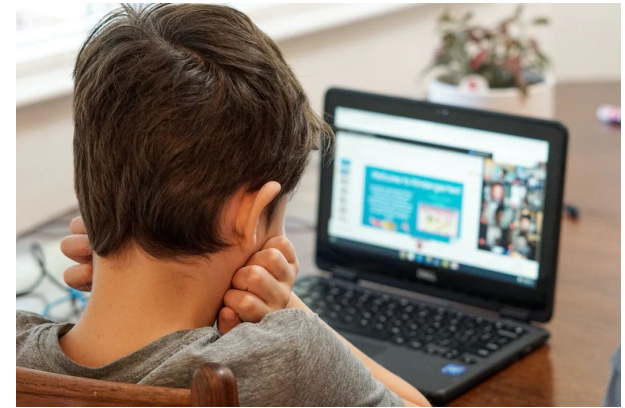


# Predicting Online Student Success

By Josh Johnson

## Online Learning

- In 2018, 35.3%: some or all classes online, And took them 16.6% online-only<sup>1</sup>
- After Covid-19: ????



*Photo by Thomas Park*  
[www.unsplash.com](https://www.unsplash.com)

## Retention

- Online university courses have a 10-20% higher dropout rate than in person
- Other online courses have a drop out rate between 40% and 80%<sup>2</sup>

1. <https://nces.ed.gov/fastfacts/display.asp?id=80>

2. [https://journals.sagepub.com/doi/pdf/10.1177/2158244015621777#:~:text=Online%20courses%20have%20a%2010,Smith%2C%202010\).](https://journals.sagepub.com/doi/pdf/10.1177/2158244015621777#:~:text=Online%20courses%20have%20a%2010,Smith%2C%202010).)

# Can Predictive Modeling Improve Student Success?



Only if we know who needs it!

Photo by Frank Romero [www.unsplash.com](https://www.unsplash.com)

# The Solution

**Identify students in danger of failing.**

- Halfway through the course
- In time for interventions to be successful

# The Data

## Online University: Years 2012/2013 - 2013/2014

- 24743 registrations
- 22424 unique students
- 7 course modules
- 22 cohorts

**10,655,280 Student VLE interactions**

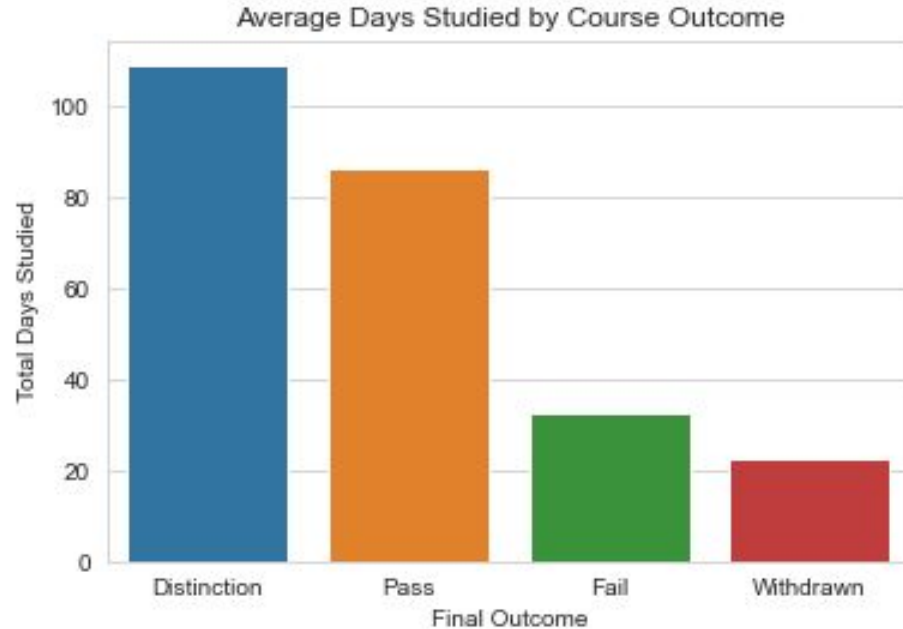
# Features to Model



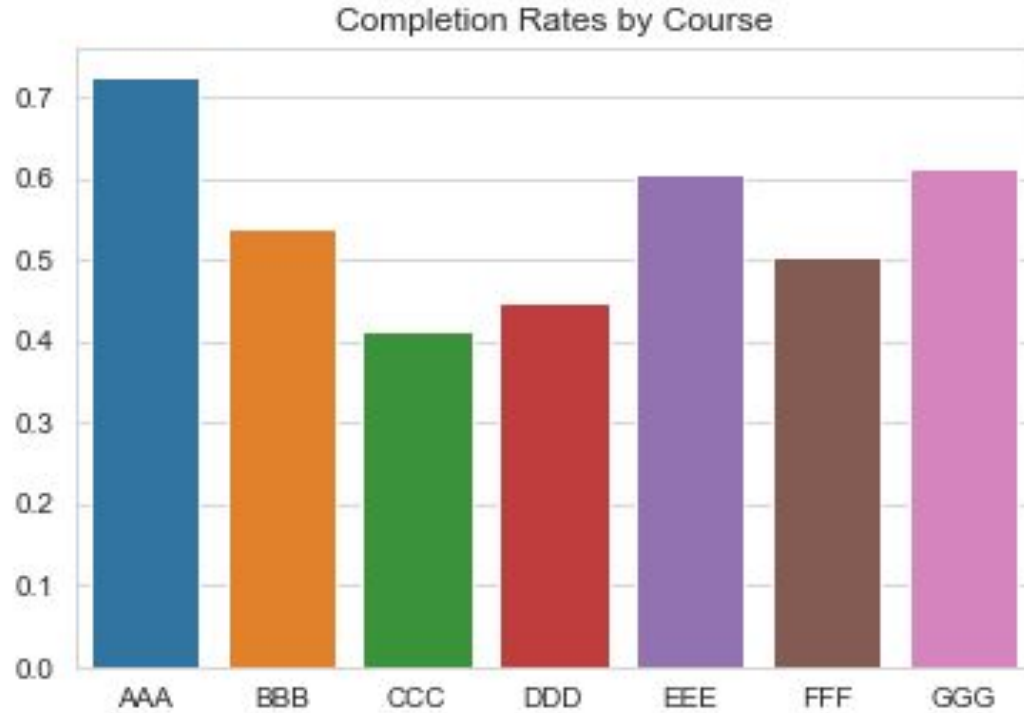
Image by Steinar Engeland, courtesy of [Unsplash.com](https://unsplash.com)

1. Average assessment scores
2. Number of assessments completed
3. Number of days studied
4. Number of activities engaged
5. Total number of clicks
6. Times repeated the course

# More Days Studying Correlates to Success

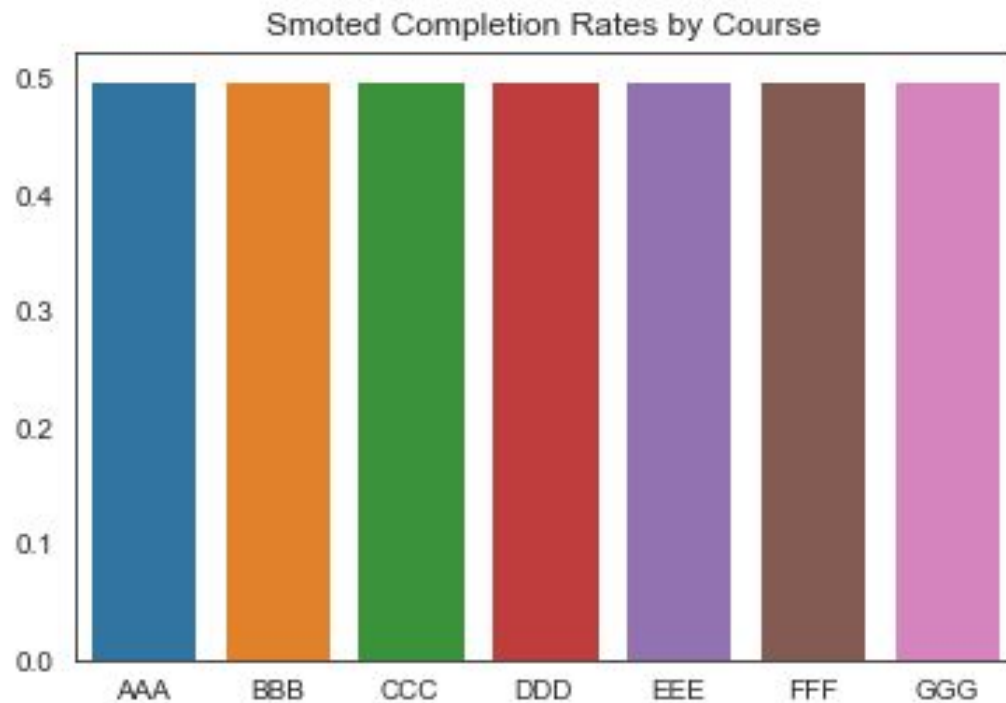


# Some Courses are Harder Than Others.



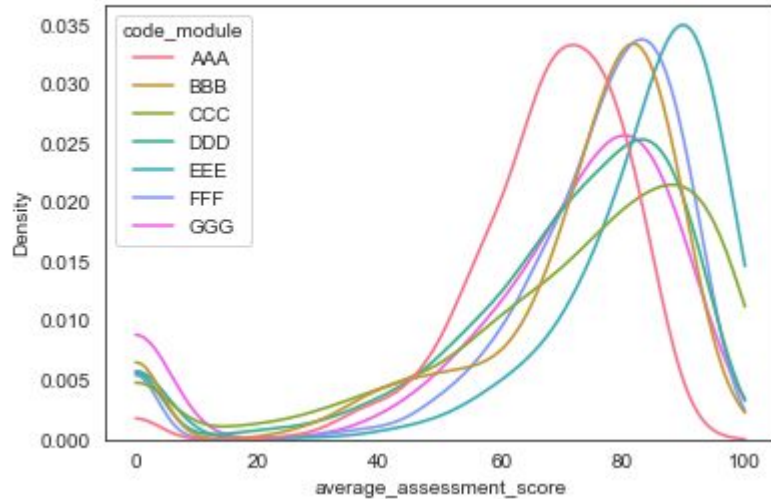


...But My Model Shouldn't Know That.

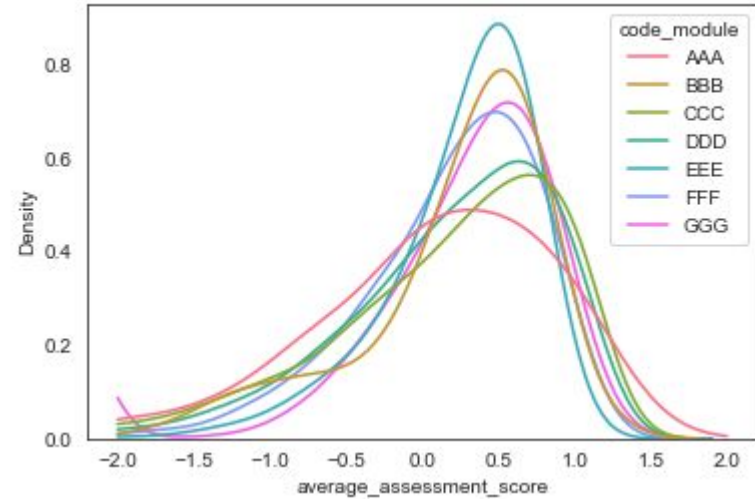


# Distribution of Assessment Scores by Course

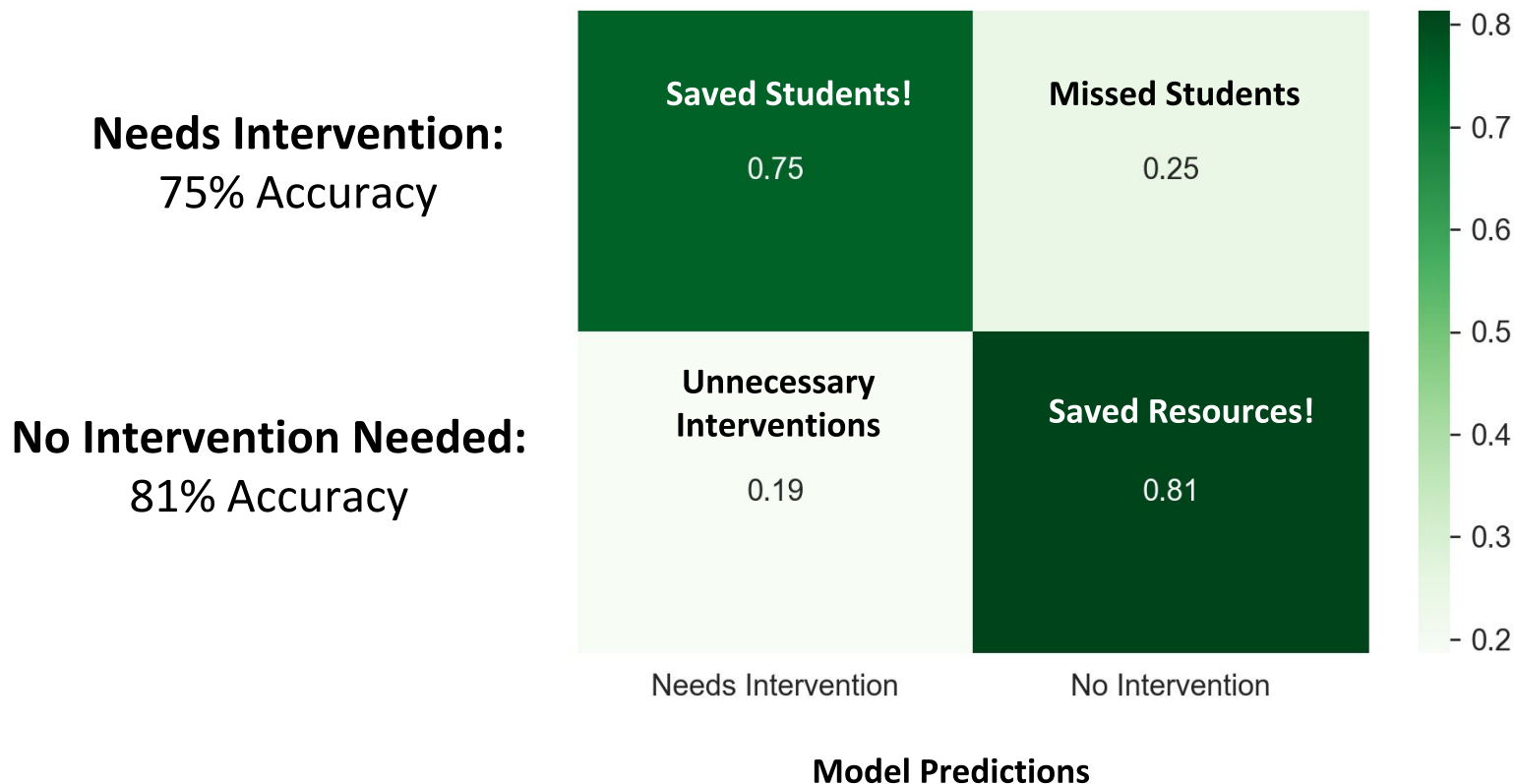
Comparative Distribution of average\_assessment\_score Between Courses



Scaled Distribution of average\_assessment\_score Between Courses



## XGBoost Accuracy After the First Half of Courses: 79%



## Next Steps:

- Try more model types to improve accuracy
- Try different prediction windows
- Evaluate model results to find insights to help more students succeed
  - For instance, will doing fewer activities but over more days still help you succeed compared to doing more on fewer days?
- Apply model to larger and more diverse datasets.
- Deploy model in a live scenario.

# Contact

**Josh Johnson**

**LinkedIn:** <https://www.linkedin.com/in/josh-johnson-049a2619/>

**Github:** <https://github.com/Caellwyn>

**Email:** [caellwyn@gmail.com](mailto:caellwyn@gmail.com)