

JC2066 Assignment 1: College Admissions Algorithms

Learning objectives:

- Understand what algorithms are, how they might be used
- Understand how algorithms can reflect bias
- Understanding outliers, missing information in data, and human judgements in algorithmic decisions
- Reflect on that algorithm's design and implications
- Reflect on that discussion with AI tools

Requirements

Answer all the following questions in PARTs 1 to 3, compile all your answers into a single PDF or a docx file, and submit via Canvas. The deadline to submit your answers to Canvas is **Nov. 7th (Tuesday), 23:59.** Please note that this assignment accounts for 20% of your final course score.

PART 1 SURVEY

Please respond to the questionnaire (link: <https://forms.gle/zV5JHKn1QFL21dHP8>) based on your personal experiences BEFORE finishing other parts.

PART 2 REFLECTION ON ALGORITHMIC DECISION MAKING

Background

As you know, the college admissions process involves a lot of types of data from prospective students to make decisions. With the number of applicants increasing, colleges may begin relying on algorithms to select which applications should receive more intensive human review. An algorithm could use *quantitative data* – such as GPA and SAT or other standardized test score – to provide initial recommendations. In fact, there is more data available than ever. [Many colleges even track data about prospective student engagement](#) – e.g., whether they open emails, visit the college website, engage on social media, etc. This creates a “*demonstrated interest*” value.

[Based on a recent survey of college admissions officers](#), we know some of the weights that humans tend to give to these different types of data.

Your task will be to describe an algorithm that provides a recommendation for which prospective students are likely to be the best candidates for admission.

[IMPORTANT!] When answering the following questions in PART 2, please DO NOT use any AI tools. You will be able to use AI tools in PART 3 and will be asked to document your conversations with AI tools.

Q1: How should this algorithm work? (10 points)

In plain language, write out how you think an algorithm based on student applications should work. e.g., “It should find the students with the highest grades but also take into account what high school they went to...”

Q2: How much weight should be given to each factor, what should the cut-off score be, and what does this algorithm look like? (10 points)

To make things simpler for this algorithm, we have turned this prospective student data into a set of factors, all with a value on a **0 to 10 scale**, with 0 being the worst and 10 being the best.

- **Test Scores:** standardized tests such as HKDSE, Gao Kao, SAT, and ACT
- **Grades:** grade point average with more weight for advanced placement courses
- **Strength of Curriculum:** how difficult the curriculum at their high school is (or, how “good” of a high school they went to). Note that this factor is often correlated with socio-economic status (e.g., students whose families have a lot of money are more likely to go to very “good” schools).
- **Demonstrated Interest:** based on online engagement data (see description above)

An algorithm like this would actually be attaching weights to different factors. Let’s create an algorithm that does that.

Decide on a weight for each of the four factors, so that the weights add up to 100. Basically, this means how much each of these factors should matter to the algorithm’s decision. For example, if test scores are VERY important, give it a weight of 50%, or if demonstrated interest is not important, give it a weight of 10%. The total of the 4 weights should add up to 100.

Test scores: ____

Grades: ____

Strength of curriculum: ____

Demonstrated interest: ____

All of these factors are on a 0 to 10 scale with 0 being the worst and 10 being the best. If you averaged all of these together, the very highest overall score you could get would be 10!

Choose a number between 0 and 10 that you think should be the cut-off to further consider a student for admission:

Cut-off score: ____

Here is an example of the type of admissions algorithm you might have just created, if you decided that 5 should be the cut-off. This algorithm has test scores as 20%, grades as 40%, strength as 30%, and interest as 10%.

Consider this student IF $(\text{SAT} \times 0.2 + \text{GPA} \times 0.4 + \text{Strength} \times 0.3 + \text{Interest} \times 0.1) > 5$

Fill in the blanks for your own admissions algorithm, based on your answers above:

IF $(\text{SAT (or other test)} \times \text{____} + \text{GPA} \times \text{____} + \text{Strength} \times \text{____} + \text{Interest} \times \text{____}) > \text{_____}$

Q3: Reflect on the outcomes (10 points)

First, create at least three *hypothetical* candidates for admission, make up data for all of them (with each on a 0 to 10 scale), and then calculate whether or not they would be further considered for admission by your algorithm. For example, let's say that Jack has a Test Score value of 5, a Grades value of 8, a Strength value of 10, and an Interest value of 2.

In 100 to 200 words, reflect on the outcomes of this algorithm. What are some examples of ways this algorithm might give undesirable results for certain types of prospective students?

Q4: Algorithmic bias (10 points)

Assume that when implementing this algorithm in real world, the programmer has heard about “protected” characteristics, such as age, gender, and race, and that they try to avoid using these features when implementing the algorithm. Do you believe that the algorithm includes any biases? Briefly explain your answer, and give examples if possible.

Q5: Mitigating bias (10 points)

- a. What group of students do you think might be the privileged group in your algorithm? What about the unprivileged group? Please briefly explain why.
- b. What strategies do you think we can use to mitigate the bias in your algorithm? Please explain your strategies with examples and details (e.g., what constraints you may want to use, and how you may reweight group/label).

Q6: How should and shouldn't algorithms be used in college admissions? Why? (10 points)

Given the realities of time and resources (i.e., only so many person hours can be involved in the admissions process), how *should* an algorithm like this be used in the admissions process (assuming an algorithm that is far more sophisticated than the one you just created!)? How *shouldn't* it be used? Explain your thinking, taking into account possible undesirable outcomes.

PART 3 REFLECTION ON AI OUTCOMES

Large Language Models (LLMs)' fast and articulate answers to professional questions can enable efficient searching for relevant expert knowledge. They further enable users to engage in discussions and encourage critical thinking as they can be prompted to act as different roles when interacting with users.

In this part, you are required to first discuss with an AI assistant (e.g., ChatGPT, CityU Chatbot, or POE Chat) on several tasks, and then answer the questions based on your experience.

You can choose to use one of the following AI tools:

- OpenAI: <https://chat.openai.com/>
- POE: <https://poe.com/>
- CityU GPT: <https://www.cityu.edu.hk/GenAI/usage-quota.htm>

Q6: Conversation Flows (15 points)

Engage in a conversation with the AI assistant of your choice to discuss the questions in PART 2 (Q1-Q6). Please document your conversations by:

- Taking screenshots of the full conversation flow from start to finish.
- Clearly indicating when you are posing each original question from Q1-Q6 to the AI assistant.
- Showing both your questions and the AI's responses in the sequence they occurred.
- Exporting and compiling the full conversation into a separate document if needed.
- Annotating parts of the conversation to point out key sections, insights, or limitations.
- Providing additional context on the flow and progression of the dialogue for clarity.

Your documentation should allow the reader to clearly follow your full conversation with the AI assistant and see how the interaction progressed in addressing each original question. The goal is to illustrate how the AI tool responded and contributed insights on the case study and reflection topics.

Q7: Reflect on the AI tool's outcomes (15 points)

In 200 to 400 words, reflect on the outcomes of this discussion process with AI.

- a. What were the main takeaways from the discussion? What new perspectives or considerations did the AI assistant introduce compared to your own initial thoughts? Were there any surprises or limitations? Please briefly explain why.
- b. In what ways did this change or reinforce your understanding of the admission algorithms and their issues related to ethics and algorithmic fairness?

Q8: Algorithmic fairness (10 points)

In 100 to 200 words, reflect on the outcomes of this discussion process with AI.

- a. Do you think the AI assistant exhibited any biases when discussing the admissions algorithm? Why or why not?
- b. How did the AI's perspectives compare to your own views on algorithmic fairness and mitigating bias?
- c. Based on this experience, what do you think about AI assistants' abilities to have nuanced discussions about ethics and fairness? Briefly explain your answer.