iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** III **Month of publication:** March 2023

DOI: https://doi.org/10.22214/ijraset.2023.49744

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

Book Cave: A Bookstore for Everyone

Jatin Goel[1], Harshita Jain[2], Khushi Rastogi[3], Dimple Sharma[4]
Inderprastha Engineering College

Abstract: ***India as we know is a densely populated country and Every year more than 6 crores of Indians graduates from diverse backgrounds and with diversity in education. Almost similar number of students enter into colleges for taking various education to help them in seeking jobs. Many sectors have experienced tremendous growth in employment and thus masses opt for those sectors whereas in many sectors there is huge unemployment either due to low jobs availability or demand of skilled workers is required.***

Thinking of the each and every branch and when comparing it with the current employment in India and abroad, we will definitely find some points that will help in predicting the admissions and jobs scenarios in the fields of engineering and technology, management and pharmacy. Due to the changing technology and its requirement for getting employed in India and abroad, there has to improvements suggested by experts for predicting the Prediction of Admission & Jobs in Engineering & Technology /Management/Pharmacy with respect to demographic locations. This is not a one time process and needs to be done frequently as trends in the industry keep changing. Addressing this problem will introduce the required changes that would bring the current youth and upcoming generations in parallel with the students of other countries in terms of knowledge and skills in that domain. There is a need to forecast the current trend in the admissions and job sectors so as to blend the courses and syllabus accordingly to keep the youth employed and skilled with rapidly changing world. Here we will achieve it by using MaAhine Learning algorithm.

I. INTRODUCTION

A. Motivation

India, as we all know, is a highly populated nation, and every year, more than 6 crore Indians graduate from a variety of educational backgrounds and socioeconomic backgrounds. A nearly equal percentage of students enroll in institutions to pursue various degrees that will aid in their career search. In comparison to ten years ago, the demands of the work market have changed significantly with the advancement of technology. While there is significant unemployment in many industries due to either a lack of jobs or a need for qualified workers, several sectors have seen remarkable development in employment and as a result, many people choose to work in those sectors. When considering each and every branch and contrasting it with the employment trends both in India and abroad, we will undoubtedly discover some factors that will assist in predicting admissions and employment scenarios in the fields of management, pharmacy, and engineering and technology.

B. Problem Statement

Improvements have been suggested by experts for predicting the Prediction of Admission & Jobs in Engineering & Technology, Management, and Pharmacy with respect to demographic locations due to the changing technology and its requirement for employment in India and abroad. Due to the constant shift in industry trends, this is not a one-time process that needs to be completed frequently.

If this problem is solved, the necessary changes will be made, putting the young people of today and the generations to come up with the same knowledge and skills as students in other countries. However, despite AICTE's approval of the proposed courses, graduates of such institutions, engineering divisions, management departments, pharmacy schools, and others continue to be unemployed. There is currently no method for estimating or forecasting the short- and long-term employment potential of any Engineering, Management, Pharmacy, or other course by combining data from various sources and developing a computer program or application. In order to keep young people employed and skilled in a world that is rapidly changing, it is necessary to forecast the current trend in the admissions and employment sectors in order to blend the courses and syllabus accordingly. Educational institutions have always been crucial to a person's personal growth and development in society. There are currently a number of college prediction apps and websites, but using them is somewhat tedious due to the lack of clear information about colleges and the time required to find the best worthy school.

©IJRASET: All Rights are Reserved

C. Objective

The AICTE grants approval for the launch of new courses and the establishment of new technical institutions in the fields of management, architecture, engineering, and technology, among others. The AICTE's (Approval Process Handbook) minimum standards for infrastructure are currently the primary criterion for granting approval for such courses. However, despite AICTE's approval of the proposed courses, graduates of such institutions, engineering divisions, management departments, pharmacy schools, and others continue to be unemployed. Correlating data from various sources and developing a computer program or application using neural networks OR any other programming tools is currently the only method available to estimate or forecast employment potential for any branch(s) or course(s) in engineering, management, pharmacy, etc. on a short-term and long-term basis. The creation of software for estimating and forecasting graduate and postgraduate employment potential in various fields and courses on a short-term and long- term basis by combining data from various sources could be useful for approving infrastructure and other facilities. Current time simulated intelligence and figuring frameworks can be utilized to help simultaneously and any huge advancement in this field is exceptionally valued which thinks about different factors and gauges the affirmations and occupations area situations. In order to keep young people employed and skilled in a world that is rapidly changing, it is necessary to forecast the current trend in the admissions and employment sectors in order to blend the courses and syllabus accordingly. We will use a machine learning algorithm to accomplish this here.

II. LITERATURE SURVEY

A. Existing System

Admissions have been predicted generally by ranking of the college. These rankings will vary yearly according to the pass out batch performance and results of that particular year. Due to this variation every student who are trying to get in these colleges will be in confusion. We have some sites in online for prediction but most of them are either with not good prediction because of algorithm or due to lack of training data. To improve and get the perfect and accurate prediction there is a need of better data set. Due to this variation every student who are trying to get in these colleges will be in confusion so to with good accuracy and correct prediction we trained a model.

B. Proposed System

One can enter their scores in the corresponding fields of the system that has been proposed. The system then evaluates the data entered and generates a list of universities a person could enroll in based on their test results. This is reasonably swift and saves both time and money. We have suggested an innovative approach using machine learning techniques to do this. We have considered not one, but multiple machine learning techniques to increase the accuracy of our model. Neural networks, linear regression, decision trees, and random forests are some of these techniques. The Methods portion of this article will go into further detail on these algorithms. The algorithm with the best key performance indicators is then chosen after comparison of these algorithms.

C. Related Work

The issue of student admittance is crucial for educational institutions. In this study, machine learning algorithms are used to forecast a student's likelihood of admission to a master's degree. Students will benefit from knowing in advance whether they stand a chance of being admitted.

Multiple linear regression, k-nearest neighbor, random forest, and multilayer perceptron are the machine learning models. The Multilayer Perceptron model outperforms other models, according to experiments.

A person can enter their rank in the provided fields of the proposed system. After that, the data that was entered is processed by the system, and an output of the list of colleges that a person could apply to, along with their scores, is produced. This is quick and helps save money as well as time. We have proposed a novel approach that makes use of Machine Learning algorithms to accomplish this. We have taken into account all of the following to ensure that our model is as accurate as possible: but a number of algorithms for machine learning.

Neural Networks, Linear Regression, Decision Tree, and Random Forest are some of these algorithms. The Algorithms section of this paper will discuss these algorithms in greater detail. The Prediction System will be built using the algorithm with the best key performance indicators after these algorithms are compared. We also anticipate incorporating university clustering based on a profile and classifying them as less likely, highly likely, or both.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

III. ALGORITHM FOR ADMISSION PREDICTION

A. Linear Regression

Regression models are used to describe a relation between different variables by using the observed data into a line. Straight lines are used in linear regression models, whereas curved line is used in logistic and non-linear regression models. Linear regression model is a method used as response for only a single feature, it is based on supervised learning. Regression models always target a prediction value which is based on independent variables. It is used to calculate the relationship between two quantitative variables. Regression models differ upon the type of relation between independent and dependent variables, the number of variables being used and the ones they are considering. It is assumed that these variables have a linear relationship. From this point forward, we attempt to define a linear function that accurately predicts the response value(y) as a function of the feature or independent variable (x). cut-offs of the colleges, admission intake and preferences of students. Also, it helps students avoid spending time and money on counsellor and stressful research related to finding a suitable college.

B. Decision Trees

Classification is a two step process, learning and prediction. At first, model is developed based upon the given training data in its learning step. Then the model is used to predict response for the given data in prediction step. One of the most popular classification algorithms and easiest to learn and understand in decision tree. Decision tree algorithm is also used for solving classification and regression problems. Decision trees use a class label for predicting, for a record it starts from the root of tree. Then compare the values of the root with its record attribute. After the comparison, it follows the branch which is corresponding to the value and jump upon to the next node. There are two types of decision trees, Categorical variable and continuous variable. Categorical variable has a categorical target variable and continuous variable has a continuous target variable. Decision tree has three types of nodes, decision nodes, chance nodes and end nodes. Decision tree assigns a class label for each leaf node. Even the non-terminal nodes, the root and internal nodes, also contain attribute test conditions to separate records that have different characteristics.

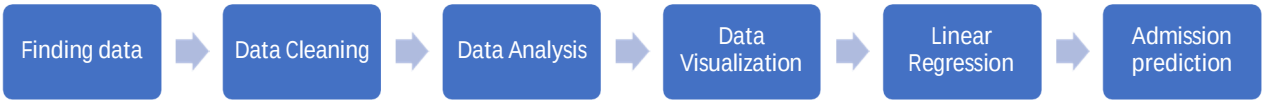
C. Random Forests

The random forest is a machine learning algorithm which is widely used in regression and classification problems. Decision trees are built upon multiple different samples and then take their majority vote for average and bifurcation in case of regression. Random forest has the ability to handle a data set which contains continuous variables in case of regression and categorical variables in case of classification. Hence, it provides good results for classification problems. In industry lingo, reason behind forest works algorithm works so well is: Any huge quantity of moderately uncorrelated trees working as a body will outperform any of the individual constituent models.

IV. METHODOLOGY AND IMPLEMENTATION

The primitive step to building a model for our use case is choosing the right dataset. For our predictions, we chose a dataset which contains all the important attributes that would affect the chances of admit. This is followed by data cleaning where we handle missing values present in various fields.

Once the data is ready to be analysed, we use various tools and libraries to visualize the data and perform analysis. This includes visualizing bar graphs and the correlation matrix. Once the data is ready to be processed, we split it into training and testing data. For this, we will be using 3 machine learning algorithms; linear regression, random forest and neural network. Once these models are built over the dataset, we compare them using key performance indicators. These indicators help us choose the right model for predicting whether an applicant has chances of admission.



Block Diagram

©IJRASET: All Rights are Reserved

V. DATASET OF ADMISSION PREDICTION

The data set comprises of different factors attributed towards picking the right university. It contains data of 1000 different students. Data set is classified into 6 different parameters which are considered important during the application for Engineering.
Those parameters are: Year, 10th Marks, 12th Marks, 12th Division, AIEEE Rank, College.

	Year	10th Marks	12th Marks	12th Division	AIEEE Rank	College
0	2015	95	92	2	100	IIT delhi
1	2015	75	88	3	1023	VIT vellore
2	2015	83	84	6	2935	Ahemedabad IT
3	2015	75	91	8	5647	University college of ENGG
4	2015	94	94	9	3564	SRMIST chennai

Dataset of Admission Prediction

VI. OUTPUT OF ADMISSION PREDICTION

A. *Linear Regression Output Graph*

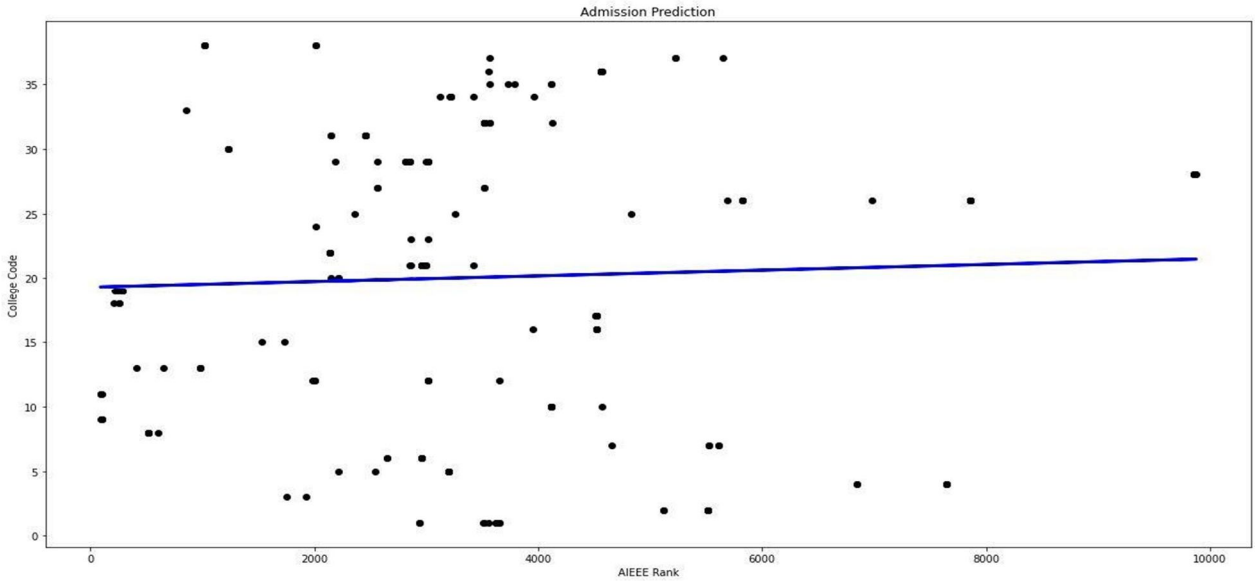


Figure 7.1.1 Linear Regression Output Graph

B. *Admission Prediction Output*

```
col=df.columns.tolist()[4:5]
print(col)
usrip=[]
for i in col:
    print("==================================================")
    usrip.append(eval(input(i+": ")))

userpreddt=model.predict([usrip])
print("You may have change to get entrance in: ",college[clg_code.index(int(userpreddt[0]))])
```

```
['AIEEE Rank']
==================================================
AIEEE Rank: 57
You may have change to get entrance in:  IIT kharagpur
```

Output

Here in the output 1 the admission is predicted in the IIT Kharagpur college for the AIEEE Rank of 57.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

C. Job Prediction With Respect To Demographic Location

Here we study the trends and see the availability of jobs at a particular location or state. Here classification is done based on the branches/streams in a state. Based on the previous data we analyze the cumulative salary over the years.

The predictions made are shown in the form of graphs. Due to unavailability of data, the graphs are the static based on predictions from previous data.

VII. ALGORITHMS

D. LSTM Algorithm

Long Short Term Memory networks are referred to as the LSTM Network model. These particular neural networks are capable of comprehending long-term dependencies in general. Long-term dependencies were generally intended to cause problems, and LSTM models successfully address these issues in the majority of cases.

E. Min-Max Scaler

Another approach to data scaling sets the minimum feature value at zero and the maximum feature value at one. The MinMax Scaler reduces the data inside the specified range, often between 0 and 1. By scaling features to a predetermined range, it changes data. It scales the values to a particular value range while preserving the original distribution's shape.

F. Tensor-flow

An open-source software library is TensorFlow. In its simplest form, TensorFlow is a software library for numerical computation utilising data flow graphs, where nodes in the network stand in for mathematical operations. edges in the graph represent the multidimensional data arrays (called tensors) communicated between them. (Please note that tensor is the central unit of data in TensorFlow).

VIII. DATASET

The data set comprises of different factors attributed towards picking the right university. It contains data of 164 different data.
Data set is classified into 4 different parameters which are considered important during the application for Engineering.

Those parameters are: College State, Specialization, Graduation Year, Salary

	CollegeState	Specialization	GraduationYear	Salary
0	Delhi	instrumentation and control engineering	2013	445000
1	Delhi	computer science & engineering	2013	420000
2	Delhi	computer science & engineering	2013	200000
3	Delhi	computer science & engineering	2013	375000
4	Delhi	computer science & engineering	2013	350000

Figure 10.1 Dataset of Job Prediction

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IX. OUTPUT OF JOB PREDICTION

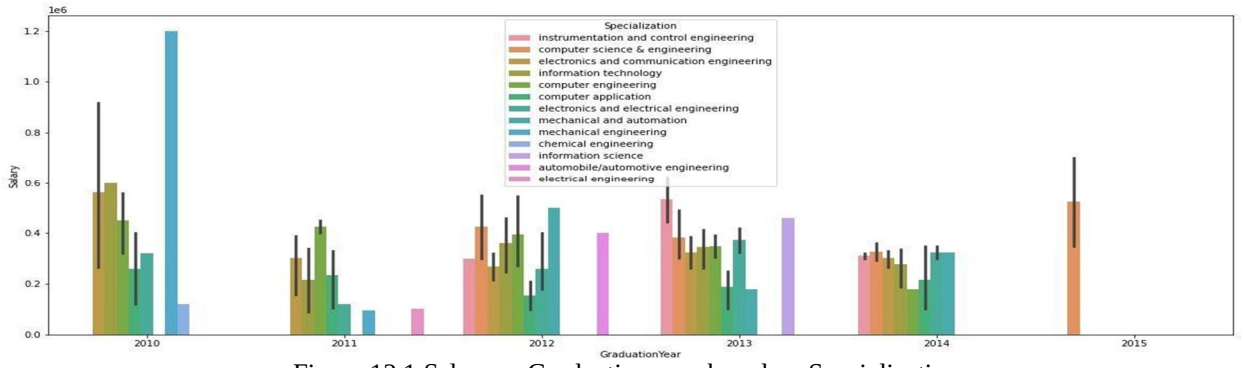


Figure 12.1 Salary vs Graduation year based on Specialization

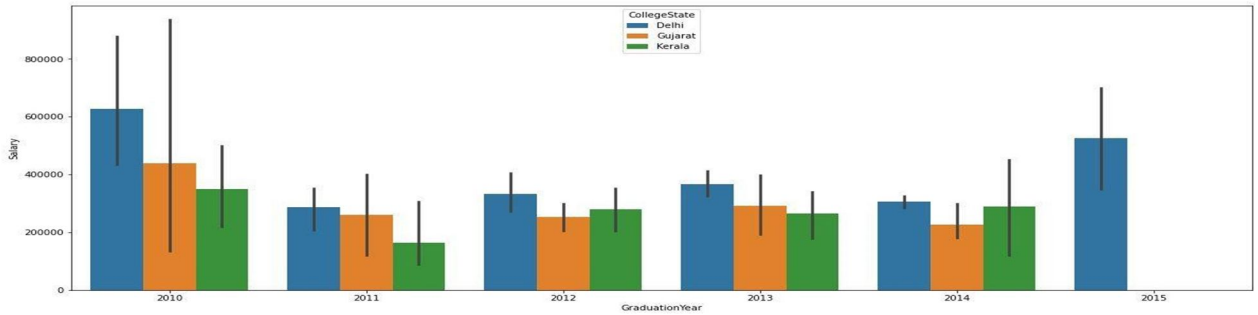


Figure 12.2 Area wise Salary vs Graduation year

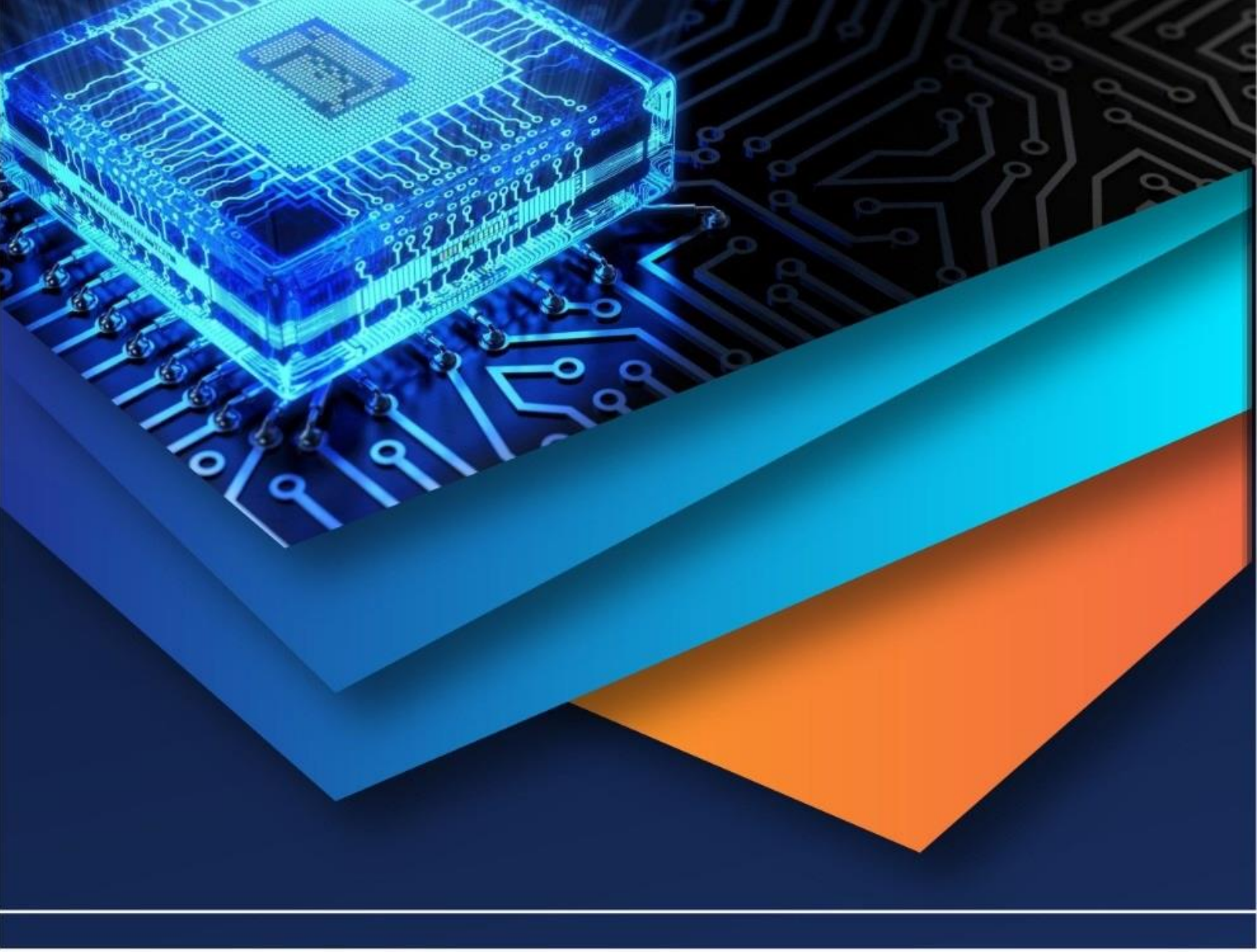
X. CONCLUSION

Every year millions of students apply to universities to begin their educational life. Most of them don't have proper resources, prior knowledge and are not cautious, which in turn creates a lot of problems as applying to the wrong university/college, which further wastes their time, money and energy. With the help of our project, we have tried to help out such students who are finding difficulty in finding the right university for them. It is very important that a candidate should apply to colleges that he/she has a good chance of getting into, instead of applying to colleges that they may never get into. This will help in reduction of cost as students will be applying to only those universities that they are highly likely to get into our prepared models work to a satisfactory level of accuracy and may be of great assistance to such people. This is a project with good future scope, especially for students of our age group who want to pursue their higher education in their dream college. Results show us that the highest accuracy is achieved through the linear regression model.

REFERENCES

[1] Subba Reddy.Y and Prof. P. Govindarajulu," A survey on data mining and machine learning techniques for internet voting and product/service selection", IJCSNS International Journal of Computer Science and Network Security, VOL.17 No.9 September 2017.
[2] Zhibo Wang, Jilong Liao, Qing Cao, Hairong Qi, and Zhi Wang, "Friend book:A Semanticbased Friend Recommendation System for Social Networks IEEE Transactions on Mobile Computing.
[3] J. Bobadilla et al. "Knowledge-Based System" Elsevier B.V.
[4] Hector Nunez, Miquel sanchez-Marre, Ulises Cortes, Joaquim Comas, Montse Martinez, Ignasi RodriguezRoda, Manel Poch, "A Comaprative study on the use of similarity measure in case based reasoning to improve the classification of environmental system situations,", ELSEVIER, Environmental Modeling and Software (2003)
[5] DINO IENCO, RUGGERO G. PENSA and ROSA MEO, "From Context to Distance: Learning Dissimilarity for categorical Data Clustering," Journal Vol. X. 10 2009, pages 1- 10

Group Members:
N.Indrasena Reddy , 19311A04X3
K.Sumanth , 19311A04Y2
K.Vijay Bhargav , 19311A04AV



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)