



## College Admission Prediction Using Machine Learning

<sup>1</sup>Anmol Pawar, <sup>2</sup>Rushikesh Patil, <sup>3</sup>Kadayya Mathapati, <sup>4</sup> Pratik Lonare, <sup>5</sup>Prof.Laximikant Malphedwar

<sup>1</sup>Student, <sup>2</sup> Student, <sup>3</sup> Student, <sup>4</sup> Student, <sup>5</sup>Professor  
Department of computer engineering Dr. D Y Patil College of engineering  
and innovation Varale, Pune, Maharashtra

**Abstract :** Making a short list of institutions to apply to might be challenging for prospective students. Students frequently question if their profile fits the requirements of a certain institution because applications are so dynamic. Additionally, because applying to college is so expensive, it is crucial that students narrow down their selection of colleges based on their profiles. Students can utilize a university admission prediction technique to figure out their odds of getting into a certain college. Data about prior applicants to other universities and their acceptance or rejection status may be used by the system. Numerous issues plagued the early iterations of these prediction algorithms, including their inability to consider important variables like research experience or GRE (Graduate Record Exam) scores. The decision to attend college is an important one, and with the admissions process becoming more competitive, there is increased interest in using machine learning to forecast admission results. This study provides a thorough review of how machine learning is being used to forecast college admissions. The research talks about the common data sources in this area, such as previous admission statistics, test scores, high school grades, extracurricular activities, recommendation letters, and personal statements. Research is conducted to determine how well various machine learning techniques—such as logistic regression, neural networks, decision trees, and random forests—model and forecast admission outcomes. The research also investigates the difficulties and constraints of applying machine learning for predictions of college acceptance, including problems with data fairness, model fairness, and the intrinsically subjective character of the admissions process. According to the findings, machine learning models should be utilized in the admissions process as a supplementary tool rather than as the main factor, even if they may offer insightful forecasts and important insights. Predictive modeling fairness and ethical issues are emphasized as being essential components of this application.

**Keywords –** College admissions, GRE scores, Selection criteria, Neural networks, Decision trees

### I. INTRODUCTION

When choosing a college, students who want to continue higher education are a constant source of confusion. In addition to the college's rating, students must take into account a number of other considerations. The college's admission rate is one of these variables. Only a very tiny fraction of applicants are accepted into certain highly regarded institutions' programs. These are also schools that receive a ton of applications. However, because the cost of each college's application is so costly, students are discouraged from applying to many of them. Students often only apply to a small number of institutions because of the intense competition for admission and the hefty expense of each application. This is also the most practical choice. Therefore, a prediction system that evaluated students' prospects of enrolling in a college based on their test results and academic background would be of great use to students. A crucial and very competitive stage of any student's educational path is being accepted to college. Educational institutions must choose the most qualified applicants from a wide pool of applicants as the number of college applications rises. The use of machine learning techniques for forecasting college admission results has attracted a lot of attention in this domain. In-depth analysis of machine learning's use in the context of college admission predictions is the primary objective of this study. In order to set tone, this introduction emphasizes the rising significance of data-driven decision-making in education as well as the particular difficulties presented by the college admissions process.[1] Predicting university admission can be accomplished by taking into account student and prior admission data, then generating a prediction based on this understanding. A simpleton a model utilizing simple machine learning techniques in order to tackle this. picking up algorithms. In order to attain an elevated level of accuracy, One option is to use deep learning techniques. Many recommender [2] systems were initially only very fundamental content-based recommender systems, also referred to as query-based information retrieval systems. For example, search engines recommend webpages whose content corresponds with user queries. [4] so many students must enroll in the institution each year, this decision-making process has grown to be a very challenging one. Because this procedure takes into account more than just test results, it also considers the backgrounds of the students and other factors that are weighted according to how well they performed in their postsecondary education. [5] Currently, many institutions project their admission lines primarily using the past admittance rates throughout time and the estimated scores of their students. Naturally, the foundation of this approach is predicated on the idea that academic standards at the schools would not significantly alter and that admission rates would remain constant.[7] This study uses data from Kaggle to discuss the application of three machine learning algorithms to forecast prospective university ratings to

admit the students. In terms of the least error, we evaluate the performance of the three models (the decision tree model, the logistic regression model, and the linear regression model) and provide some suggestions for additional study. [8] This study uses data from Kaggle to discuss the application of three machine learning algorithms to forecast prospective university ratings to admit the students. In terms of the least error, We compare the three models' performances—the decision tree model, the logistic regression model, and the linear regression model—and make some suggestions for additional study.

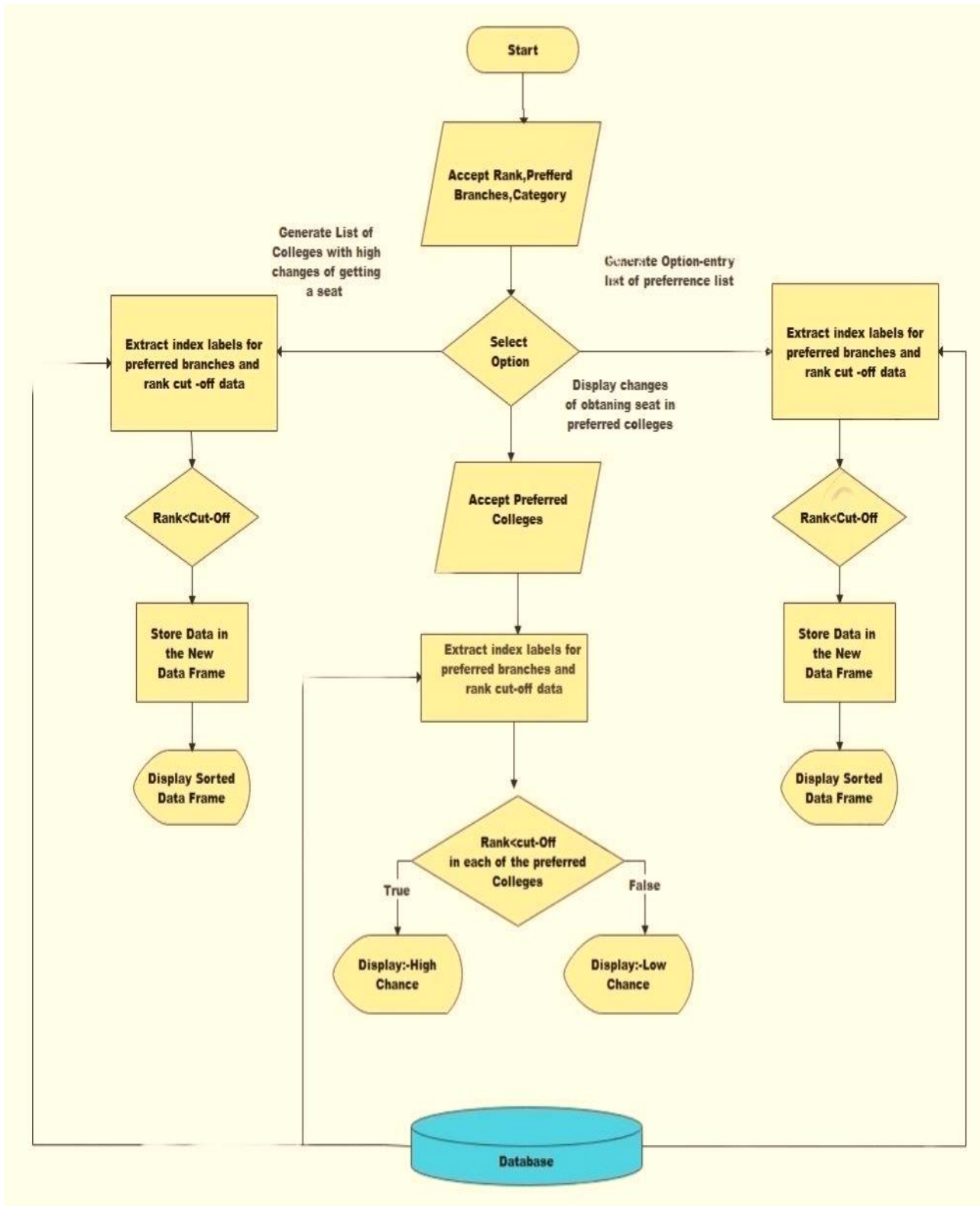
## II. LITERATURE SURVEY:

Technology	Observation	Limitations	Accuracy
Multi linear regression, Random Forest	Effective for graduate admission prediction	Lack of information on dataset and features	MAE 3.37%
Hybrid recommender based on data mining	Good performance in college admission prediction	Limited to hybrid system-specific context	Rule based
Adaboost classifier	Better than random forest	Lack of dataset used	90%
Stack Ensemble Learning	Accurate prediction of university admission	Limited to specific model and dataset	91%

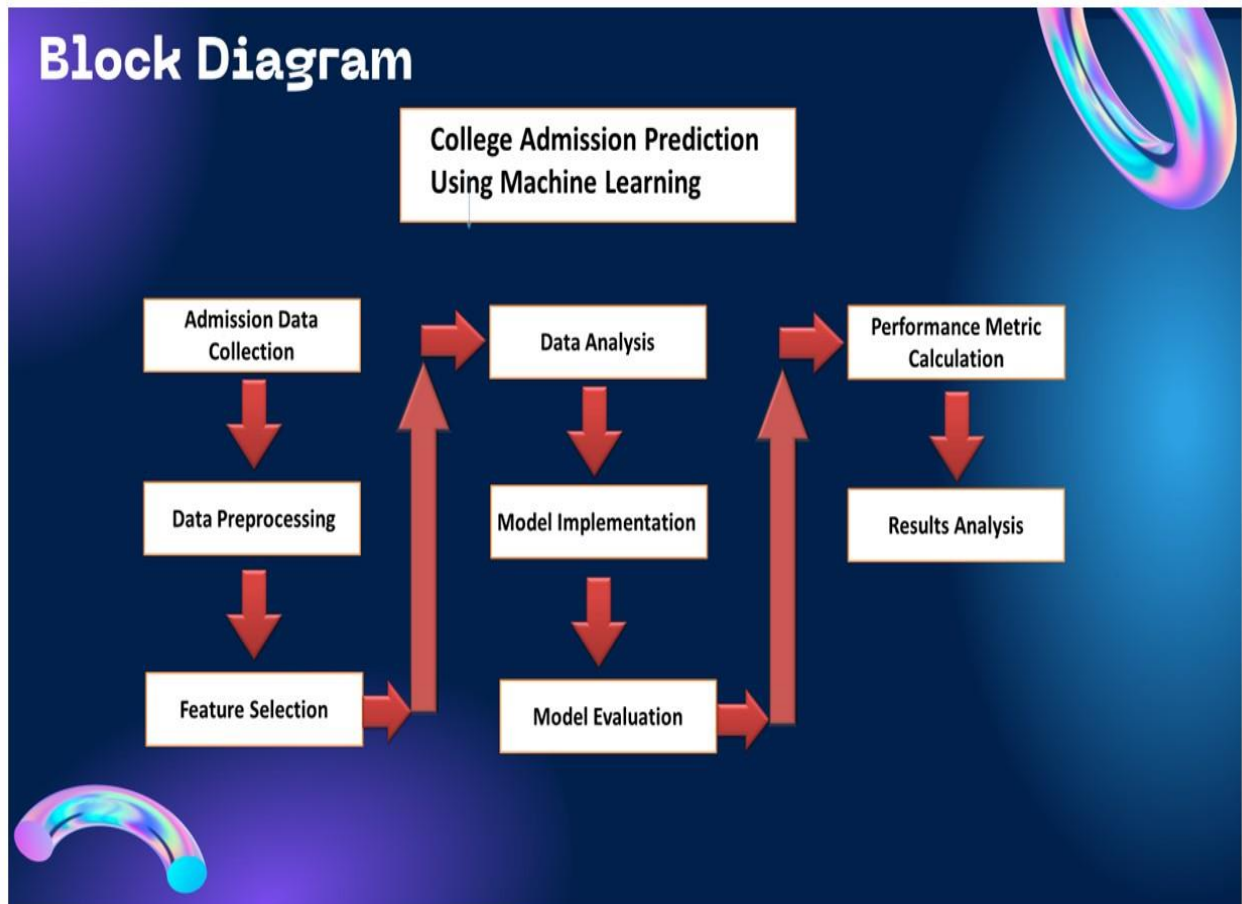
## III. PROPOSED METHODOLOGY

It is We have utilized pandas and numpy, two Python machine learning libraries, in our suggested prediction method. Additionally, we used the streamlit package to develop the web application and user interface (UI). We have also used Heroku to deploy the web application online so that it is reachable from anywhere in the world. The database also includes the average of the rank cutoff data from the preceding five years. The branch, college, and category ranks will be included in the cut-off database. A candidate will receive an approximate notion of the seat that they are most likely to be awarded based on their category and rank. The cutoff will vary depending on the course, college, and category. The row headings have the names of the colleges and their branches. The different categories are listed in the column headings. The database holds data that is of the string data type. The rank a candidate from a given category must obtain to be admitted to a given branch and college is contained in each cell (which corresponds to a branch and college, or row heading and category, or column heading) in the database. Our method uses the user's rank (mandatory), category (mandatory), preferred branches (optional), preferred colleges (optional), and preferred districts/location (optional) as input. The alphanumeric value "category" can be chosen from a drop-down list. Preferred districts, colleges, and branches are text inputs with multiple selection options that can be chosen from a drop-down list. It is mandatory for the user to enter the required fields, rank and category. Algorithms needed: For binary classification tasks, the straightforward and comprehensible algorithm known as logistic regression can be applied. It's a good starting point to understand the relationship between demographic attributes and admission probabilities. Decision Trees: Decision trees can capture nonlinear relationships and interactions between features. They are particularly useful for visualizing how different attributes contribute to admission predictions. Random Forest: An ensemble of decision trees called a random forest can handle noisy data and increase prediction accuracy. It's capable of capturing complex relationships and handling interactions between features. Gradient Boosting: Algorithms like XGBoost or LightGBM can provide accurate predictions by combining the strengths of multiple decision trees while addressing overfitting and improving interpretability.

IV. FLOWCHART



### 3.1 Architecture



### 3.2 Conclusion

The field of college admission prediction becomes increasingly relevant as more individuals choose to earn a degree. In this essay, a useful technique for estimating It has been suggested what a student's chances are of getting into a particular college. Additionally, we evaluated how well the various machine learning methods performed in predicting admissions when compared to the suggested approach. It has been shown that the suggested approach offers the best accuracy and performance.

### 3.3 Future Scope

As we gather more information on students, the model can be enhanced in the future. Information from the applicant's Letters of Recommendation (which can be scored) and Statement of Purpose essay can be used to improve the prediction accuracy. Natural language processing methods can also be used to assess the essays and letters. In addition, past trends may be taken into consideration when adjusting the attribute weighting.

## REFERENCES

- [1] S. Sridhar, S. Mootha and S. Kolagati, "A University Admission Prediction System using Stacked Ensemble Learning," 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), Cochin, India, 2020, pp. 162-167, doi: 10.1109/ACCTHPA49271.2020.9213205.
- [2] Li, Qing & Kim, Byeong. (2003). Clustering approach for hybrid recommender system. Proceedings of the International Conference on Web Intelligence. 33- 38. 10.1109/WI.2003.1241167.
- [3] Aljasmī, Sara & Nassif, Ali & Shahin, Ismail & Elnagar, Ashraf. (2020). Graduate Admission Prediction Using Machine Learning. 14. 10.46300/91013.2020.14.13.
- [4] A. H. M. Ragab, A. F. S. Mashat and A. M. Khedra, "HRSPCA: Hybrid recommender system for predicting college admission," 2012 12th International Conference on Intelligent Systems Design and Applications (ISDA), Kochi, India, 2012, pp. 107-113, doi: 10.1109/ISDA.2012.6416521.
- [5] Zhenru Wang and Yijie Shi, "Prediction of the admission lines of college entrance examination based on machine learning," 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 2016, pp. 332-335, doi: 10.1109/CompComm.2016.7924718.
- [6] R. Sanchez-Salmeron, J.L. Gomez-Urquiza, L. Albendin-Garcia, M. Correa-Rodriguez, M.B. Martos-Cabrera, A. Velando-Soriano, et al., "Machine learning methods applied to triage in emergency services: A systematic review", International Emergency Nursing, vol. 60, pp. 101109, 2022.
- [7] Amal AlGhamdi, Amal Barsheed, Hanadi AlMshjary, and Hanan AlGhamdi. 2020. A Machine Learning Approach for Graduate Admission Prediction. In Proceedings of the 2020 2nd International Conference on Image, Video and Signal Processing (IVSP '20). Association for Computing Machinery, New York, NY, USA, 155–158. <https://doi.org/10.1145/3388818.3393716>
- [8] A. Sivasangari, V. Shivani, Y. Bindhu, D. Deepa and R. Vignesh, "Prediction Probability of Getting an Admission into a University using Machine Learning," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 1706-1709, doi: 10.1109/ICCMC51019.2021.9418279.
- [9] I. Sengupta, C. Koner, N. K. Bhattacharjee and S. Gupta, "Automated Student Merit Prediction using Machine Learning," 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), Sonbhadra, India, 2022, pp. 556-560, doi: 10.1109/AIC55036.2022.9848976.
- [10] Devarapalli, D.J. (2021). Classification Method to Predict Chances of Students' Admission in a Particular College. In: Gunjan, V.K., Zurada, J.M. (eds) Proceedings of International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications. Advances in Intelligent Systems and Computing, vol 1245. Springer, Singapore. [https://doi.org/10.1007/978-981-15-7234-0\\_19](https://doi.org/10.1007/978-981-15-7234-0_19)