**Title of Project:**

**“"Predictive Modeling and Analysis of Graduate School Admissions Using Logistic Regression”**

Mini project report on

**Data Science**

By

1. Anish Nimbal
2. Prathmesh Nikam
3. Umesh Nemane
4. Suhail Pathan
5. Mudassar Sayyed

## ABSTRACT

This study focuses on predicting graduate school admission outcomes using logistic regression and analysing the model's performance through a classification report. The dataset comprises various parameters such as test scores, GPA, and research experience, with the target variable indicating whether an applicant's chance of admission exceeds a predefined threshold. The methodology involves splitting the dataset into training and testing sets, training a logistic regression model on the training data, and evaluating its performance on the test set. The classification report provides insights into the model's precision, recall, F1-score, and support for predicting admission outcomes. This analysis sheds light on the effectiveness of logistic regression in predicting graduate school admission and identifies areas for potential improvement in the model's predictive accuracy.

Top of Form

# Table of Content

|  |  |  |
| --- | --- | --- |
| **Sr no.** | **Content** |  |
| **1** | Introduction |
| **2** | Requirement |
| **3** | Proposed Work |
| **4** | Literature Review |
| **5** | Architecture |
| **6** | Conclusion |
| **7** | **Reference** |

## INTRODUCTION

The admission process for graduate school programs is a critical aspect of higher education, with universities and institutions carefully evaluating applicants based on various parameters. In recent years, data-driven approaches, particularly machine learning techniques, have gained prominence in aiding decision-making processes, including admission selection. Logistic regression, a fundamental algorithm in classification tasks, offers a powerful tool for predicting binary outcomes, making it suitable for modeling admission probabilities.

In this study, we delve into the realm of graduate school admission prediction using logistic regression and analyze the model's performance through a comprehensive classification report. Our aim is to explore the predictive capabilities of logistic regression in determining an applicant's likelihood of admission and to provide insights into the strengths and limitations of the model.

The dataset utilized in this study comprises diverse parameters typically considered during the graduate school admission process, such as standardized test scores, undergraduate GPA, statement of purpose, letters of recommendation, and research experience. The target variable represents whether an applicant's chance of admission surpasses a predefined threshold, allowing us to frame the problem as a binary classification task

**TECHNIQUES REQUIREMENT**

**—-------------------------------------------------------------------**

### Library Requirement

1. Pandas
2. NumPy
3. sklearn

### Model Requirement

1. Logistic Regression
2. Label Encoding

**MOTIVATION**

#### ----------------------------------------------------------------------------------

The motivation behind this study stems from the increasing importance of data-driven decision-making in the domain of graduate school admissions. As competition for admission to prestigious programs intensifies and the number of applicants continues to rise, there is a growing need for efficient and effective methods to assess and select candidates.

Traditional methods of evaluating applicants, while informative, may be subjective and prone to biases. By contrast, machine learning techniques offer a systematic approach to analyzing large volumes of data and extracting valuable insights to aid in decision-making processes.

Logistic regression, a widely-used classification algorithm, provides a transparent and interpretable framework for predicting binary outcomes, making it well-suited for modeling admission probabilities. By leveraging logistic regression, we can harness the power of predictive analytics to assess an applicant's likelihood of admission based on a diverse range of factors, including academic credentials, test scores, and extracurricular achievements.

## PROPOSED WORK

**-------------------------------------------------------------------------**

The proposed work aims to explore the effectiveness of logistic regression in predicting graduate school admission outcomes based on various parameters such as test scores, GPA, and research experience. The project will begin with a comprehensive review of existing literature on predicting admission outcomes in higher education, focusing on studies utilizing logistic regression and other relevant techniques.

Subsequently, a detailed description of the dataset used in the project will be provided, including an overview of features and the target variable, along with preprocessing steps applied to the data. The methodology section will outline the process of splitting the dataset into training and testing sets, training the logistic regression model on the training data, and evaluating its performance on the test set.

Results obtained from the logistic regression model will be presented, with a focus on classification report metrics such as precision, recall, F1-score, and support. These results will be thoroughly discussed to interpret the effectiveness of logistic regression in predicting graduate school admission outcomes, as well as to identify potential areas for improvement in the model's predictive accuracy.

The project will conclude with a summary of key findings, emphasizing the significance of the study and its contributions to the field. Furthermore, the report will outline potential future research directions and provide a comprehensive list of references cited throughout the project.

**Literature Review**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Publisher | Author | Year | Paper Name | Objective |
| Research Gate | **1.Sara Aljasmi**  **2.Ashraf M Elnagar** | **2020** | Graduate Admission Prediction Using Machine Learning | **This will assist students to know in advance if they have a chance to get accepted. The machine learning models are multiple linear regression, k-nearest neighbor, random forest, and Multilayer Perceptron.** |
| JETIR | **1.Kadayya Mathapati, 2. Pratik Lonare, 3. Prof.Laximikant Malphedwar** | **2023** | College Admission Prediction Using Machine Learning |  |
| Research Gate | **1.Amal AlGhamdi**  **2.Hanan AlGhamdi** | **2020** | A Machine Learning Approach for Graduate Admission Prediction | **The evaluation compares three regression models—linear regression, decision tree, and logistic regression—to determine the most accurate predictor for future applicant admissions.** |

**Architecture**

**-------------------------------------------------------------------------**

## 

## CONCLUSION

In conclusion, this study underscores the potential of logistic regression as a valuable tool for predicting graduate school admission outcomes. By combining predictive modeling with comprehensive analysis and visualization, we can gain valuable insights into the admission process and inform decision-making strategies for stakeholders in the higher education sector. Further research and refinement of predictive models are warranted to enhance the accuracy and reliability of admission predictions and promote equitable access to educational opportunities.

#### REFERENCE

**------------------------------------------------------------------------------------**

* https://www.kaggle.com/datasets/tanmoyie/us-graduate-schools-admission-parameters?rvi=
* https://www.geeksforgeeks.org/data-preprocessing-in-data-mining/
* <https://www.javatpoint.com/data-preprocessing-machine-learning>
* ChatGPT