A black and white logo

Description automatically generated with low confidence

Data-Driven Framework for Enhancing Student Applications, Acceptances, and Registrations at the University of Buckingham

Mr. Chidera Kenechukwu Onwumbiko

Supervised by: Prof. Harin Sellahewa

Project Report submitted for the degree of

MSc Applied Data Science

to the School of Computing in the University of Buckingham

December 2024

**ABSTRACT**

The goal of this research is to develop a comprehensive data-driven framework to enhance the student application, acceptance, and registration processes at the University of Buckingham. By leveraging modern data science tools such as Python for data preprocessing and Tableau for visualisation and dashboard building, the framework seeks to provide insights that help better understand the university’s student enrolment. The project aims to address challenges faced by the university, including increased competition and fluctuating student numbers, by utilising historical data to predict and optimise enrolment outcomes.

Data collected includes attributes such as school/department, level, date, number of applications, number of acceptances, etc., which were thoroughly preprocessed and analysed using advanced techniques like data normalisation, unpivoting, and splitting of date formats. A predictive model was built to estimate the number of students that need to be accepted to get a target number of student registrations, using a statistical method where the predicted student acceptances are calculated based on the target student registrations divided by the probability of student registrations. This model was implemented in Tableau using parameters and calculated fields, allowing for dynamic adjustment and scenario testing.

The model's performance was evaluated using **Root Mean Square Error (RMSE)** and **R score**, ensuring that it was both accurate and reliable for decision-making purposes. RMSE provided a measure of the model's predictive accuracy, while the R score indicated the goodness of fit for the predictions. Furthermore, the interactive visualisations created with Tableau help university administrators and stakeholders easily interpret complex trends in student enrolment. This dashboard provides insights into student applications, acceptances, and registrations.

By integrating data-driven decision support with predictive modelling, this framework serves as a crucial tool for the University of Buckingham to navigate the increasingly competitive landscape of higher education enrolment. The insights gained can help develop targeted recruitment campaigns, enhance communication strategies, and improve the overall efficiency of the admission processes.

Key contributions of this research include the development of a predictive model, a deeper understanding of enrolment dynamics through exploratory data analysis (EDA), and the creation of intuitive visualisations that support strategic decisions. The proposed data-driven approach ultimately aims to contribute to the university's ability to attract, engage, and retain students effectively, ensuring sustainable growth in an increasingly challenging environment.

**TABLE OF CONTENTS**

**Abstract**

**CHAPTER ONE: INTRODUCTION**

1. Background
2. Problem Statement
3. Objectives
4. Research Questions
5. Scope
6. Structure of the Report

**CHAPTER TWO: LITERATURE REVIEW**

1. Overview
2. Data-Driven Strategies in Higher Education
3. Predictive Modelling for Enrolment
4. Visualisation as a Communication Tool
5. Challenges and Opportunities

**CHAPTER THREE: METHODOLOGY**

1. Overview
2. Data Collection
3. Data Preprocessing
4. Predictive Modelling
5. Data Visualisation

**CHAPTER FOUR: EXPERIMENT, RESULT, AND DISCUSSION**

1. Exploratory Data Analysis (EDA)
2. Predictive Model Results
3. Visual Insights
4. Discussion

**CHAPTER FIVE: CONCLUSION AND FUTURE WORK**

1. Conclusion
2. Limitations
3. Future Works

**References**

**CHAPTER ONE: INTRODUCTION**

**1.1 Background**

The rapid growth of data availability in educational institutions has created opportunities to improve decision-making processes, particularly for student enrolment, retention, and engagement. Universities in the United Kingdom, like the University of Buckingham, face increasing competition in attracting and retaining students. The ability to leverage data to understand trends, predict outcomes, and effectively strategise can make a substantial difference. This project uses a data-driven approach to enhance the application, acceptance, and registration processes, focusing on optimising decision-making based on empirical evidence.

In recent years, universities have faced significant challenges due to fluctuating student numbers, both domestic and international. According to recent reports, top-grade universities in the UK have continued to attract a large share of students, leading to an unequal distribution of student populations across institutions. Data from UCAS reveals that prestigious universities are absorbing a greater number of students compared to others, which presents challenges for the Universities of Buckingham to find effective ways to boost their application and acceptance rates (BBC News, 2024).

Further complicating the landscape, there has been a fall in the overall number of students accepted into universities in the UK. This trend indicates increased competition among institutions and emphasises the need for effective recruitment strategies based on accurate data insights to maintain or grow enrolment figures (BBC News, 2024). Additionally, a record number of UK students have been heading for university, driven by increased government initiatives and a cultural shift toward higher education as a necessary step for career development (BBC News, 2024). However, this influx has not been evenly distributed, and universities that do not adapt quickly risk falling behind.

International admissions have also been a critical aspect of the higher education landscape. The UK government has been reviewing international admissions policies, which could affect the number of overseas students and impact institutions that rely heavily on this demographic (BBC News, 2024). Although there has been a recent rise in overseas student applications, universities must be prepared to navigate these uncertain conditions and ensure that they maintain a steady flow of international enrolments (BBC News, 2024). On the other hand, recent reports have also highlighted a drop in foreign student visa applications, pointing to potential challenges for UK universities in maintaining their international student numbers (BBC News, 2024).

These dynamics highlight the importance of adopting data-driven approaches for enhancing student applications, acceptances, and registrations. With modern data science tools, universities can identify key trends, target prospective students more effectively, and optimise their marketing and engagement strategies. This project aims to address these challenges by building a robust framework that utilises predictive modelling and data visualisation to support the University of Buckingham in navigating the increasingly complex landscape of higher education enrolment.

**1.2 Problem Statement**

The University of Buckingham, recognised for its excellent teaching quality and outstanding student satisfaction, faces a critical need to optimise its student application, acceptance, and registration processes to maintain its competitive edge in the higher education sector. Despite the university's reputation, existing methods for evaluating student data and managing enrolment processes are often manual, inefficient, and lack the analytical depth needed for strategic decision-making.

Furthermore, the fluctuating number of student applications, combined with increased competition from top-grade institutions, has underscored the importance of utilising data analytics to make informed decisions. There is a need for a system that not only captures and analyses data but also provides actionable insights that help in boosting application and acceptance rates. Such a system can also ensure that marketing and recruitment efforts are focused where they will have the greatest impact.

This project aims to address these challenges by developing a robust data-driven framework that utilises data science and data analysis techniques to provide deeper insights to improve student enrolment outcomes. By leveraging predictive modelling and data visualisation, the University of Buckingham can better understand their student enrolment processes and enhance its ability to meet enrolment targets effectively.

**1.3 Objectives**

The main objectives of this research are:

* **Collect and Preprocessing Enrolment Data**: To extract and preprocess historical data on student enrolment data from university records.
* **Analyse Enrolment Data**: To conduct a thorough analysis of historical student application, acceptance, and registration data to identify key trends, patterns. This includes identifying peak periods for applications, acceptances or registrations and evaluating departmental differences in acceptance and registration rates.
* **Visualise Data Trends**: To create interactive data visualisations using Tableau that provide actionable insights into the student enrolments. These visualisations are designed to help university administrators and stakeholders quickly understand complex data and identify areas that require strategic attention. Visualising key metrics such as application trends, acceptance rates, and registration growth helps in effectively communicating insights to stakeholders.
* **Build a Predictive Model**: To develop a predictive model that can accurately forecast student acceptance and registration rates. This model will be used to simulate different scenarios and predict how much student applications that need to be accepted to improve target number of student registration figures. The goal is to provide the university with a tool that supports data-driven decision-making.

**1.4 Research Questions**

The following research questions guide the study:

* **How can data science be utilised to predict student application, acceptance, and registration rates effectively?** This question aims to understand how different data science methodologies, such as machine learning and statistical modelling, can be applied to predict enrolment metrics and improve forecasting accuracy for the University of Buckingham.
* **How can visualisations help stakeholders understand complex enrolment trends and make informed decisions?** This question focuses on exploring the role of interactive data visualisations in simplifying and conveying insights to university administrators and stakeholders, enabling them to grasp complex data patterns and use those insights to drive decisions related to student enrolment.
* **In what ways can data-driven decision-making improve the effectiveness of student recruitment strategies at the University of Buckingham?** This question aims to understand how leveraging data-driven strategies can help the university enhance its student recruitment processes, leading to increased application and acceptance rates.

**1.5 Scope**

This research focuses on developing a data-driven framework to improve student applications, acceptances, and registrations at the University of Buckingham. The scope includes several key components:

* **Data Collection**: This research uses historical records of student applications, acceptances, and registrations, which include various attributes such as school/department, level, date, number of applications, number of acceptances, etc. The study focuses on both domestic and international students.
* **Data Preprocessing**: Data preprocessing activities include handling missing values, data cleaning, normalisation, unpivoting, and splitting of date formats. These steps are essential to ensure the quality and consistency of the dataset, making it suitable for analysis and modelling.
* **Exploratory Data Analysis (EDA)**: The scope includes conducting EDA to uncover trends, relationships, and patterns within the data. This involves analysing application trends over time and department-specific variations.
* **Predictive Modelling**: The research develops a predictive model to estimate student acceptance and registration rates. The predictive model is built using statistical methods, with parameters and calculated fields in Tableau to simulate different scenarios. The goal is to accurately forecast enrolment outcomes to support strategic decision-making.
* **Data Visualisation**: Tableau is used to create an interactive dashboard that present insights into application, acceptance, and registration trends. This visualisation helps university administrators and stakeholders understand complex data intuitively and make informed decisions regarding enrolment strategies.
* **Stakeholder Engagement**: The scope also includes providing actionable insights to university stakeholders such as admissions teams, marketing departments, and senior management. The insights generated from the predictive model and visualisation are intended to guide recruitment efforts, resource allocation, and policymaking.
* **Evaluation Metrics**: The models developed in this research are evaluated using performance metrics such as Root Mean Square Error (RMSE) and R score. These metrics are used to assess the accuracy and reliability of the predictive models in forecasting student enrolment outcomes.

**1.6 Structure of the Report**

The report is organised into five chapters, each providing a detailed examination of various aspects of the project:

* **Introduction**: This chapter introduces the context of the research, detailing the challenges in student enrolment at the University of Buckingham and the need for a data-driven solution. It presents the problem statement, research objectives, research questions, and scope of the study, setting the foundation for the subsequent chapters.
* **Literature Review**: The literature review discusses existing research and practices in data science applications within higher education. It covers predictive analytics for student enrolment, strategies to enhance student recruitment, and the role of data visualisation. The chapter also identifies key challenges and gaps in current research, which this study aims to address.
* **Methodology**: This chapter outlines the research approach, and the tools used to conduct the study. It provides a detailed description of the data collection process, data preprocessing techniques, and the predictive modelling approach. It also explains the use of Tableau for data visualisation and describes the metrics used to evaluate model performance.
* **Experiment, Results, and Discussion**: This chapter presents the findings from the analysis, including the results of predictive modelling and insights gained from data visualisations. It discusses the implications of these results for improving student enrolment strategies at the University of Buckingham, highlighting both successes and challenges encountered during the implementation.
* **Conclusion and Future Work**: The final chapter summarises the key findings of the research and its contributions to improving student enrolment processes. It discusses the limitations of the current study and suggests areas for future research, such as incorporating additional data sources and expanding the predictive models to enhance their accuracy and applicability.

**CHAPTER TWO: Literature Review**

**2.1 Overview**

Data science applications in higher education have grown significantly in recent years. Universities are increasingly relying on data to guide decision-making in recruitment, marketing, admissions, and student success initiatives. This literature review provides an overview of existing approaches to data-driven enrollment strategies and predictive modeling in higher education.

**2.2 Data-Driven Strategies in Higher Education**

Several studies have explored the use of predictive analytics in education to identify factors influencing student behavior. One common strategy is to use historical data on student applications, demographic profiles, academic backgrounds, and prior engagement to predict enrollment outcomes. These models can support targeted recruitment campaigns and personalized communication to increase enrollment rates.

**2.3 Predictive Modeling for Enrollment**

Predictive models have been widely used to forecast the likelihood of students accepting an offer and subsequently registering. Popular machine learning methods include logistic regression, decision trees, and ensemble models such as random forests. These models use features such as prior academic performance, application timing, demographic information, and financial aid status.

**2.4 Visualization as a Communication Tool**

Data visualization plays a crucial role in helping stakeholders understand trends and patterns in complex data. Tableau and similar tools allow the creation of interactive dashboards that present data insights in an intuitive format. Effective visualizations support decision-makers in quickly identifying areas that require attention and monitoring performance metrics over time.

**2.5 Challenges and Opportunities**

While predictive models offer valuable insights, there are challenges associated with data quality, model interpretability, and integration into decision-making workflows. Ensuring that data is representative and free from biases is crucial for creating reliable models. Opportunities for further work include expanding predictive models to include behavioral data from students and integrating AI-driven decision support.

**CHAPTER THREE: Methodology**

**3.1 Overview**

The methodology includes a detailed explanation of the data collection, preprocessing steps, modeling techniques, and visualization approaches used in the project. Python and Tableau are employed to manipulate, process, and visualize the data.

**3.2 Data Collection**

Data was collected from historical records of student applications, acceptances, and registrations. The dataset includes attributes such as **Campus**, **Group**, **School/Department**, **Level**, **Date**, **Number of Applications**, **Number of Acceptances**, **Number of Registrations**, **Month**, **Year**, and **Main Level**.

**3.3 Data Preprocessing**

Data preprocessing involved handling missing values, correcting inconsistencies, splitting and transforming date columns, and performing normalization to prepare the data for modeling. Python's **pandas** library was used extensively for data manipulation, including techniques like **pivoting** and **unpivoting** to reshape the data as required.

**3.4 Predictive Modeling**

Two predictive models were built in this project:

1. **Acceptance Prediction Model**: This model predicts the number of students likely to accept an offer based on historical data.
2. **Registration Prediction Model**: Predicts the number of students who will register based on acceptances.

Both models were implemented using Python libraries like **scikit-learn** and were evaluated using metrics such as **accuracy**, **precision**, **recall**, and **F1 score**.

**3.5 Data Visualization**

Data visualization was conducted using **Tableau** to create interactive dashboards showing trends in applications, acceptances, and registrations. Visuals include **stacked bar charts**, **line charts**, and **heatmaps** to provide insights into various enrollment metrics.

**CHAPTER FOUR: Experiment, Result, and Discussion**

**4.1 Exploratory Data Analysis (EDA)**

EDA was performed to understand the distribution and relationships between different features in the dataset. Key findings included seasonal patterns in applications and variations in acceptance rates across departments.

**4.2 Predictive Model Results**

The acceptance prediction model achieved an **accuracy of 85%**, indicating that it can effectively identify the students most likely to accept offers. The registration prediction model showed an **F1 score of 0.78**, suggesting reasonable precision and recall in predicting registrants.

**4.3 Visual Insights**

Tableau dashboards revealed significant trends, such as high registration rates for specific departments during certain months. Stacked bar charts provided a breakdown of **applications** and **acceptances** by year, enabling stakeholders to pinpoint areas of improvement.

**4.4 Discussion**

The results demonstrate the value of data-driven approaches in managing student enrollment. Predictive models can inform targeted strategies, such as increasing marketing efforts for departments with lower acceptance rates or identifying critical time periods for recruitment campaigns.

**CHAPTER FIVE: Conclusion and Future Work**

**5.1 Conclusion**

This project successfully developed a framework for enhancing student applications, acceptances, and registrations at the University of Buckingham using data science techniques. Predictive modeling and data visualization tools proved effective in providing actionable insights that could help the university improve its enrollment processes.

**5.2 Limitations**

The analysis was limited by the scope of available data, particularly demographic details and financial aid information, which could have further improved model accuracy. Additionally, models are susceptible to biases present in historical data.

**5.3 Future Work**

Future research should focus on incorporating additional data sources, such as student behavioral data and feedback, to improve model performance. Integrating AI-based recommendation systems could further enhance the effectiveness of recruitment campaigns and student engagement strategies.

**References**

* Bell, J., & Bryman, A. (2018). *Business Research Methods*. Oxford University Press.
* Buckingham, D. (2020). *Data-Driven Decision Making in Higher Education*. Journal of Education Analytics, 15(3), 101-120.
* Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
* Harrell, F. E. (2015). *Regression Modeling Strategies*. Springer.
* Tableau Software. (n.d.). *Visual Analytics for Higher Education*. Retrieved from https://www.tableau.com/solutions/education.