IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 3, June 2023

Students Placement Sign Anticipation System Using Artificial Intelligence

Nilesh Khedekar and Dr Priti Shende

Department of Electronics and Telecommunication Engineering

Dr. D. Y. Patil Institute of Technology, Pimpri, Pune, India

nileshkhedekar@gmail.com, pritijawale76@gmail.com

Abstract: *This paper presents an overview of artificial intelligence algorithms that use neural networks to predict a student's placement and give technical data. All educational institutions have a very crucial duty: the ability to predict a student's success. This will not be chosen only based on a student's academic standing. It is crucial to consider a student's behaviour, including their aptitude, attitude, communication style, technical competence, problem-solving skills, and demonstration-related information, to predict how they will do in the actual demonstration. Since this is the issue with predicting where undergraduate students will be placed, this study may be used to calculate the probability that an undergraduate student will be employed using different AI algorithms.*

Keywords*:* Artificial intelligence, Placement prediction

I. INTRODUCTION

Engineering students' final or third year of school is when the stress of the placement starts to show. They need to know where they stand and how to hone their skills to improve their chances of landing a job. This research provides an overview of artificial intelligence techniques that might be used to predict a student's placement success. For many educational institutions, predicting a student's success is essential. This will not be chosen only based on a student's academic standing. Considerations such as aptitude, attitude, communication, technical proficiency, aptitude for problem-solving, etc., should be considered to predict a student's success. The chance that first- and second-year students will be placed may be predicted using this essay. Assess the student's readiness for the placement within the first year. The student's weak area should be improved in the second and third years. Please make an effort to put more students in their last year. This approach may help to improve the likelihood of student placement. Artificial intelligence: The use of artificial intelligence (AI) that gives systems the capacity to automatically pick up new skills and refine existing ones based on experience without having to be explicitly programmed. Artificial intelligence aims to create computer systems that can access and utilize data to learn independently. It enables computers to learn autonomously and modify their behaviour without human input.

Naive Baye's Classifier and K- Nearest Neighbors [KNN] method are two machine learning classification algorithms proposed by Shreyas Harinath et al. (2019). These systems use artificial intelligence to forecast student placement on their own. The effectiveness of the algorithms was then compared using the results. Their technical abilities are strengthened by using this model. The proposed model by NeelamSWaroopa et al. (2019) uses an algorithm to anticipate the same. The same institution where the placement forecast is made provided the data, and appropriate pre-processing techniques were also used. Regarding accuracy, fidelity, and recall, this model also contrasts with conventional classification methods like a Decision tree and Random Forest. The findings show that the suggested approach performs noticeably better when compared to the other methods stated.

According to Shubham Bavane et al. (2019), this technique can quickly identify student knowledge and may be helpful to several educational institutions. The Nave Bayes, SVM, and KNN algorithms are used in this system to forecast student performance. Education organizations may provide their students with better training based on the student's knowledge set for placement and non-placement classes.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 3, June 2023

Tansen Patel et al. presented research on several clustering algorithms and compared the outcomes of placement prediction for higher education in 2017. Behrouz Minaei-Bidgoli, Deborah A. Kashy, et al. (2013) have demonstrated how combining several classifiers improves accuracy noticeably. The advantages of using the LON-CAPA data to forecast the students' final grades based on their attributes, which are retrieved from the homework data, are shown by the successful optimization of student categorization in all three scenarios.

By incorporating a common topic of artificial intelligence, the NSF-funded research led by Zdravko Markov, Ingrid Russell, et al. (2005) hopes to improve student learning outcomes in the introductory artificial intelligence course. In their 2017 study, Mukesh Kumar et al. examined how to forecast student academic performance, educational dropout rates shortly, institute placement, and admission to a new academic year. This approach is utilized to enhance the teaching and learning process.

Compared to the Decision Tree, Naive Bayes, Bayes Network, and CART, the Random Forest method for predictive modelling by Mukesh Kumar et al. (2017) produced the best results. In 2018, Amandeep Kaur, Nitin Umesh, et al. developed a system that reduces college dropout rates. Analysis of varied educational data was aided by education mining. In this situation, the hybrid artificial intelligence classification strategy performs better in accuracy and obtaining accurate prediction outcomes.

An NN prediction model has been given by PauziahMohdArsad, NorlidaBuniyamin, et al. (2013) to forecast the academic performance of Electrical Degree students based on multiple entrance levels, particularly Matriculation and Diploma entry levels. A review paper on a comprehensive review of Deep Learning techniques for educational data mining was proposed by David Tomás et al. in 2019. The primary objectives of this study were to pinpoint the EDM activities that Deep Learning has made more effective. K.

Prasada Rao et al. (2016) developed system classification approaches to forecast learning behaviour on a student database. This study aids in early failure detection and identification of slow learners. The J48, Naive Bayes, and Random Forest algorithms are also contrasted in this work. Karishma

A study on categorization algorithms to suggest well-behaved carriers for students was given by B. Bhegade et al. (2016). Unruly and aggressive students have an impact on their careers. Decision trees are known for producing classification rules that are simple to understand. Different classifier types are tested for accurate calculation and demonstration. Students' presentation outperforms other methods by making insightful predictions and getting the best outcomes. M. Usha Rani et al. presented a method utilizing EDM in 2019. The classification was performed to anticipate pupils in various class groups, such as High, Medium, and Low—Support Vector Machine classifiers (SVM). In this article Bendangnu Ksung and others (2018), a deep neural network model for forecasting student demonstration was put out. With this model, accuracy was 84.3% with more records and attributes in the dataset. E. Chandra and K. Nandhini et al. (2007) presented a study on the issue of demonstration prediction and discovered that it is feasible to forecast students' demonstrations automatically. Furthermore, the ability to quickly and consistently incorporate such knowledge into the learning job is made feasible using extensible classification formalisms like Bayesian networks.

The primary contribution of the article by Akhilesh P. Patil et al. (2017) was the comparison of several recurrent neural architectures. In contrast to a decision tree, SVM, and feed-forward neural network-based algorithms that have previously been used to solve this problem of student score prediction, the novel aspect of the proposed method is that it has memory to distinguish tuples with different orders of scores and learns to assign the weights of the relationship between nodes by scanning the sequence in both directions.

A work employing Tensor flow for deep learning and artificial intelligence direction to solve classification problems and anticipate YS-reported non-linear outcomes. He et al. (2018). The final results were between 80% and 91 %. A strategy for forecasting students' future demonstration in degree programmes based on their present and historical demonstration was put forth by Jie Xu et al. (2018). The study by Paul Akangah et al. (2018) examined the correlations between predictor factors such as GPA>3.0, "PASSED PHYS241", "PASSED Q," and "PASSED RQ" and the pass rate in MEEN241 (64.1%). Assessments and class assignments were created.

Different mathematical models have reportedly been used to anticipate students' demonstrations, according to Sahar Al-Sudani et al. (2019). In order to categorize students' degrees into either a good or introductory degree class, He employed a mix of institutional, academic, demographic, psychological, and economic aspects in this study. He used a multi-layered neural network (NN) to anticipate students' displays.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 3, June 2023

In order to anticipate real-time student performance, which has been deemed the most prominent significance in MOOCs (but under-explored), Byung-Hak Kim et al. (2019) introduced a beneficial method. This method allows predicting student outcomes while a course is still in progress. Four modelling methodologies—neural networks, logistic regression, discriminant analysis, and structural equation modelling—were compared in a paper by Joe J.J. Lin et al. 2009.

To examine the placement information of the students, Karan Pruthi and Dr Parteek Bhatia et al. (2015) submitted a study based on the decision tree algorithm. This process helped identify the responsible Department and the placement coordinator to build ways to enhance students' academic performance, coding abilities, and soft skills. This approach will be crucial in enhancing the institute's overall placement rates.

The creation of a placement prediction system (PPS) employing a logistic regression model was proposed by Keshav Kumar et al. in 2014. Lakshmi Priya K et al. (2017) provided a model to examine data mining approaches to research the behaviour of graduates regarding their job preferences, by analyzing data such as mobile phone usage, internet connectivity, course assignments, study hours, interestingness, and grade.

A study on the prediction of student academic demonstration using different categorization algorithms was given by Jai Ruby et al. in 2014. For the student dataset, all classification methods, MLP, ID3, J48, REP Tree, NB Tree, Simple Cart, and Decision Table,the exhibit prediction accuracy of over 68%. Ankita A. Nichat et al. (2017) established a model for data mining approaches, and the findings of their study show that these techniques may be used to improve analytical student demonstrations. 6 methods were provided in a paper by Oktariani Nurul Pratiwi et al. (2013) that may be used to categorize student data.

A report on applying several classification algorithms utilizing Data Mining tools (WEKA) for the examination of students' departmental placement was given by GetanechBerieTarekegn et al. in 2016. The study developed a prediction model for student placement using J48, Naive Bayes, and Random Forest algorithms. According to a model developed by Namita Puri et al. (2015), ID3 is the best method for classifying students and predicting their placement in engineering colleges. The best classifier, with a 95 percent accuracy rate, is the ID3 decision tree method.

In order to validate the methods, Mangasuli Sheetal B et al. (2016) presented a work that had been examined and forecasted utilizing the fuzzy logic and KNN algorithms. K. Nasaramma et al. (2017) demonstrated how C5.0 properly categorizes student data, assisting the placement coordinator in identifying students who fall short in particular areas. D. Ganesh Gopal et al. (2014) developed a model Sum of a different approach to accomplish the objective and extract patterns from the provided dataset. Praveen Rani et al. (2015) reported research that used the j-48 classification method with straightforward K-Mean clustering to separate students. J. Jayanthi et al. (2017) proposed a student prediction method and utilized it to separate the student data and information based on the student demonstration.

Siddhi Parekh et al. (2016) introduced a dashboard system that shows statistics in graphs and charts, enabling simple comprehension of students' academic standing at any given time. Mansi Gera et al. (2015) have created a proposed methodology to forecast a student's placement eligibility so that they may only prepare for organizations for which they will be qualified. AnimeshGiri et al. (2016) suggested a placement prediction system that uses artificial intelligence to use the k-nearest neighbour's classification model to forecast the likelihood that an IT business would hire an undergraduate student. Professor Ashok M Assistant Professor Apoorva A, 2016. "Data Mining Approach for Predicting Student and Institution's Placement Percentage." The author employed various data mining methods in this study, including decision trees, Naive Bayes, neural networks, and the suggested approach. Decisions were made with the use of confusion matrices.

A TPO management system to forecast qualified candidates for campus drives was proposed by Syed Ahmed et al. in 2017. Liya Claire Joy et al. (2019) presented a placement predictor method to forecast the likelihood or the kind of employer where a pre-final year student would be placed. The work provided by ChandiniLulla, Yash Agarwal, et al. (2017) was based on the findings of an evaluation procedure used in conjunction with data mining tools. As a consequence, table 1 provides the formulation of a predictive model. The accuracy of each algorithm is then analyzed.

Table1: List of attributes of the dataset

Attributes	**Type of data**
Branch	Nominal

Gender	Nominal{M/F}
10 th%	Numeric
10thBoard	Nominal
12 th%	Numeric
12thBoard	Nominal
Diploma%	Numeric
FE-I SEM%	Numeric
FE-IISEM%	Numeric
SE-I SEM%	Numeric
SE-IISEM%	Numeric
TE-ISEM%	Numeric
Aggregate Engineering%	Numeric
Live Backlogs	Numeric

Classification Techniques for Prediction

Naïve Bayes Classifier

Numerous real data applications benefit significantly from the Naive Bayes Classifier. Building a classifier model using this method is simple. The naive bees classification approach relies on the "Bayes Theorem" and assumes that predictors are independent. It is relatively straightforward and supposes that the classification attributes are independent. The maximum likelihood incidence and conditional probability form the foundation of the creative Naive Bayesian approach.

K-nearest Neighbour (KNN)

KNN (IBK) is a simple and lazy classification method that trains on the entire dataset. By figuring out a sample's class, one may categorize an unknown sample by finding its nearest neighbours. Due to how rapidly and effortlessly it converges, KNN is favoured over other classification algorithms. The data with the most comparable cases are shown after exploring the training dataset for the k-most similar occurrences.

Support Vector Machine

The Support Vector Machine (SVM / SMO) supervised learning model is used for categorization. SVM is suitable for binary classification, after all. Using the available training data, SVM creates a model. This data is represented as points in space by the SVM model. The SVM tries to map test data to the same space as soon as it gets it.

Logistic Regression:

Logistic regression is a statistical method for analyzing data that includes one or more independent factors influencing an outcome. The result of this process is measured using a dichotomous variable. Given that its input values can range from negative to positive infinity and its output is restricted to zero and one, the logistic function can be used to determine probabilities.

Decision Tree

A decision tree is used in artificial intelligence choice tree learning as a prediction model that connects observations about an object to assessments of the object's intended value. Regression trees or classification trees are more poetic names for these tree models. The branches of these "tree-like" structures represent feature conjunctions, while the leaves represent classes. Building decision trees facilitates decision-making.

J48Algorithm:

The WEKA project team of Ross Quinlan developed a variant of ID3 called J48. Additional features of J48 include accounting for missing data, decision tree pruning, continuous attribute value ranges, the creation of rules, etc. To

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 3, June 2023

create decision trees, a preset set of examples is employed. The tree that results are used to categorize subsequent samples. The example belongs to a class and exhibits various traits (like Yes or No). A decision node is a non-leaf node, whereas the leaf nodes of the decision tree have the class label value applied to them. Each node branch represents a potential value for the attribute being checked at the decision node. Similar to ID3, C4.5 builds decision trees from a training data set using the information entropy concept.

Random forest:

The random forest method may also be considered an AI ensemble strategy. A random forest method receives as input a dataset of records with attributes. The input is divided into created random subgroups. A decision tree will be constructed for each generated random subset. While forecasting for both the test sample and the supplied data, Random Forest constructs many decision trees.

II. TOOL AND TECHNIQUES

For the sake of this study, MS Excel and WEKA (Waikato Environment for Knowledge Analysis) software applications can be used. The well-known artificial intelligence software package WEKA, written in Java, was developed at the University of Waikato in New Zealand. WEKA is free software as defined by the GNU General Public License. The WEKA workbench has several visualization tools, algorithms, and graphical user interfaces, allowing easy access to these capabilities.

The WEKA tool has many packages, including Filters, Classifiers, Clusters, Associations, and Attribute Selection. Using the WEKA Visualization tool, you may visualize datasets and classifier predictions by calling a dataset from your own Java code. The ARFF should be followed while creating WEKA datasets.

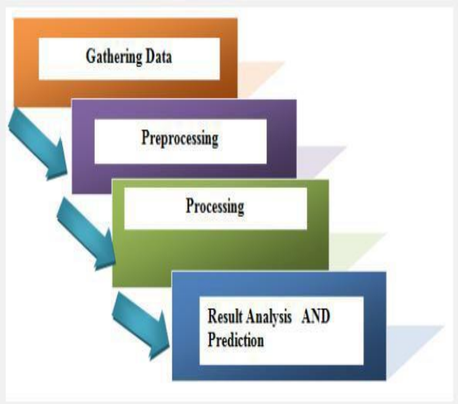

Figure 2: Implementation of the System

III. RESULT AND ANALYSIS

The gathered information includes the outcomes from 1000 pupils during the last two years. There are 1000 occurrences in the collection, along with 19 characteristics. The data file must use either "CSV" or "ARFF" format. Data for this experiment was converted from and kept in MS Excel. arff (Attribute-Relation File Format). This file served as the input for the WEKA 3.8.4 tool, which produced the results. The data processing is shown in Figure 2. Data Pre-processing is the initial stage of this project's assessment. Select WEKA Explorer interfaces for the classification model for this project.

Table 2 displays the results of a 10-fold cross-validation of student data. Here, contrast the LWL, Naive Bayes, Logistic, Multilayer Perceptron, SMO, J48, Random Forest, and LMT demos. A classification problem's predicted outcomes are compiled in a confusion matrix. Count values describe the number of accurate and inaccurate predictions for each class. This is the confusion matrix's secret. The confusion matrix demonstrates how your classification model produces predictions while being confused. Table 3 provides the confusion matrix for various classifiers' prediction results.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 3, June 2023

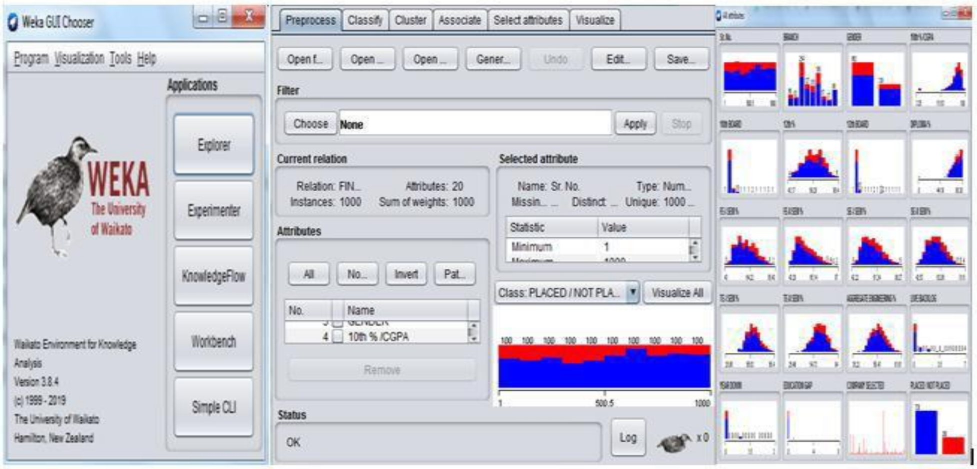

Figure 3 WEKA Explorer with processing database

Table 2: The confusion matrix for the prediction of different classifiers

Classifiers		a	b
LWL	a	719	1
	b	2	278
Logistic	a	715	5
	b	49	231
MultilayerPerceptron	a	717	3
	b	2	278
SMO	a	719	1
	b	39	241
LMT	a	719	1
	b	39	241

The prediction model/classifier accuracy is defined as the total number of correctly predicted/classified instances.

$$Accuracy = TP + TN / TP + FP + FN + TN * 100$$

TP, TN, FN, and FP represent the number of true positive, truenegative, falsenegative, and falsepositivecases. Table 3 shows the comparison of accuracy for classifiers.

Table 3: Students' placement prediction system results with accuracy

Classifiers	Accuracy	Time i s taken to build a model in seconds
LWL	99.7%	0
Naïve Bayes	69%	0.1
Logistic	94.6%	0.95
Multilayer Perception	99.5%	116.75
SMO	96%	0.94
J48	71.6%	0.17
Random Forest	70.8%	1.31

Model	LWL	Naïve Bayes	Logistic	Multilayer Perception	SMO	J48	Random Forest	LMT
Correctly classified instances	0.997	0.69	0.946	0.995	0.96	0.716	0.708	0.96
Incorrectly classified instances	0.303	0.31	0.054	0.05	0.04	0.284	0.292	0.404
kappa statistic	0.99	0.273	0.859	0.987	0.896	0.1833	0.1363	0.896

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 3, June 2023

Mean absolute error	0.0063	0.337	0.081	0.0068	0.04	0.367	0.349	0.166
Root mean squared error	0.0557	0.490	0.217	0.071	0.2	0.455	0.433	0.218
Relative absolute error	0.05	0.8366	0.2023	0.0168	0.099	0.911	0.8651	0.411
Correctly classified instances	0.997	0.69	0.946	0.995	0.96	0.716	0.708	0.96
Incorrectly classified instances	0.03	0.31	0.054	0.05	0.404	0.284	0.292	0.404
kappa statistic	0.99	0.273	0.859	0.987	0.896	0.1833	0.1363	0.896
Mean absolute error	0.0063	0.337	0.081	0.0068	0.04	0.367	0.349	0.166
Root mean squared error	0.0557	0.490	0.217	0.071	0.2	0.455	0.433	0.218
Relative absolute error	0.01563	0.8366	0.20237	0.0168	0.099	0.9112	0.8951	0.411

IV. CONCLUSION

This research paper shows evolution results obtained after 10-fold cross-validation on classifiers. Here LWL and Multilayer Perceptron classifiers give accuracy up to 99.7% and 99.5%,respectively. The accuracy of these two algorithms is better than other algorithms. Also, the Logistic, SMO and LMT give94.6 %, 96%, and 96 % accuracy, respectively. The LWL classifier suits this model because it builds the model in less time with reasonable accuracy.

REFERENCES

[1]. PauziahMohdArsad, "A Neural Network Students' Demonstration Prediction Model (NNSPPM)," Proc. of the IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), 26-27 November 2013.

[2]. David Tomás, "A Systematic Review of Deep Learning Approaches to Educational Data Mining,"Hindawi Complexity Volume, 2019.

[3]. K.Prasada Rao, "Predicting Learning Behavior of Students using Classification Techniques,"International Journal of Computer Applications (0975 – 8887) Volume 139 – No.7, April 2016

[4]. Karishma B. Bhegade, "Student Demonstration Prediction System with Educational Data Mining,"International Journal of Computer Applications (0975 – 8887) Volume 146 – No.5, July 2016

[5]. Usha Rani, "Student's Demonstration Prediction using Deep Learning and Data Mining Methods," International Journal of Recent Technology and Engineering (IJRTE)ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019

[6]. Bendangnuksung, Dr Prabu P "Students' Demonstration Prediction Using Deep Neural Network," International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 2,2018

[7]. E.Chandra and K.Nandhini, "Predicting Student Demonstration using Classification Techniques," Proceedings of SPIT-IEEE Colloquium and International Conference,

[8]. Mumbai, India,Vol.5,83

[9]. Akhilesh P Patil, Karthik Ganesan, Anita Kanavalli, "Effective Deep Learning Model to PredictStudentGrade Point Averages," IEEE,2017

[10]. Wilton W.T. Fok, Y.S. He, H.H. Au Yeung, K.Y. Law, KH Cheung, YY. Ai, P. Ho, "Prediction Model for Students' Future Development by Deep Learning and Tensorflow Artificial Intelligence Engine," 4th IEEE International Conference on Information Management2018

[11]. Shreyas Harinath, Aksha Prasad, Suma H S, "StudentPlacementPredictionUsingArtificial Intelligence," International Research Journal of Engineering and Technology(IRJET), e-ISSN: 2395-0056, Vol. 06, Issue 04, Apr 2019.