4 Year New Transfer Student Graduation Prediction

Data Science Team

Fall 2023

Transfer students are important & a distinct subpopulation

- Transfer students are typically more racially, ethnically and socioeconomically diverse than their traditional counterparts
- They're more likely to be "nontraditional," and may be first-generation, working, or with significant financial need. All these factors may very well make transfer students more price sensitive.
- They may experience "transfer shock" and other transition issues
- They may be less engaged than their traditional first-year peers. Less engagement might well make transfers more likely to bolt if they experience a poor student experience or lack of success
- They may be veterans
- They may have competing obligations, such as family and careers

Transfer students are trending downward nationwide

- A lot of people are pointing to the pandemic. Between 2020 and 2022 transfer students are down 7% nationwide
- “Upward transfers” (moving to a 4 year from a 2 year, which is what we’re focused on today) are down over 14% in the same period
- White and Asian students are down the most, but all race and ethnicity categories are down
- Women are down more than men
- 20+ are way down. Under 20’s are up
- Is it high prices? Is it low unemployment? Is it poor hybrid and remote experiences?

Transfer students are an important sub population @ UMBC

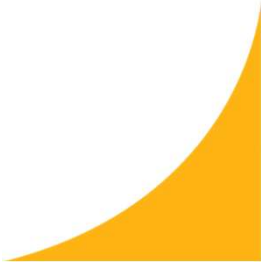
- For lots of reasons, including state policy that encourages upward transfers as a strategy to make college more accessible and affordable
- But transfers are down at UMBC, as they are across the nation
- The old stylized fact:
  - "Transfer students are about half the population" (a little less)
- The new stylized fact"
  - "Transfer students are about a third of the population" (a little more)
- New transfers are down by 5.7% in Fall 2023 (about 50 students)
  - "Note here that new transfer enrollments reflect the 15-25% decline in enrollments at our feeder community colleges"

Transfer students are important to support

- For their future…and for ours
- Improving support can increase retention making UMBC more accessible and affordable, and help mitigate declining new transfer enrollments
- The tools and processes will continue to be useful when the trend reverses
- Transfer students, especially those in the 20+ category, will be helpful as the age-based demographics lead to a declining pool of 18 year olds
- We'll continue to honor our principles and values of inclusive excellence by supporting this sub population
- We generally do a good job with transfers. On the data analytics side of the house, our roadmap includes better models to help transfer students succeed

Special Thanks to Emily Wang

Graduate Assistant in Data Science



UMBC

Overview

- Motivation behind developing a graduation prediction model
- Overview of the tool
- How to use the tool
- Model creation and feature engineering
- Conclusions

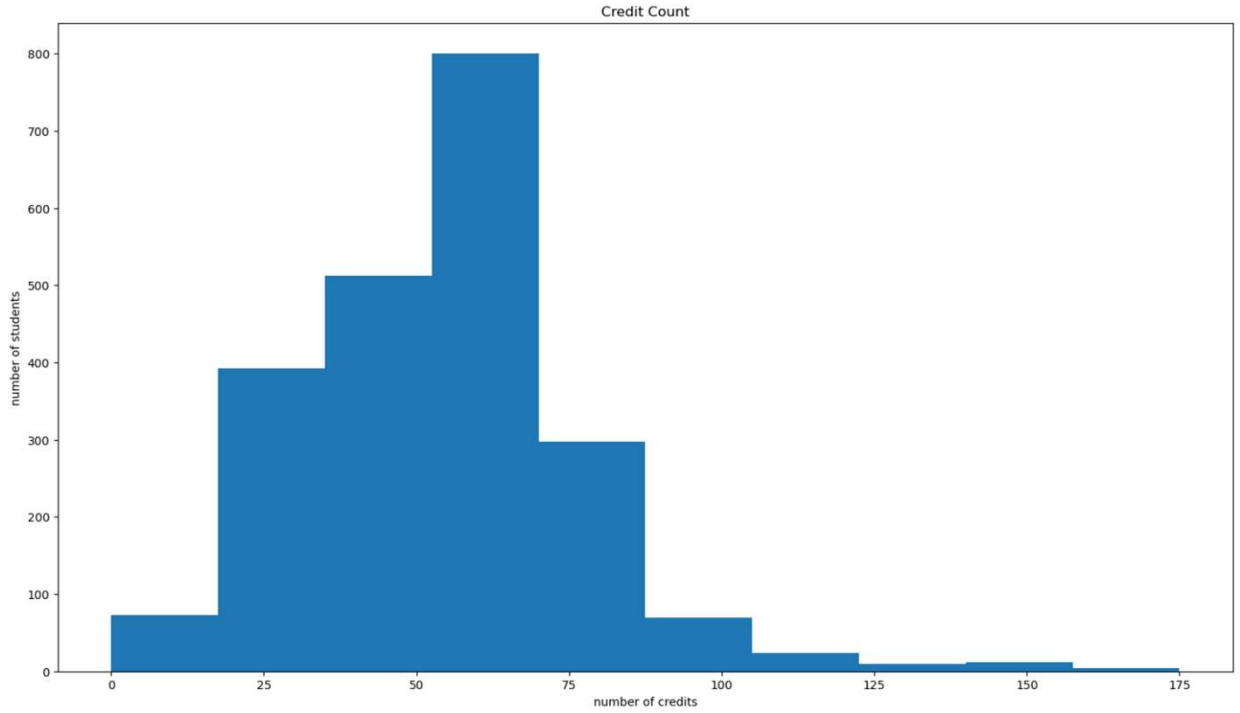
Motivation

- Our group has spent most of our time developing tools and models for First Time Full Time Freshman (FTFTF)
- Transfer students represent a significant population of students that tend to be more conscious of time to degree and cost.
- Developing a model and tool to help these students complete on time is in line with tools we have already developed
- A small lift in graduation rates for transfer students could mean transfer graduation meets or exceeds that of First Time Full Time Freshman.

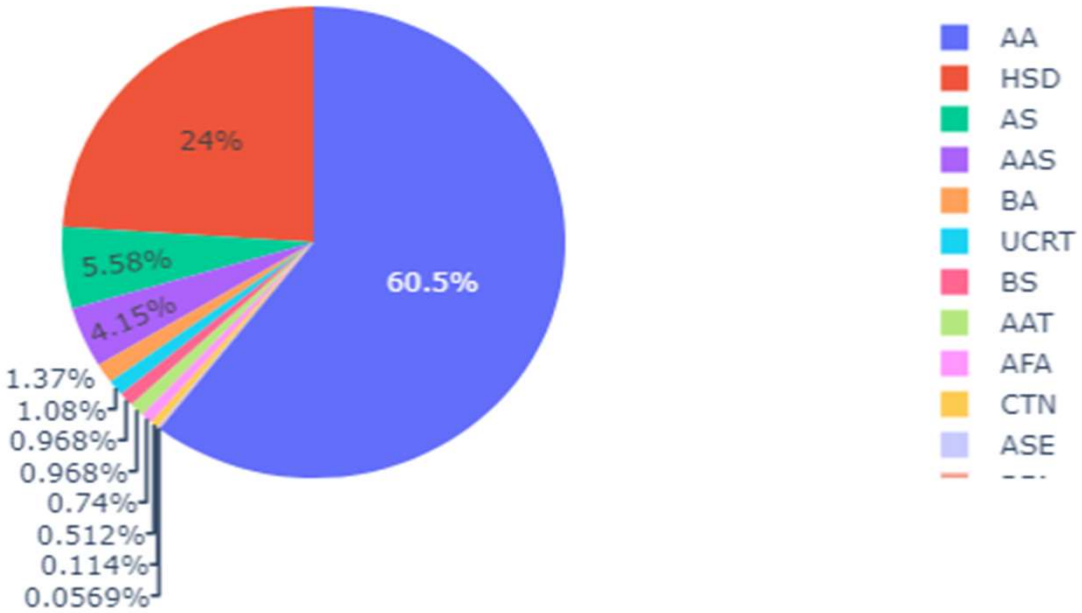
The Data: Focusing on 30

- Training and testing data consists of First Time Transfer students from Fall 2012 to Fall 2017.
- Includes only students entering with < 30 credits (students with >30 credits behave much more similarly to FTFTF)
- 2193 students in total
- 1378 students in this data set graduated in 4 years (62.8%)
- We now have models to predict graduation for the entire undergraduate student population.

Number of Transfer Credits 2012-2017 New Transfer Students

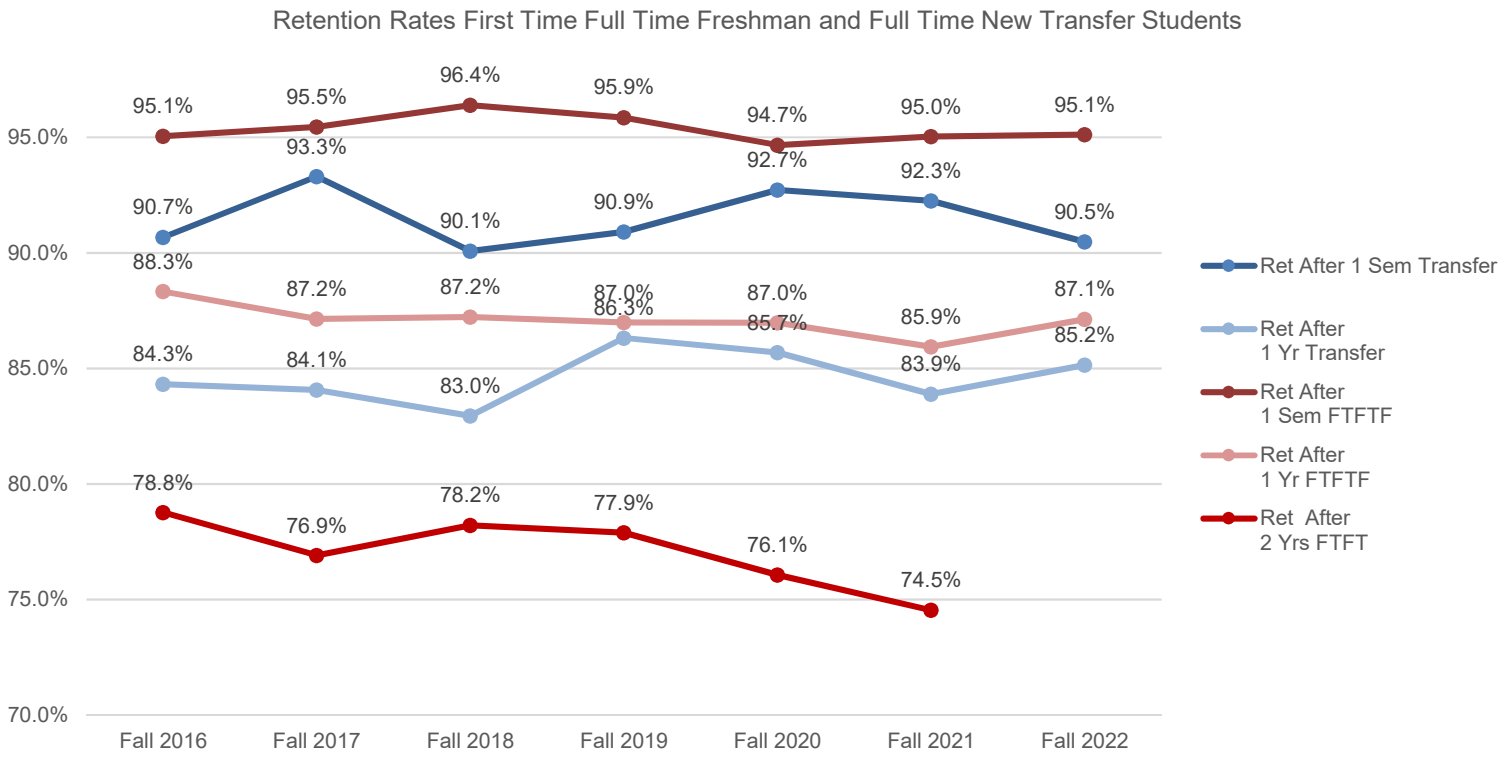


Students that transfer to UMBC come in with experience



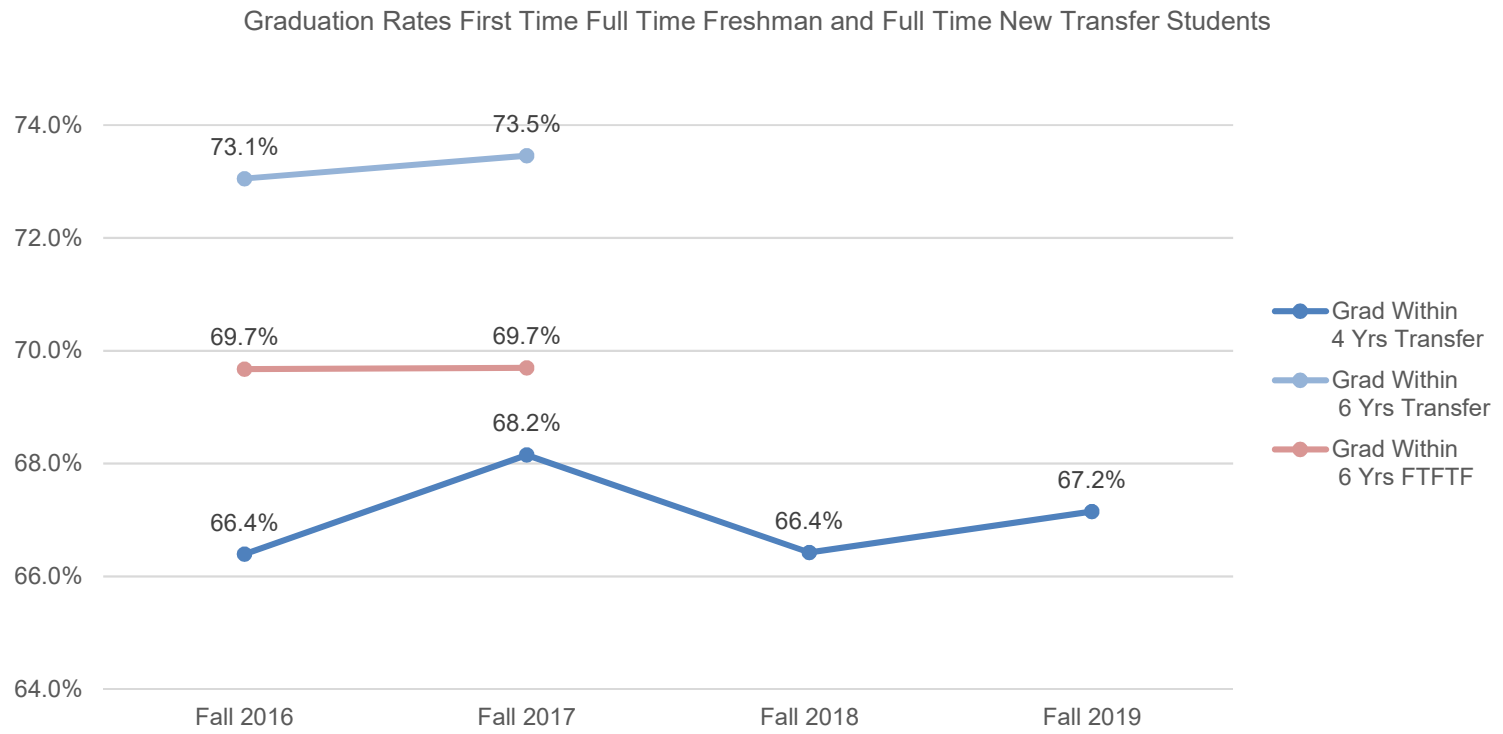
2012-2017 New Transfer Students

Retention rates for New Transfer students match those of their New Freshman counterparts



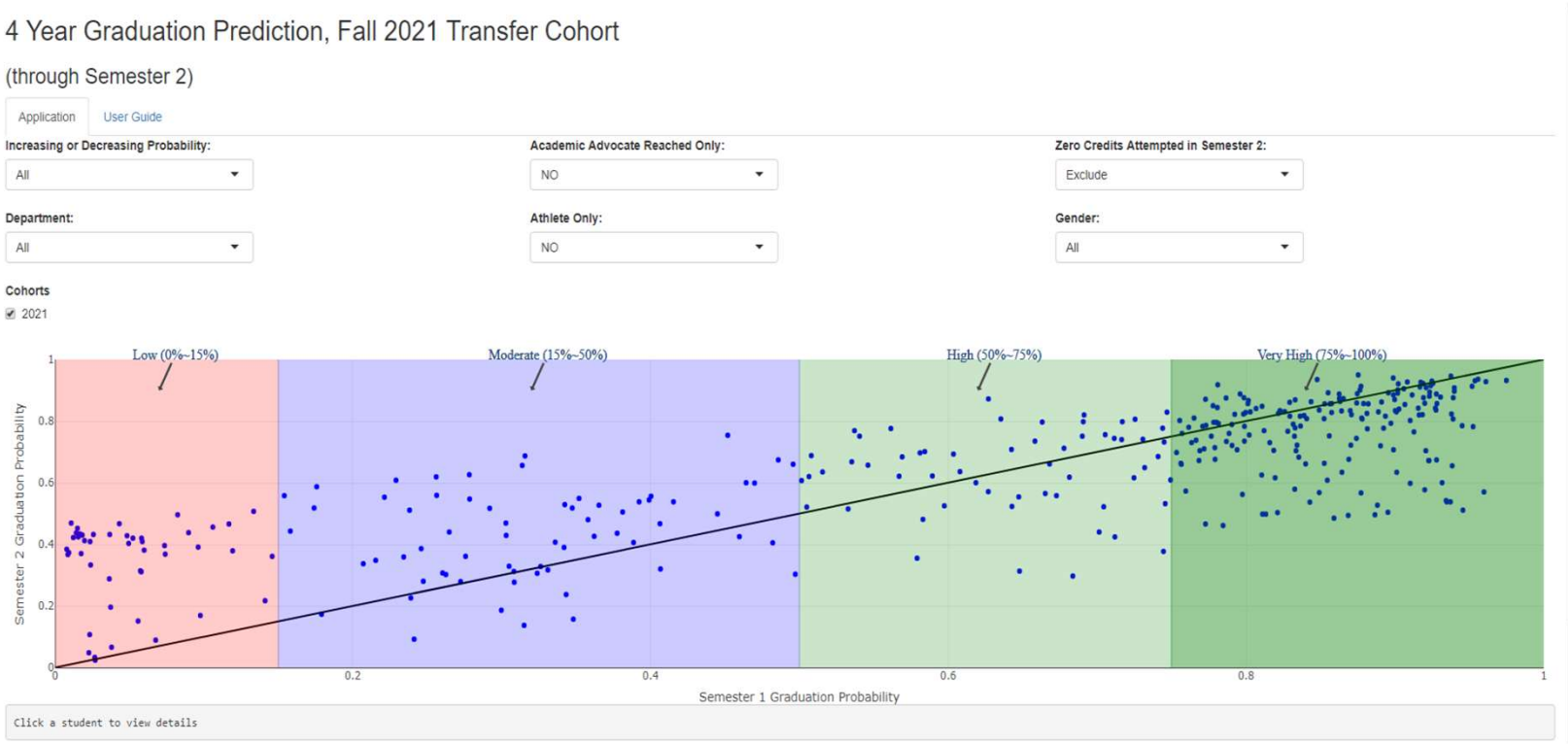
Data sourced from IRADS Census Reporting

4 Year Transfer graduation rates only slightly lag New Freshman and exceed Freshman at 6 years

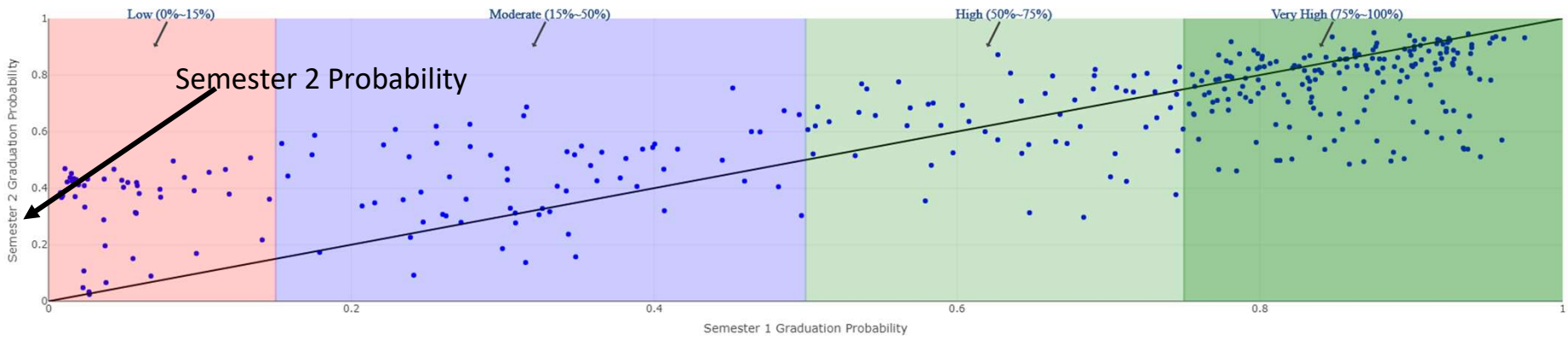


Data sourced from IRADS Census Reporting

The Tool: Predicting 1st and 2nd Semester Graduation Probability for New Transfer Students

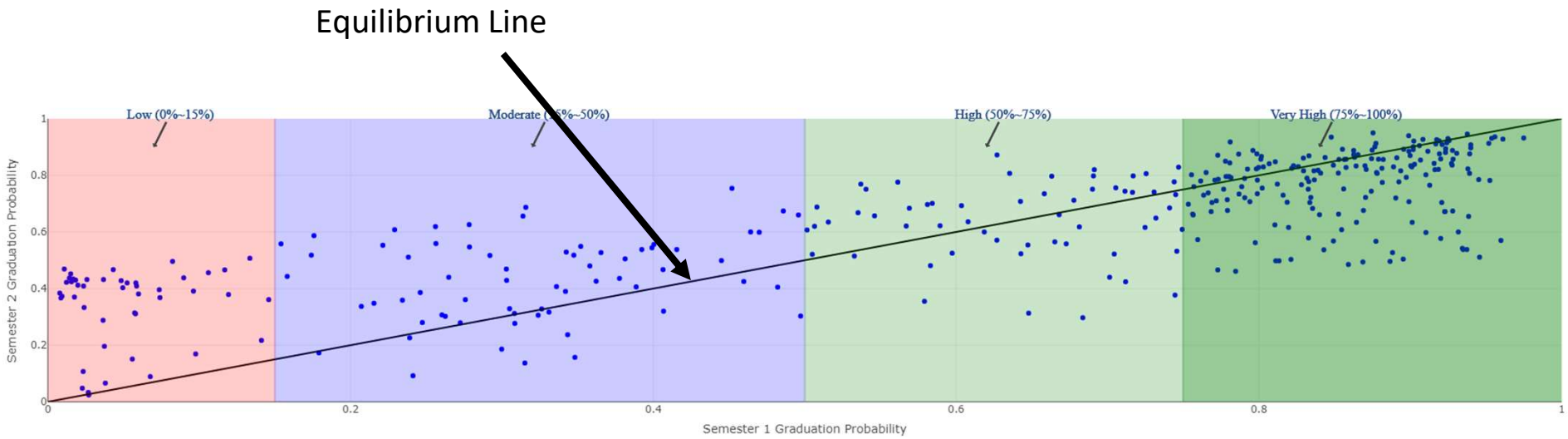


How to interpret the chart



Semester 1 Probability

How to interpret the chart



Points above this line are students whose probability of graduating in 4 years decreased from semester 1 to semester 2 and points below this line are students whose probability has increased.

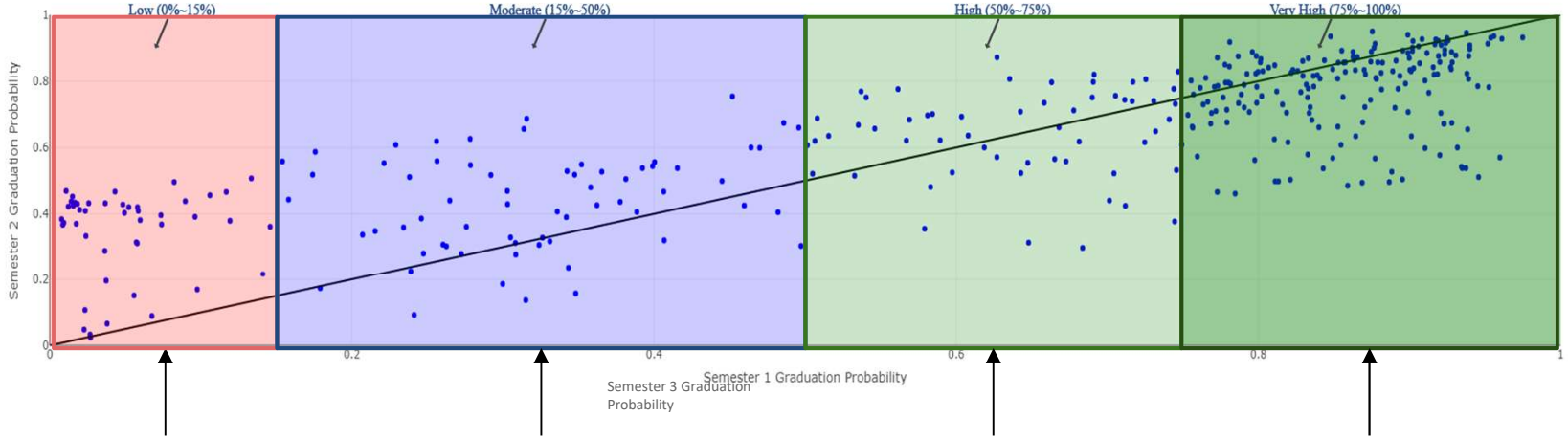
How to interpret the chart

3 % Graduation Rate

32 % Graduation Rate

71 % Graduation Rate

91 % Graduation Rate



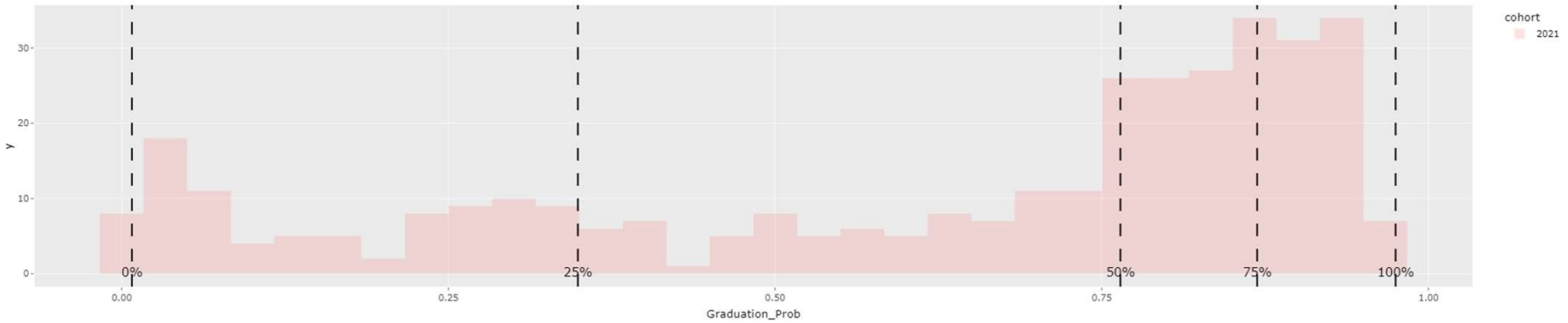
Low : 0%-15%

Medium : 15%-50%

High : 50%-75%

Very High : 75%-100%

Most Students are highly likely to graduate



How to use this tool

- Lasso or Box Select students in the moderate and/or low likelihood range.
- List of students can then be downloaded for contacting or intervention.
- Could eventually be tied into CRM action
- Students whose Semester 2 probability has dropped significantly (above the black line) can be individually selected and analyzed for risk factors.

A Use Case for an API?

Student Information

```
StudentKey :
************************************************************
DEMOGRAPHIC :
Ethnicity :
Gender :
Citizenship :
Residency :
************************************************************
ACADEMIC :
College : College of Arts, Humanities and Social Sciences
Major :
Cumulative GPA :  Rank :
Cumulative Credits Earned  Rank :
High School GPA  Rank :
Term GPA:  Rank :
Term Credits  Rank :
```

Box Select Information

```
Number of Students : 32
Average Cumulative GPA : 2.26
Average Term GPA : 1.2
Average HighSchool GPA : 3.59
Average Cumulative Credits Earned : 31.83
Average Term Credits Earned : 6.31
Average AP Credits : 5.56
Average AQI : 193.56
```

The Model: Ensemble Learning

- Different from the typical learning model we use, we combine the predictions from several models
- We use 2 different learning algorithms: Logistic regression, and random forest to generate the predictions.
- Each model generates a prediction and the predictions are then averaged. A cutoff determines the outcome of the prediction from the averaged models.
  - Final Prediction = (LG_result + RF_result) / 2
- The model makes predictions for both the second and third semesters of the cohort. While there is a difference in performance, the change in prediction becomes meaningful

UMBC

Model Training Performance on Semester 2 Data

	Non-Graduation Prediction		Graduation Prediction		
Model	**Precision**	**Recall**	**Precision**	**Recall**	**Accuracy**
Logistic Regression	.77	.50	.81	.93	.80
Random Forest	.78	.60	.84	.92	.82
Ensemble Learning Model	**.89**	**.50**	**.81**	**.97**	**.83**

*10 Fold Cross Validation

How does the model perform across early transfer semesters? It performs well!

Average Model Performance(Semester 1)

	Precision	Recall
Not Graduated in 4 years	80%	61%
Graduated in 4 years	79%	91%

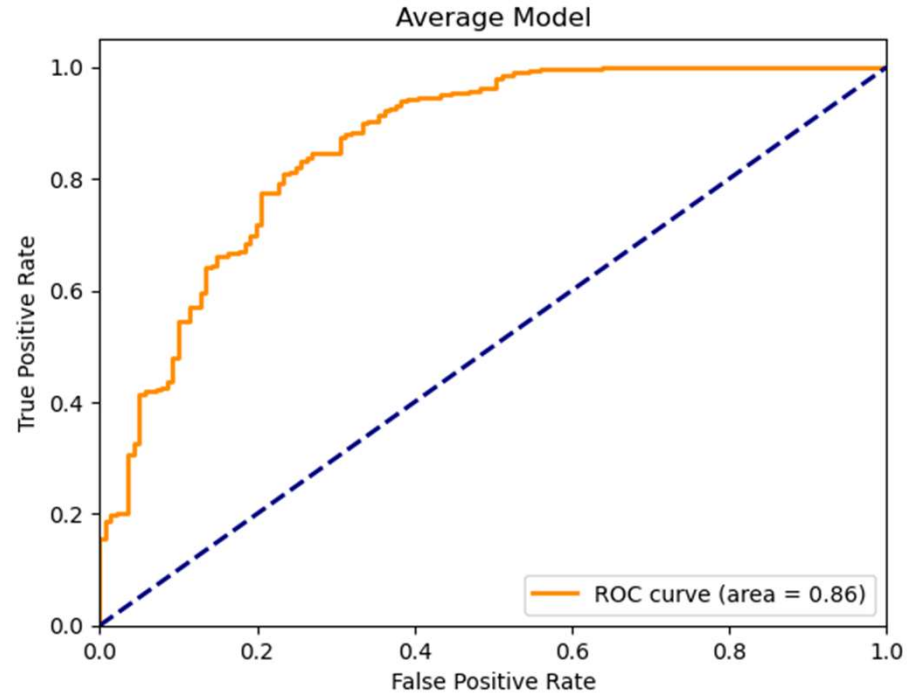
Average Model Performance(Semester 2)

	Precision	Recall
Not Graduated in 4 years.	82%	61%
Graduated in 4 years.	80%	92%

Average Model Performance(Semester 3)

	Precision	Recall
Not Graduated in 4 years.	86%	51%
Graduated in 4 years.	82%	96%

Average Model ROC Curve



Features in the Model

Demographic	Academic
Last College Type	Transfer GPA
Degree Earned	Cumulative Transfer Credits
	Credits taken at UMBC
	Cumulative GPA at UMBC
	Current Term GPA
	Repeated Courses
	Number of Withdrawn Courses

Feature Importance in RF