Student Success Analysis
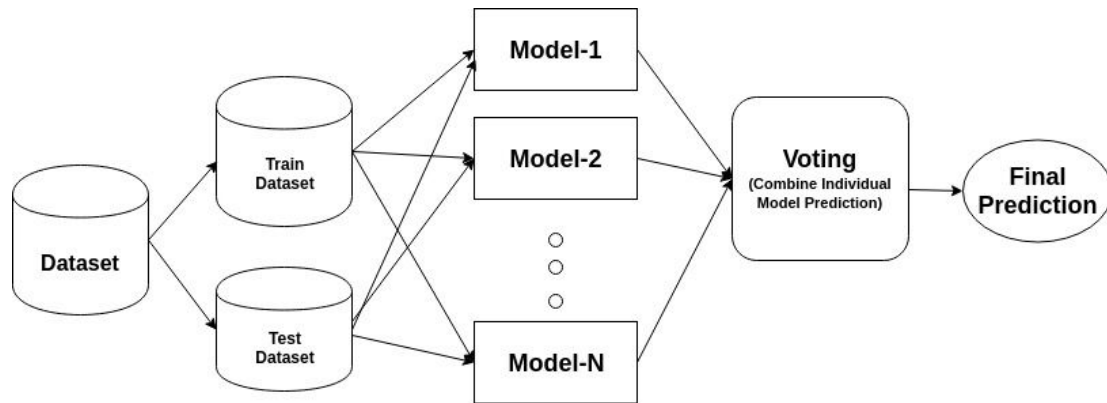
Adam Meadows, Jennifer Robert, Kyra Hull

Student Success Analysis - Model

- Ada Boost Classifier - How it works

"AdaBoost Classifier Algorithms Using Python Sklearn Tutorial." *DataCamp Community*, https://www.datacamp.com/community/tutorials/adaboost-classifier-python.

Student Success Analysis - Classification Matrix

	precision	recall	f1-score	support
fail	0.67	0.46	0.54	804
pass	0.72	0.86	0.78	1285
accuracy			0.70	2089
macro avg	0.69	0.66	0.66	2089
weighted avg	0.70	0.70	0.69	2089

Training Score: 0.686941251596424
Testing Score: 0.7027285782671134

Student Success Analysis - GitHub

- [Student Success Repository](#)
- ReadMe:
  - Description of topic, data source and questions
  - Communication protocols are included
  - Data Management section with a link to the sql commands for table generation
  - Description of Front End with a link
  - Description of the model

Student Success Analysis - Pre-Processing/Cleaning

Steps Taken:

- Narrowed down original csv to potential columns that comprised 16,446 rows
- Removed missing data reducing the dataframe to 8353 rows and 56 columns
- Binned several categories
- Utilized OneHotEncoder
- Reviewed Feature Importance
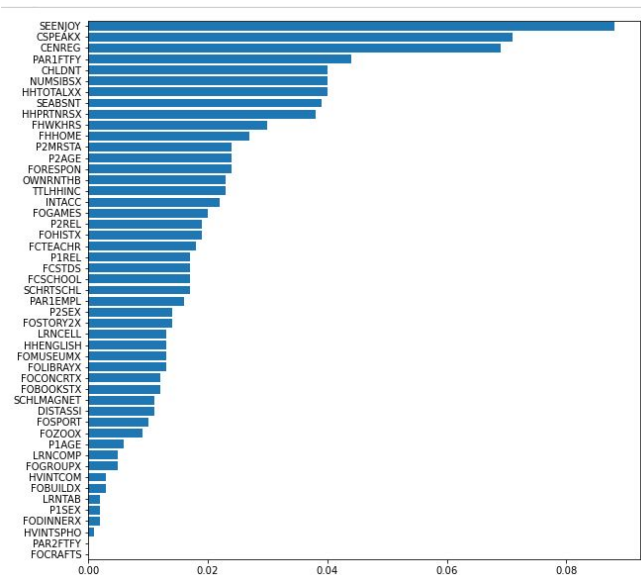
Student Success Analysis - Database

- All data for the analysis is stored in a Postgres database hosted with Heroku
- There are 3 tables in the database. They are:
  - "NHES_19_PFI" - Stores the initial dataset
  - "Features" - Stores significant features, their significance and their group
  - "Surveys" - This stores all questions submitted by the front end for later processing
- Below is a screenshot of the first few fields in the ERD:

NHES_19_PFI			Surveys			Features		
id	🔑 varchar		id	varchar		feature	text	
ALLGRADEX	int		updated	date		value	float	
SEGRADES	int		ALLGRADEX	int		group	text	
CENREG	int		SEGRADES	int				
SCH_TYPE	int		CENREG	int				
EDCPUB	int		SCH_TYPE	int				
EDCCAT	int		EDCPUB	int				

Student Success Analysis - Dashboard

- Combined categories and averages

- Region: 0.078
- School Behavior: 0.042
- School Sentiment: 0.041
- Enrichment Activities: 0.032
- Family Demographics: 0.027333
- Technology: 0.016
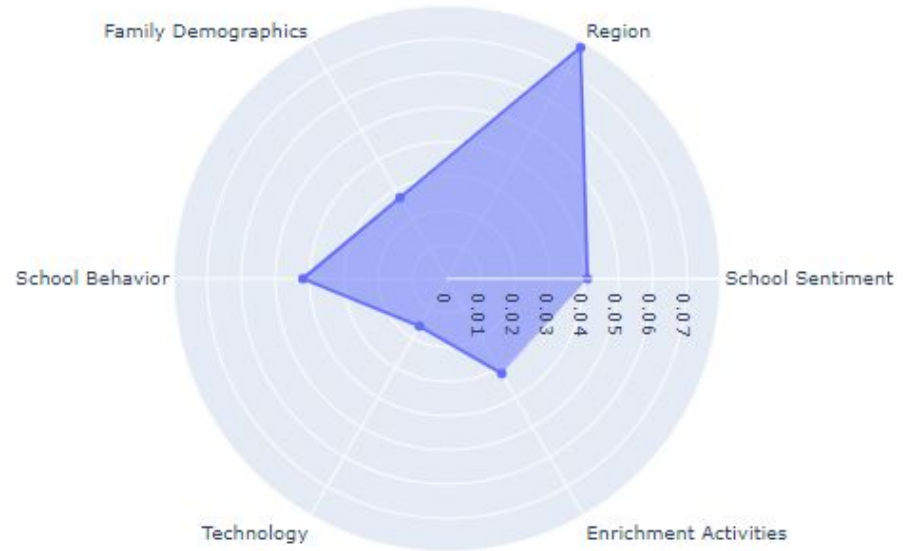
Student Success Analysis - Next Steps

- Highlighting region that were categorized together
- Creating a scale that can highlight
- Creating source information for student in improvement