Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

Predicting Student's campus placement chance using machine learning algorithms

Vinod S.Agrawal
[1] Jawaharlal Nehru Engineering College, MGM University
Email:vagrawal@mgmu.ac.in
Sarika Shelke Kadam
Dr.G.Y.Pathrikar College of Computer Science and IT,MGM University
Email:sarikashelke2032@gmail.com

----------**ABSTRACT**----------

This research paper focuses on the prediction of placement chance for student. This research paper works on students academic records like 10th, 12th, Graduation score, Graduation trade, PG score, work experience etc. In this research paper we are going to predict the placement chance of student using various machine learning models (algorithms) based on the historical academic records and to validate hypotheses regarding significant relationships between these academic records and a student's placement success. To accomplish this,we have used SPSS (Statistical package for Social Sciences) tool for analyzing the dataset. By using SPSS, we assess the statistical significance of each variable in relation to the placement outcome. The results not only validate the hypotheses but also contribute to the understanding of the factors influencing a student's placement chances. However, this endeavor is not without its challenges. obstacle is the potential presence of outliers and anomalies in academic records, which could impact the reliability of the predictive models. Furthermore, the integration of work experience into the predictive framework adds another layer of complexity. Capturing the impact of work experience on placement success requires careful consideration.

Keywords - **Machine learning algorithms, campus placement, Statistical significance, predictive analytics, Data driven decision making.**

Date of Submission: Date of Acceptance: 15 Nov 2023

I. INTRODUCTION

Getting a job after finishing studies is a big deal for students. Imagine if there was a way to know beforehand how likely you are to get a job right after college. That's what we're exploring in this research.Predicting whether a student will get placed in campus drives is important for two main reasons. First, for students, it helps them plan better for their future. Knowing the likelihood of getting a job in campus drive.Second, for colleges and universities, to improve placement sucess percentage.we'll talk about how we did this prediction, what factors we considered, and how these factors affect the placement chance.

II. PROBLEM STATEMENT

This research paper explores following research questions.
How can we predict the placement status based on the previous academic details of student?
Whether Gender affects job placement chance and job salary?
Whether SSC & HSC board affect job placement chance?
Whether PG% affect job salary?
Whether work experience affect job placement.
Whether Gender affect on job salary

III. DATASET USED

Dataset downloaded from Kaggle having 119 rows and 15 columns

Serial Number	Gender	Secondary Education Percentage	Board of Education (Secondary)	Higher Secondary Education Percentage	Board of Education (Higher Secondary)	Higher Secondary Education Specialization	Degree Percentage	Degree Specialization	Work Experience	Employability Test Percentage	MBA Specialization	MBA Percentage	Status	Salary Offered By Corporate
1	M	67	Others	91	Others	ommerce	58	Sci&Tech	No	55	Mkt&HR	58.8	Placed	270000
2	M	79.33	Central	78.33	Others	Science	77.48	Sci&Tech	Yes	86.5	Mkt&Fin	66.28	Placed	200000
3	M	65	Central	68	Central	Arts	64	m&Mgmt	No	75	Mkt&Fin	57.8	Placed	250000
4	M	56	Central	52	Central	Science	52	Sci&Tech	No	66	Mkt&HR	59.43	ot Placed	
5	M	85.8	Central	73.6	Central	ommerce	73.3	m&Mgmt	No	96.8	Mkt&Fin	55.5	Placed	425000
6	M	55	Others	49.8	Others	Science	67.25	Sci&Tech	Yes	55	Mkt&Fin	51.58	ot Placed	
7	F	46	Others	49.2	Others	ommerce	79	m&Mgmt	No	74.28	Mkt&Fin	53.29	ot Placed	
8	M	82	Central	64	Central	Science	66	Sci&Tech	Yes	67	Mkt&Fin	62.14	Placed	252000
9	M	73	Central	79	Central	ommerce	72	m&Mgmt	No	91.34	Mkt&Fin	61.29	Placed	231000
10	M	58	Central	70	Central	ommerce	61	m&Mgmt	No	54	Mkt&Fin	52.21	ot Placed	
11	M	58	Central	61	Central	ommerce	60	m&Mgmt	Yes	62	Mkt&HR	60.85	Placed	260000
12	M	69.6	Central	68.4	Central	ommerce	78.3	m&Mgmt	Yes	60	Mkt&Fin	63.7	Placed	250000
13	F	47	Central	55	Others	Science	65	m&Mgmt	No	62	Mkt&HR	65.04	ot Placed	
14	F	77	Central	87	Central	ommerce	59	m&Mgmt	No	68	Mkt&Fin	68.63	Placed	218000
15	M	62	Central	47	Central	ommerce	50	m&Mgmt	No	76	Mkt&HR	54.96	ot Placed	
16	F	65	Central	75	Central	ommerce	69	m&Mgmt	Yes	72	Mkt&Fin	64.66	Placed	200000
17	M	63	Central	66.2	Central	ommerce	65.6	m&Mgmt	Yes	60	Mkt&Fin	62.54	Placed	300000

IV. OBJECTIVES

To predict the placement of student before completion of the degree course

Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

To check whether Gender affects job placement chance.
To check whether SSC & HSC board affect job placement chance.
To check whether PG% affect job salary.
To check whether work experience affect job placement.
To check whether Gender affect on job salary.

V. LITERATURE REVIEW

1. Different machine learning algorithms used to find out trend of placement like decision tree, Artificial neural network,Support vector machine etc. They used deep neural network classifiers with 1000, 2000, 5000 iterations with 71%, 77% and 91% of accuracy[1].

2.In this research paper hypothesis test performed and proved the hypothesis like to check impact of gender for placement chance, impact of PG CGPA on placement, impact of PG specialization on placement, impact of UG CGPA on placement, impact of soft skill competency on placement. Based on the confusion matrix accuracy observed[2].

3.Different 5 machine learning classification models used in this study. The accuracy of Decision tree algorithm found as 72.83%[3].

4.Different algorithms for the prediction of placement of final year B.Tech students. SVM algorithm accuracy observed as 100% and logistic regression as 97.59%[4].

5.Decision Tree machine learning algorithm used to predict the placement chance of student.100% accuracy found in this algorithm[5].

6.Algorithms used to check the accuracy of prediction of campus placement of student. Logistic regression and Decision tree classifiers accuracy observed as 83%[6].

Merits & Demerits in existing work:

Merits: Machine learning algorithms accuracy calculated and presented analysis of algorithms.

Demerits: Demographic factors could not seen for validity of hypotheses

VI. HYPOTHESES FORMULATION

H1: We can predict the placement of student before completion of degree course
H2: Gender does not affect on job placement chance
H3: SSC & HSC board affect job placement chance
H4: PG% does not affect job salary
H5: Work experience does affect job placement
H6: Gender does not affect on job salary.

VII. RESEARCH METHOD

a. Introduction to Methods:

Overview:
This study employs a mixed-methods approach to investigate the relationship between students' academic records and their success in campus placements.

Objective:
The primary objective is to develop a predictive model using machine learning algorithms and validate the model's outcomes through statistical analysis.

b.Population:
MBA students Alumni

c. Data Collection:
Variables:

Dependent Variable: Placement Success (Binary: Yes/No)
Independent Variables: 10th%, 12th%, Grad%,work experience,mba specialization,degree %, mba%,10th board, 12th board, 12th specialization, Degree Specialization etc

d. Data Analysis:
Statistical Methods:Chi-square test for assessing the association between categorical variables.
Software:
SPSS for statistical analysis, scikit-learn library for machine learning.

e. Procedure for Machine Learning:
Algorithms: 08 different machine learning algorithms used. (Refer Fig. 2)

Training and Testing:
80% of the dataset used for training, 20% for testing.

f. Procedure for Statistical Analysis:
Tests Used: Chi-square

g.Significance Level:
$\alpha = 0.05$
8. Data Preprocessing:
h.Cleaning: Removal of outliers Handling missing data
Normalization/Scaling:

Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

In case of hypothesis-1,

H1: We can predict the placement of student before completion of degree course
Null hypothesis: we can not predict the placement of student before completion of degree course.
Alternative hypothesis: we can not predict the placement of student before completion of degree course.

```
Decision Tree Accuracy: 0.75
Random Forest Accuracy: 0.7916666666666666
Logistic Regression Accuracy: 0.8333333333333
Naive Bayes Accuracy: 0.8333333333333334
KNN Accuracy: 0.75
SVM Accuracy: 0.8333333333333334
XGBoost Accuracy: 0.7916666666666666
AdaBoost Accuracy: 0.75
```

Fig. 2. Prediction result using 8 different machine learning algorithms

Based on the above result, Various algorithms applied on the same dataset giving different accuracy. Random forest, Naive byes and Support Vector Machine giving highest accuracy. From above it is clear that we can predict the placement of student before completion of degree course.

In case of hypothesis-2,

H2: Gender does not affect on job placement chance
Null hypothesis: Gender and campus placement status have no significant difference.
Alternative hypothesis: Gender and campus placement status have significant difference.
In above case both the parameters are of Nominal so statistical test chi-square can be applied.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.965[a]	1	.326		
Continuity Correction[b]	.603	1	.438		
Likelihood Ratio	.955	1	.329		
Fisher's Exact Test				.411	.218
N of Valid Cases	119				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.61.
b. Computed only for a 2x2 table

Fig. 3. Chi-square test result for hypothesis -2

From the above table, it is found that pearson chi square is 0.326 > 0.05 means Gender and campus placement status have no significant difference.

In case of hypothesis-3,
H3: SSC & HSC board affect job placement chance
Null hypothesis: SSC, HSC board and campus placement status have no significant difference.
Alternative hypothesis: SSC, HSC board and campus placement status have significant difference.
In above case both the parameters are of Nominal so statistical test chi-square can be applied.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.350[a]	1	.245		
Continuity Correction[b]	.921	1	.337		
Likelihood Ratio	1.342	1	.247		
Fisher's Exact Test				.313	.169
N of Valid Cases	119				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 15.13.
b. Computed only for a 2x2 table

Fig. 4. Chi-square test result for hypothesis -3

From the above table, it is found that pearson chi square is 0.245 > 0.05 which means SSC, HSC board and campus placement status have no significant difference.

In case of hypothesis-4,
H4: PG% does not affect job salary.
Null hypothesis: PG% and campus placement status have no significant difference.
Alternative hypothesis: PG% and campus placement status have significant difference.
In above case both the parameters are of Nominal so statistical test chi-square can be applied.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	33.153[a]	25	.127
Likelihood Ratio	38.272	25	.044
N of Valid Cases	119		

a. 44 cells (84.6%) have expected count less than 5. The minimum expected count is .30.

Fig. 5. Chi-square test result for hypothesis -4

From the above table, it is found that pearson chi square is 0.127 > 0.05 which means PG% and campus placement status have no significant difference.

In case of hypothesis-5,
H5: Work experience does affect job placement
Null hypothesis: Work experience and campus placement status have no significant difference.
Alternative hypothesis: Work experience and campus placement status have significant difference.
In above case both the parameters are of Nominal so statistical test chi-square can be applied

Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.452[a]	1	.020		
Continuity Correction[b]	4.469	1	.035		
Likelihood Ratio	5.962	1	.015		
Fisher's Exact Test				.026	.015
N of Valid Cases	119				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.29.

b. Computed only for a 2x2 table

Fig. 6. Chi-square test result for hypothesis -5

From the above table, it is found that pearson chi square is 0.020 < 0.05 which proves the alternative hypothesis that there is significant relationship between work experience and campus placement status.

In case of hypothesis-6,
H6: Gender does not affect on job salary.
Null hypothesis: Gender and job salary have no significant difference.
Alternative hypothesis: Gender and job salary have significant difference.

Independent sample t-test can be applied in above case.
Both the parameters are of Nominal so statistical test chi-square can be applied

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	36.712[a]	33	.301
Likelihood Ratio	46.673	33	.058
N of Valid Cases	83		

a. 66 cells (97.1%) have expected count less than 5. The minimum expected count is .35.

Fig. 7. Chi-square test result for hypothesis -6

From the above table, it is found that pearson chi square is 0.301 > 0.05 which means Gender and job salary have no significant difference.

VIII. DATA ANALYSIS & INTERPRETATION

According to the results being displayed in above Figures 2 to 7, following conclusions can be made based on chi square tests.

- We can predict the placement of student before completion of the degree course
- Gender does not affect on job placement chance
- SSC & HSC board does not affect job placement chance
- PG% does not affect job salary
- Certification does not affect job placement chance.
- Gender does not affect on job salary.

Result Analysis: The results of our predictive model demonstrated a high degree of accuracy in forecasting students' placement chances. Machine learning algorithms effectively captured complex patterns within the dataset, providing valuable insights into the factors influencing placement outcomes.
Conclusion: our study successfully addressed the research questions. The integration of machine learning algorithms and statistical analyses contributed to a comprehensive understanding of the prediction and significant relationships between factors considered in the above specified hypotheses.

Overall Evaluation:

- Methodology: The methodology effectively blended machine learning and statistical analysis, offering a comprehensive approach to predicting placement chances.
- Applicability: The practical applicability of our results is significant. Implementing the predictive model in real-world scenarios could help student and institute.

REFERENCES

[1]Chandrasekhar Kumbhar 1 , Dr. S. S. Sridhar: Trend analysis of university placement by using machine learning algorithms,International Journal of Engineering & Technology, pp. 178–181 (2018)

[2]D. Satish Kumar, Zailan Bin Siri, D.S. Rao, S. Anusha: Predicting Student's Campus Placement Probability using Binary Logistic Regression,International Journal of Innovative Technology and Exploring Engineering (IJITEE), pp.2278-3075 (2019)

[3]Athreya Shetty B, Akram Pasha, Amith Singh S, Shreyas N I, Adithya R Hande: Comparative Study of Multiple Machine Learning Algorithms for Students, International Journal of Engineering and Advanced Technology (IJEAT) pp. 182-187 (2019)

[4]Irene Treesa Jose , Daibin Raju , Jeebu Abraham Aniyankunju , Joel James , Mereen Thomas Vadakkel: Placement Prediction using Various Machine Learning Models and their Efficiency Comparison,International Journal of Innovative Science and Research Technology pp. 2456-2165 (2020)

[5] Vinod S. Agrawal, Sarika Shelke Kadam: Predictive analyis of placement of students using machine learning algorithmns:Decision Tree, Journal of data mining and management, pp. 1-5 (2022)

[6]Vinod S. Agrawal, Sarika Shelke Kadam: Review on Predictive Analysis of Placement of Students Using Machine Learning Algorithms, Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics pp. 23–30 (2023)