Question:

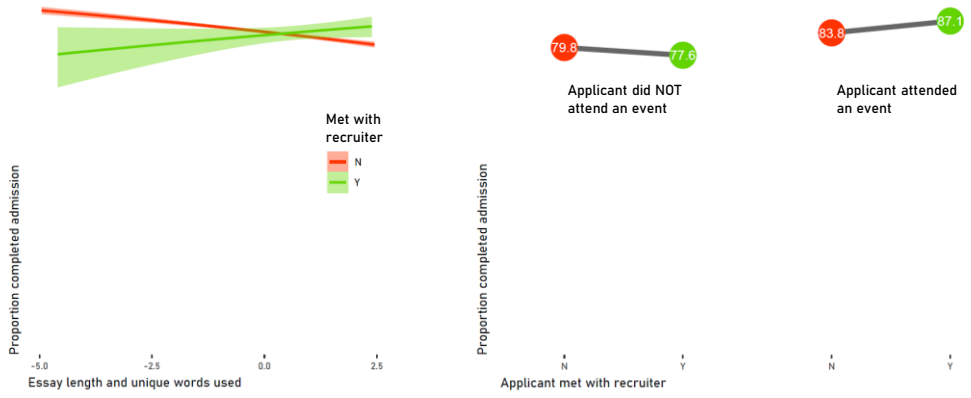
Given finite recruiting resources, which candidates should you prioritize?

Question:
Given finite recruiting resources, which candidates should you prioritize?

Approach:
What about a candidate best predicts whether meeting with a recruiter decreases the probability of withdraw?

Executive Summary:

Prioritize candidates that write longer essays with more unique words and those who attended an event.

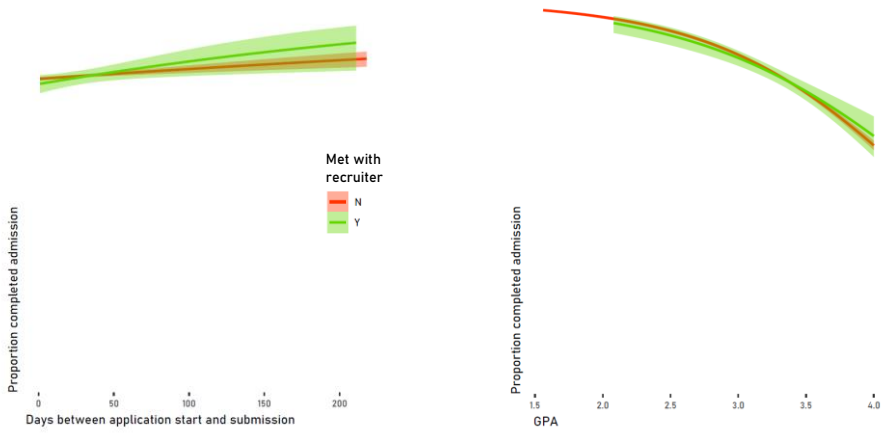


The probability of completing the application process is increases with meeting with a recruiter for applicants with above average mean essay length and number of unique words. However, recruitment meetings are associated with increase probability of withdraw for those with below average essays.
For a participant with high essay scores (PC1 = 2.4), meeting with a recruiter is expected to yield an increase of 3% in the probability of completion.

An increase in the probability of completion is also expected when recruiters meet with applicants who attended a T4F event by about 1%.

NOTE: The “essay score” is the product of a principle component dimension reduction of the three essay lengths and number of unique words. The PC accounts for >85% of variation. It is scale and centered, so 0 can be roughly interpreted as the mean average essay length and number of unique words.

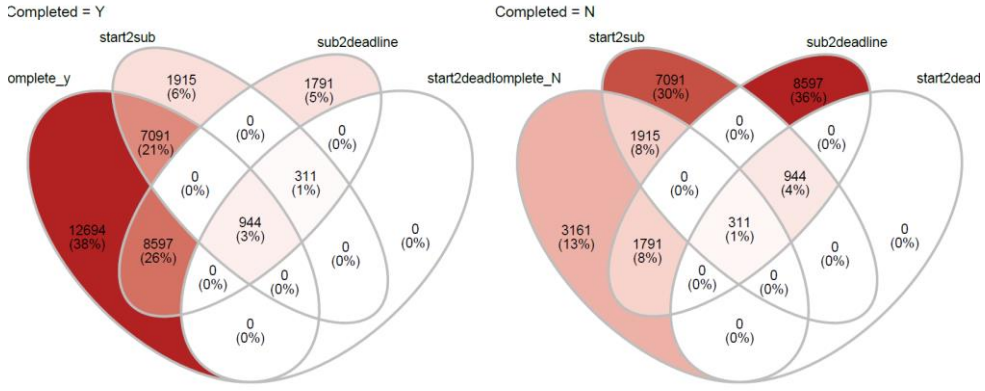
Also consider targeting applicants that start their applications earlier and those with close to 4.0 GPA.



Although, GPA and the duration from application start to submission did not have a significant interaction with whether the applicant met with a recruiter, these variables were highly important in the random forest model and may be worth experimenting with targeted experimental trials.

Ideally, target applicants that start their applications at least 50 days prior to the application deadline or those with near 4.0 GPA.

Of those that complete the admissions process, most start well before the deadline, take at least a day to work on the application before submission, and submit before the deadline.



This graph looks at the number of applications that had 0 days between application timeline points (start, submission, and deadline). Of those that complete the admissions process, most start well before the deadline, take at least a day to work on the application before submission, and submit before the deadline. However, most of those that withdraw submit on the deadline date or finish the day that they begin the application.

NOTE: Given more time, I would ordinarily fix the labels to make them more easily interpretable.

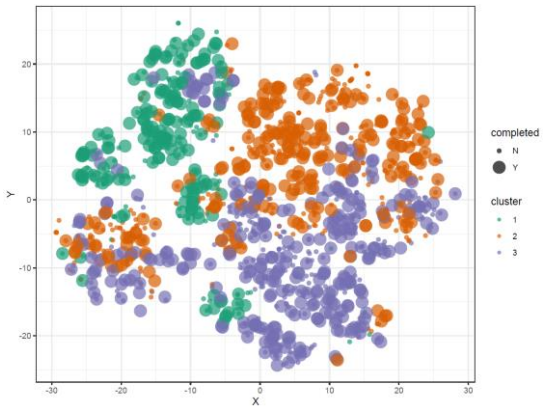
Consider de-prioritizing recruitment efforts toward business majors while prioritizing early childhood education, Spanish, and journalism majors.

Although application major did not have a significant interaction with whether an applicant met with a recruiter, the trends suggested by the model could be worth further consideration.

NOTE: Ordinarily, I would fix the labels to avoid cutting off content.

Caveats:

This analysis assumes that the applicants who met with recruiters in the dataset were randomly selected.

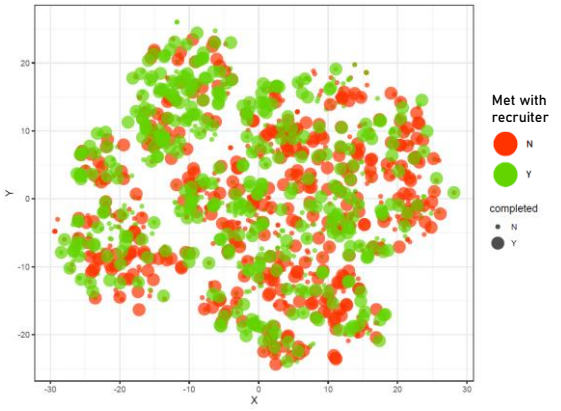


Applicants can be roughly clustered into three groups based on all of their characteristics.

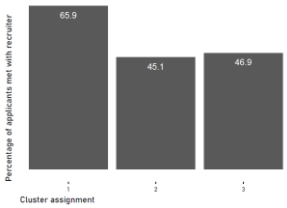
Multivariate clustering based on all applicant characteristics identified three, somewhat overlapping groups. Ideally, we would like to see recruitment effort evenly distributed over all groups. Overrepresentation of recruitment effort in one group would indicate that there is either bias in which applicants seek out recruitment meetings or bias in the type of applicants recruiters target.

Caveats:

This analysis assumes that the applicants who met with recruiters in the dataset were randomly selected.



Although recruiters met with all groups, that effort was biased toward one cluster.

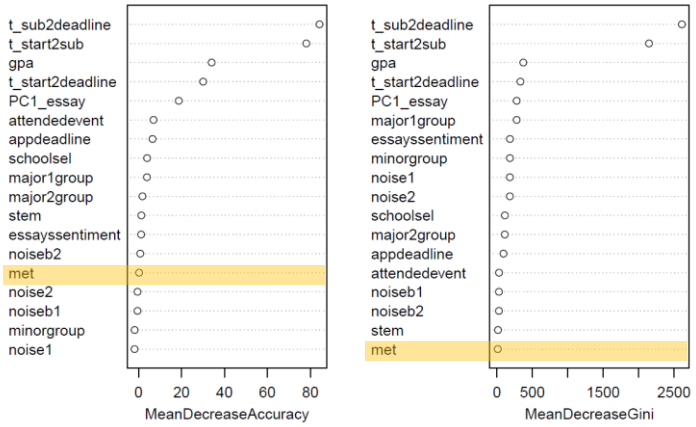


Cluster 1 is biased towards heavier recruitment effort, but it is the smallest group. Individuals in this cluster were more likely to attend an event, wrote longer essays, and started their applications much earlier compared to the other groups. Their GPA and essay sentiment falls between those of the other groups.
NOTE: This might could bias could skew earlier estimated. With more time, I would test the sensitivity of my results my simulating more or less skewed samples and rerunning my analyses.

NOTE: This is analysis was run on a subsampled and evenly weighted sample (with respect to recruitment) so we would expect the percentages in the bar chart to be at 50%.

Caveats:

The dataset does not indicate a strong effect of recruitment efforts.



Whether or not a participant met with a recruiter was not ranked as an important predictor of completion of the application process in the rando forest model. In fact, that variable was consistently less important that the random binary variables I included (i.e. “noiseb_”).

Although the prompt for this exercise makes an explicit statement that recruiters are effective in decreasing a candidate’s likelihood to withdraw, I would want to do more research on why this variable seems to have such a little effect.

A. Z. Andis Arietta
a.andis@yale.edu
azandisresearch.com