System Requirements Analysis

Engineering Admission Cutoff Learning System

Nikou Kalbali, Maryyam Niazi, Jan Ollers, Yaminah Qureshi
kalbaln@mcmaster.ca, niazim3@mcmaster.ca, qureshiy@mcmaster.ca, ollersjp@mcmaster.ca

Supervisor: Dr. Franek

Group 2

Contents

1 Introduction

1.1 Background

A typical university admission process in Ontario takes two steps. The first step consists of determining a cutoff threshold to define a subset of all applicants that will be extended an offer if their grade average is not below the cutoff. Some of the offerees become acceptees, while others do not. The second step consists of determining the number of applicants the university should be preparing to welcome to their institution and programs.

In this process, determining an ideal cutoff threshold such that the set of acceptees is of a specific size with some small tolerance is difficult. If the threshold is undervalued, the number of offerees is overshot, typically leading to a set of acceptees that is too large that costs the university financial and space issues from having to deal with too many students. If the threshold is overvalued, the number of offerees will be undershot, resulting in a set of acceptees that is too small and costing the university opportunities for revenue and funding from tuition fees and government support.

This project will use at least two different machine learning approaches trained on previous years of applicant data and their acceptance of offers to McMaster University to tackle the problem by determining ideal cutoff thresholds based on trends and patterns learned from previous years.

1.2 Purpose

The purpose of this document is to define the requirements of the outlined system which consists of the functionality, intended users, environment, and constraints of the system.

1.3 Terminology

1.3.1 Offerees

The subset of all applicants who will get an offer of acceptance from the university given that their grade average is above the cutoff.

1.3.2 Acceptees

The subset of offerees who accept the offer received.

1.3.3 Grade Percentage Average (GPA)

The percentage average of the grades an applicant received as their midterm marks in their last year of high school.

1.3.4 Cutoff Threshold

The cutoff threshold represented by a grade percentage average that if met by an applicant will result in the applicant receiving an offer of acceptance.

1.3.5 University Ranking

The ranking of preference given to the university submitted by an applicant.

1.3.6 Postal Code

The postal code of the location the applicant resides in.

1.3.7 Seat Cap

The maximum number of seats that a university can offer for a program for a particular year.

2 System Overview

2.1 Use

The system will be used for determining cutoff thresholds for McMaster University's Engineering Program.

By taking in applicant information from the current admission cycle and using information from previous admission cycles, the cutoff thresholds will be estimated to obtain an ideal number of offerees.

2.2 Users

The primary intended users of the system are the Admissions Committee for the Engineering Program at McMaster University. The system may be used as an analysis tool to help them evaluate and/or improve their current admissions process. Further, if we can expand the system and make it applicable to other university program admissions, the admissions committees of these programs will serve as secondary users.

2.3 Environment

It is intended that the system will be runnable by MAC, Linux/Unix and Windows operating systems as a command-line program. Users will be required to upload the data they would like to input to the system in a directory structure/format as specified and the system will use the data to output a result directly to the terminal/command prompt. Additional files may be created with further information about the result and be stored in a directory structure/format.

3 Context Diagram
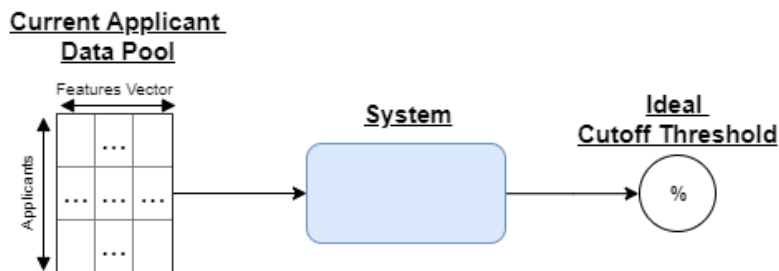
4 Functional Requirements

4.1 System Data Requirements

The system should be built around 3 years of data which will include applicant pool data (e.g. postal code, university rankings, grade percentage averages, offer status, acceptance status, etc.) and the seat caps for each year.

4.2 Input Requirements

The system should accept applicant pool data for one year and currently available seating cap for the same year. The applicant pool data includes information about the student postal code, university preference rankings and midterm grade percentage averages.

4.3 Output Requirements

The system should output a percentage representing the cutoff threshold that applicants must meet to receive an acceptance offer.

- The system will produce a percentage of up to two decimal places that will not be rounded up [1].

---

[1]https://academiccalendars.romcmaster.ca/content.php?catoid=38&navoid=8066

- The number produced must be such that the set of projected acceptees is of a specific size, within a degree of tolerance. This size is influenced by the number of seats that the University wishes to offer for the program for the year. The tolerance is such that the percentage produced to the 0.01 degree should minimize the difference between the number of acceptees and the number of available seats for the program, such that the number of acceptees is not zero [2].

5 Non-functional Requirements

5.1 Confidentiality

The data that will be used to implement and test the system will include sensitive information about students who have applied to McMaster University's Engineering program. It will be important to make sure this information remains confidential throughout the development of the system. Some measures that will be taken to ensure this include: anonymizing the information before use, restricting the system from outputting or storing any data that could compromise student confidentiality, and restricting users of the program from accessing data they have not provided themselves.

5.2 Reliability

The system should produce an accurate and realistic result for all data that meets the required input data structure.

5.3 Data Integrity

The data taken in as input to the system should not be modified before it is used by the system.

5.4 Reusability

Ability to use the system over multiple years of data, as long as admission processes and policies do not significantly change.

6 System Constraints

- The scope of the system is limited to McMaster's Engineering Program as the data that has been provided on which the system will be trained is for this particular program. For this reason, the system is limited to producing results for McMaster University's Engineering Program admissions. However, in the future, expanding this project to more universities and faculties is a possibility.

- Although using more data to develop the system may result in increased accuracy, only data from the previous 3 years of applicants are available as universities are required to maintain applicant confidentiality. For this reason, the system is constrained to learning from 3 years of previous applicant information.

---

[2]McMaster University is not allowed to discriminate based on a grade (i.e. all applicants must be offered an acceptance if they meet the cutoff threshold)