the k samples with the greatest degree of similarity are identified. Following that, the data with the most similarities is collected and shown.

Features of KNN

- Following the application of the K-Nearest Neighbors (KNN) model, the entire training dataset is saved.
- To make real-time predictions, first determine the degree of similarity between an input sample and each training case.

2. RELATED WORK

Senthil Kumar Thangavel, Divya Bharathi P, and Abhijith Shankar conducted the research, which focused on the creation of placement prediction algorithms. The individual's grade point average and the community in which they lived were taken into account. The researchers used both Decision Tree Learning and the SCI-Kit machine learning technologies for the forecasting analysis. The forecasts were based on a few characteristics, including the student's overall grade point average and whether or not they had any unpaid bills at the time of the prognosis. Wilton Y.S. He, H.H. Au Yeung, K.Y. Law, and W.T. FOK, together with their partners, used a neural network to discover the most effective technique for students to select their classes. The amount of intermediary nodes and deep learning layers present within the TensorFlow system is modified and examined.

Machine learning research comprises the development, testing, and examination of algorithms capable of discovering patterns in data on their own and then utilizing those patterns to make accurate projections of future data or intelligent choices. Machine learning streamlines the process of generating models that are applicable to real-world scenarios, which speeds up the total process. Machine learning is increasingly being used in academic domains such as bioinformatics, computer vision, robot locomotion, computational finance, and search engines.

3. METHODOLOGY

Naive Bayes

Classifiers are bits of software that, as the name implies, help computers recognize unwanted situations. These troubling occurrences are often represented as feature value vectors. There is a limited number of class name alternatives available. In the process of teaching computers to recognize categories, a set of techniques known as the Naive Bayes classifiers is used. The value of one feature does not depend on the value of any other features for a given class variable, according to these classifiers. Furthermore, it is critical to remember that no single algorithm is capable of readily performing all of the obligations. Apples

are a good example because they have unique qualities such as being red, round, and 11 cm in width. The Naive Bayes classifier evaluates an object's color, roundness, and width separately to arrive to an estimate of the object's chances of being an apple.

Working of Naive Bayes Algorithm

Before beginning any type of data mining, the data currently saved on the computer must be thoroughly reviewed. This is done in order to ensure that only relevant information is captured. This information can be gathered more quickly by using databases, cloud-based storage systems, Excel spreadsheets, and other comparable programs.

Determine the likelihood that each possible combination of attribute numbers will occur. In this situation, we will look into the elements n, n_c, m, and p. In this inquiry, the chance of each attribute being present is calculated using the mathematical equation below. The events that occurred are detailed in further detail in the following chapter. Everyone is expected to do the math correctly in every lesson.

What exactly happens in the third phase? Putting the prior numerical data into practice The formula (n_c+mp)/(n+m) can be used to calculate the probability of the attribute value (ai) given the subject value (vj). In this instance, the symbol n represents the number of training examples in which the subject value v is the same as vj. Furthermore, the value nc denotes the number of times the subject value v and trait value an are identical to the values vj and ai, respectively. In the first step of the procedure, the variable p displays an estimate of the probability P(aivj).

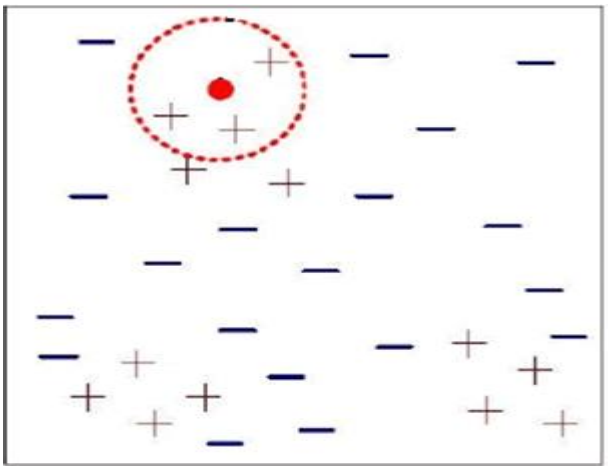
The value marked by the letter "M" represents the total number of samples gathered.

In the fourth phase, the probability value of each group is multiplied by the number of possible outcomes (p), increasing the number of outcomes. In this method, the numbers that correspond to each feature are changed by using a multiplier, marked by the letter p. The objects being examined are classified into subcategories using the whole set of results.

The fundamental goal of the research is to categorize each attribute number based on the traits that are unique to that number.

K-nearest Neighbors classifier (KNN)

Using the K-Nearest Neighbors (KNN) method to categorize objects is a straightforward, rapid, and parameter-free approach [3, 4]. The K-nearest neighbors, or KNN, method is a statistical tool for evaluating traits or features that are stable over time. K-nearest neighbors, abbreviated as KNN, are used for more than just detecting patterns and generating statistical estimates. Case-based reasoning (CBR) is a problem-solving strategy that entails applying what you've learned from prior events to help you make decisions in the present and future. Computing the nearest neighbors is a valuable strategy for gaining significant information on how to classify an unlabeled sample. Because it is simple to implement and gives reliable results, the K-Nearest Neighbors (KNN) technique is preferred over other classification algorithms the great majority of the time. The diagram depicts how "near neighbors" are placed together. There are two main stages within the scope of the K-nearest neighbors (KNN) classification approach.



- Determine which instance in the collection bears the most resemblance to the instance given to you (in this case, S).
- Instances k, as previously stated, hold a democratic vote to determine the type of the entity denoted by S.
- The performance of the K-nearest neighbors (KNN) algorithm is totally reliant on the distance metric employed and the number of closest neighbors. A considerable number of people use the cosine and Euclidean distances to determine how unlike two occurrences are to one another.

Working of KNN Algorithm

It has been reported that the proposed technique should increase the performance of the K-nearest neighbors (KNN) classifier when estimating health-related parameters. The following formula describes the method that has been proposed.

Could you please enter the file to begin?

The K-Nearest Neighbors (KNN) technique is then used to the data provided at this point in the process.

Determine the metric discrepancies that exist between the previously defined parameters and the parameters acquired during the third phase.

In the final step of the procedure, data on similarities is used to estimate the expected outcome.

The nearest neighbor algorithm measures the distance $d_E(X_i,X_j)$ between query points X_i and a set of training samples X_j to classify a new object based on majority of K-nearest neighbor category of Y attributes of training samples.

Query point $X_i = x_1, x_2, x_3, \ldots\ldots, x_n$

Training Sample $X_j = x_1, x_2, x_3, \ldots\ldots, x_n$

$$Dist(c_1,c_2) = \sqrt{\sum_{i=1}^{N}\left(attr_i(c_1) - attr_i(c_2)\right)^2}$$

$$k - NearestNeighbors = \left\{k - MIN(Dist(c_i,c_{test}))\right\}$$

$$prediction_{test} = \frac{1}{k}\sum_{i=1}^{k} class_i \ (or \ \frac{1}{k}\sum_{i=1}^{k} value_i)$$

4. FUTURE ENHANCEMENT

In order to improve the accuracy of placement predictions, the project will prioritize the incorporation of additional attributes in any future upgrades. It is possible to do more if you actively seek out and assess prospective suggestions or ideas for increasing the system's output.

5. CONCLUSION

The cutting-edge device known as the Student Placement Predictor employs machine learning technologies to generate credible placement predictions. Academics working in the field of education are primarily interested in how academic success can be assessed and predicted among pupils. All of these forecasts can be used by the university to help its students perform better and achieve greater grade point averages. Previously, experts relied on a small set of indicators, such as a person's grade point average (GPA) and record of delinquency, to make reasonable judgments about where a person would be placed. The results acquired using this method, however, are regarded as having a lesser level of credibility. This other study, on the other hand, predicts where students would be placed based on a broader range of educational criteria, resulting in more dependable outcomes.

REFERENCES

[1] "Student Placement Analyzer: A Recommendation System Using Machine Learning" 2017 International Conference on advanced computing and communication systems (ICACCS-2017), Jan 06-07,2017, Coimbatore, INDIA.

[2] "Prediction Model for Students Future Development by Deep Learning and TensorFlow Artificial Intelligence Engine" 2018 4th IEEE International Conference on Information Management.

[3] Kohavi, R and F, Provost (1998). Machine Learning30:271-274.