OU Student Performance Analysis

Sharon Li

Table of Contents

1

Introduction

Background

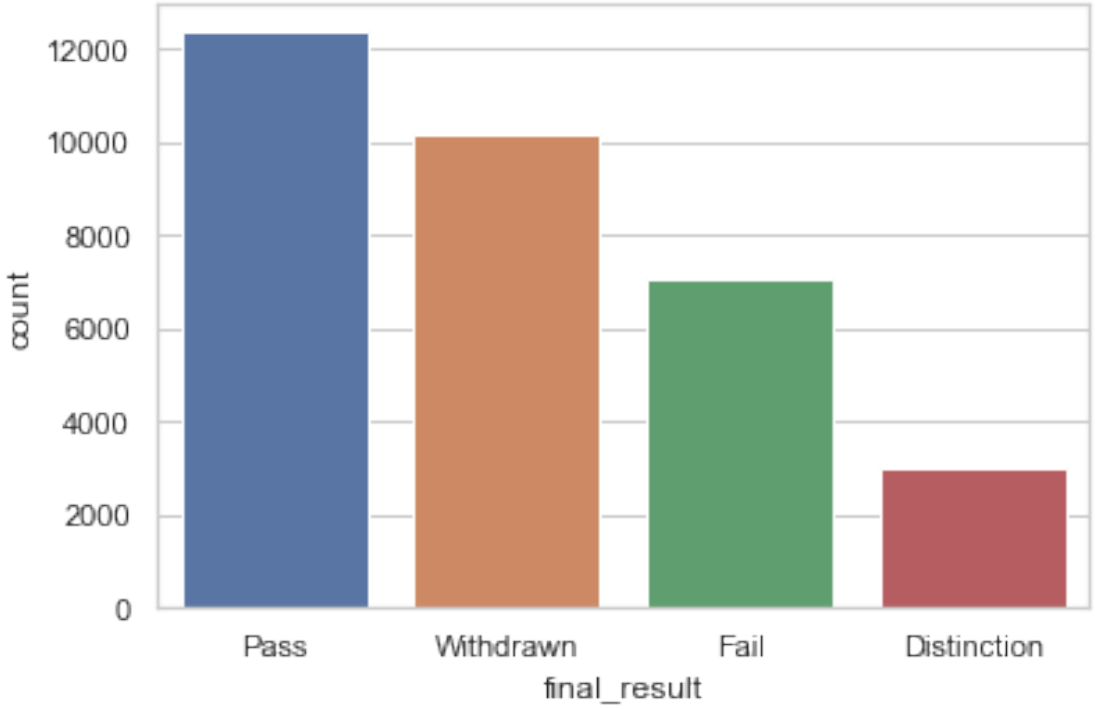
Virtual Learning Environments

170,000

Background

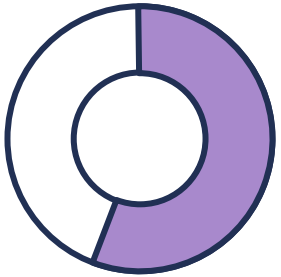
Timeline of a module

Background



However .....

Over half of students ended with **withdraw** or **fail**



52.8%

Vision & Goal

Visions:

- Help students get a better learning experience.
- Reduce fail and withdraw rate.

Goals:

- Identify potential causes of fail or withdraw a class.
- Come up with actionable suggestions.
- Come up with method to measure the impact of the solutions.
- Deliver reusable models to predict future student assessment results.

Stakeholders

- Students
- Course Administrators
- Course Platforms
- Researchers

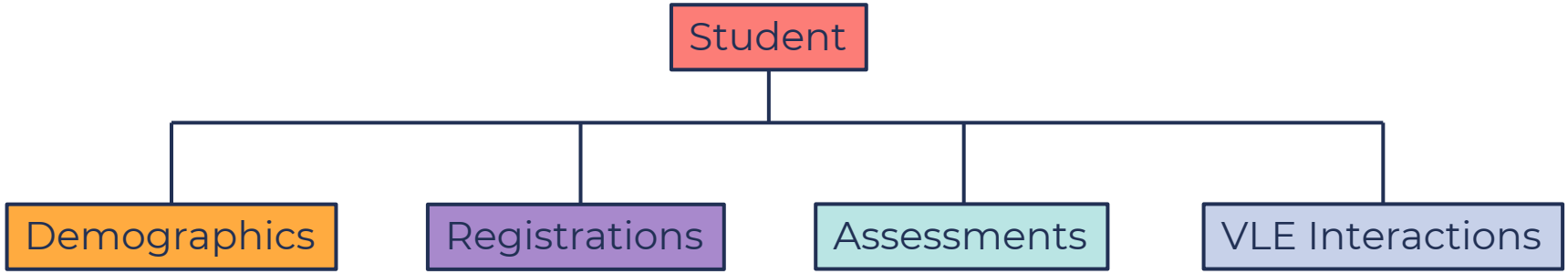
Data Science

Steps

1. Dataset Overview
2. Metrics Definition
   Feature Engineering
3. Data Cleaning
4. Machine Learning

Dataset Overview

Overall Dataset Structure



Metrics Definition

Weighted Assessment Score

Measure student overall performance of assessment

Exam Score

Measure student exam performance

Submission Before Assessment Due Date

Measure student timeliness of submission

Sum of clicks

Measure the usage of the VLE material

Metrics Definition

VLE Count

Measure the content volume of each module in each presentation

Unregistration Time

Measure how long student take to unregister a course after their registration

Module withdraw or fail rate

Measure the withdraw or fail rate of a module

Module withdraw rate

Measure the withdraw rate of a module

Module fail rate

Measure the fail rate of a module

Feature Engineering – Student Result

studentInfo
id_student
code_module
code_presentation
final_result
…

- Module withdraw or fail rate
- Module fail Rate

Feature Engineering - Assessments

assessments
code_module
code_presentation
id_assessment
date
weight
...

studentAssessmentid
assessment
id_student
date_submitted
score
...

Join by id_assessment

- Weighted Assessment Score
- Exam Score
- Submission Before Assessment Due Date

Feature Engineering - VLE (Virtual Learning Environment)

studentVle
code_module
code_presentation
id_student
id_site
sum_clicks
...

- Sum of clicks (per module per presentation per student)

Vle
id_site
code_module
code_presentation
...

- VLE counts (per module per presentation)

Feature Engineering - Registrations

studentRegistration
code_module
code_presentation
id_student
date_registration
date_unregistration

- *Unregistration Time*
- *Module Withdraw Rate*

courses
code_module
code_presentation
module_presentation_length

Data Cleaning

Wrong Value

imd_band
90-100%
20-30%
30-40%
50-60%
50-60%
80-90%
30-40%
90-100%
70-80%
?

Should be '> 100%'

Missing Value

code_module_cat	0
code_presentation_cat	0
id_student	0
num_of_prev_attempts	0
studied_credits	0
wf_rate	0
weighted_score	6668
weighted_day_before_ass	6668
exam_score	25711
sum_click	3302
vle_cnt	0
module_w_rate	0
day_un	20525
final_result_cat	0
gender_cat	0
region_cat	0
highest_education_cat	0
imd_band_cat	0
age_band_cat	0
disability_cat	0

dtype: int64

Didn't take the assessment

Didn't click material in VLE

Didn't unregister

Impute with -1

Data Cleaning

Outliers

- studied_credits > 250
- sum_click > 8
- day_un > 450
- weighted_day_before_ass > 6 or weighted_day_before_ass < -5

Remove them

Data Cleaning

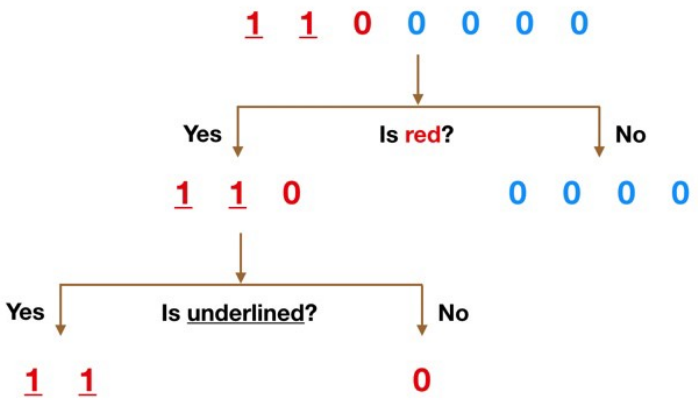
Label Encoding

code_module
AAA
BBB
CCC
DDD
EEE
FFF
GGG
……

code_module_cat
0
1
2
3
4
5
6
……

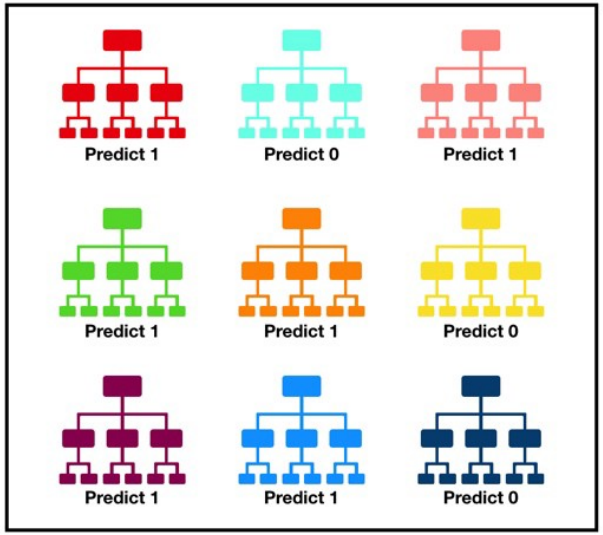
Feature Selection

Decision Tree



Random Forest

A large number of relatively uncorrelated trees.



Tally: Six 1s and Three 0s

Prediction: 1

Feature Selection

Demographical Features

+

Behavioral Features

Training Set: 80%
Test Set: 20%

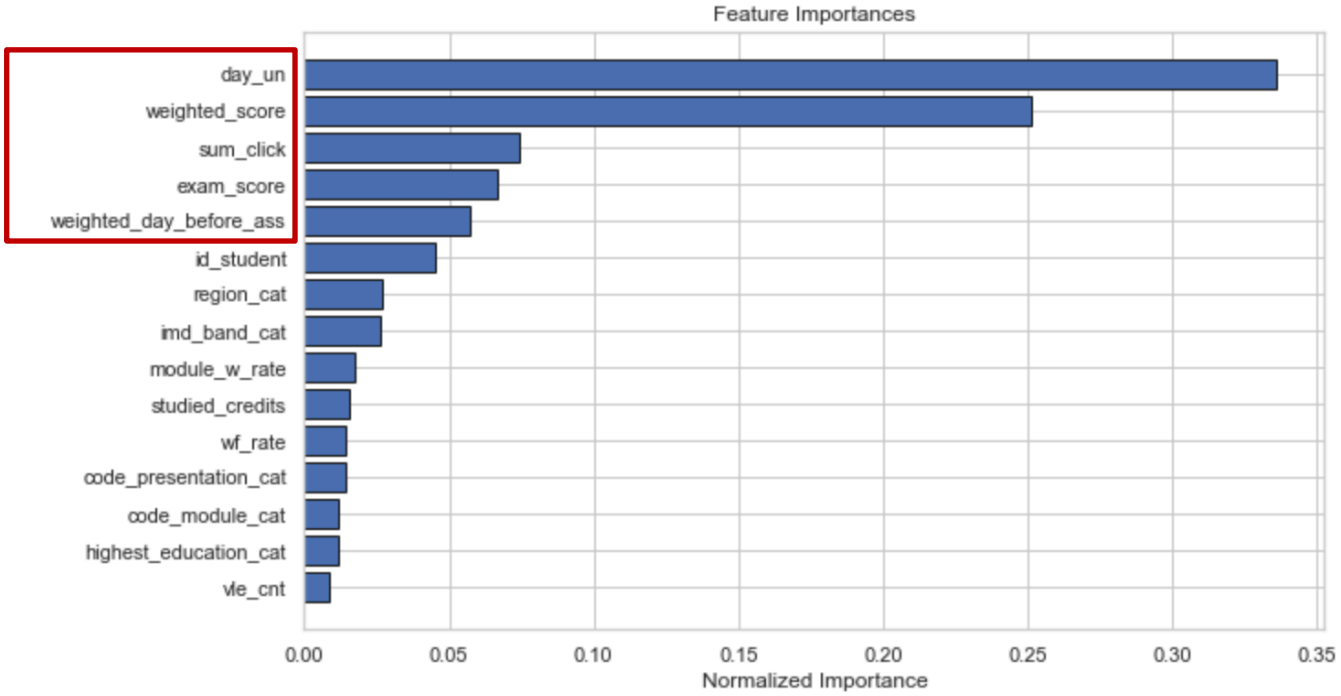
↓

Random Forest Classifier

Accuracy on Test Set: **87.2%**

Feature Selection

Could be used as ***proxies*** of student final result



Forecasting

- Directly correlated the the final result
- We can't know them in advance

Demographical Features

+

~~Behavioral Features~~

↓

Random Forest Classifier

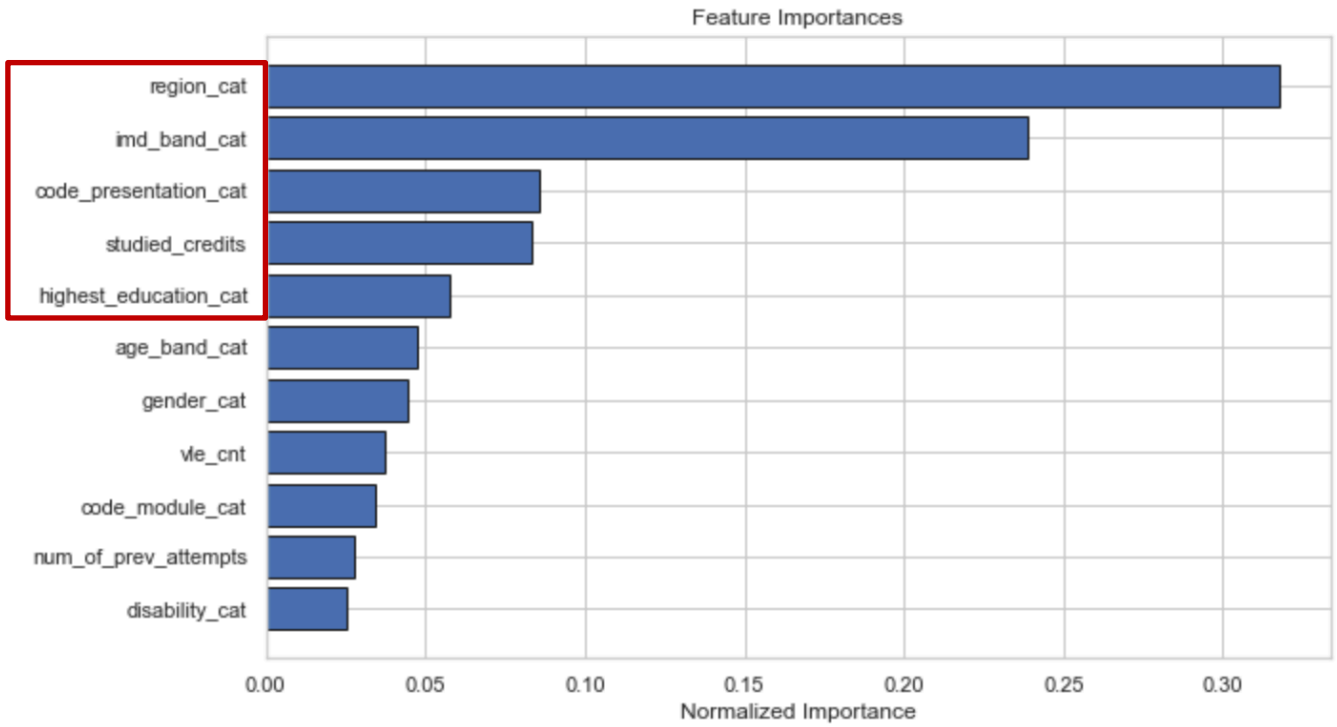
Training Set: 80%
Test Set: 20%

Accuracy on Test Set: **37.5%** ☹

Need more information to forecast!

Forecasting

Let's explore how they could impact on student final result!



Analysis

A Timeline Always Works Well

Demographical Features

Behavioral Features

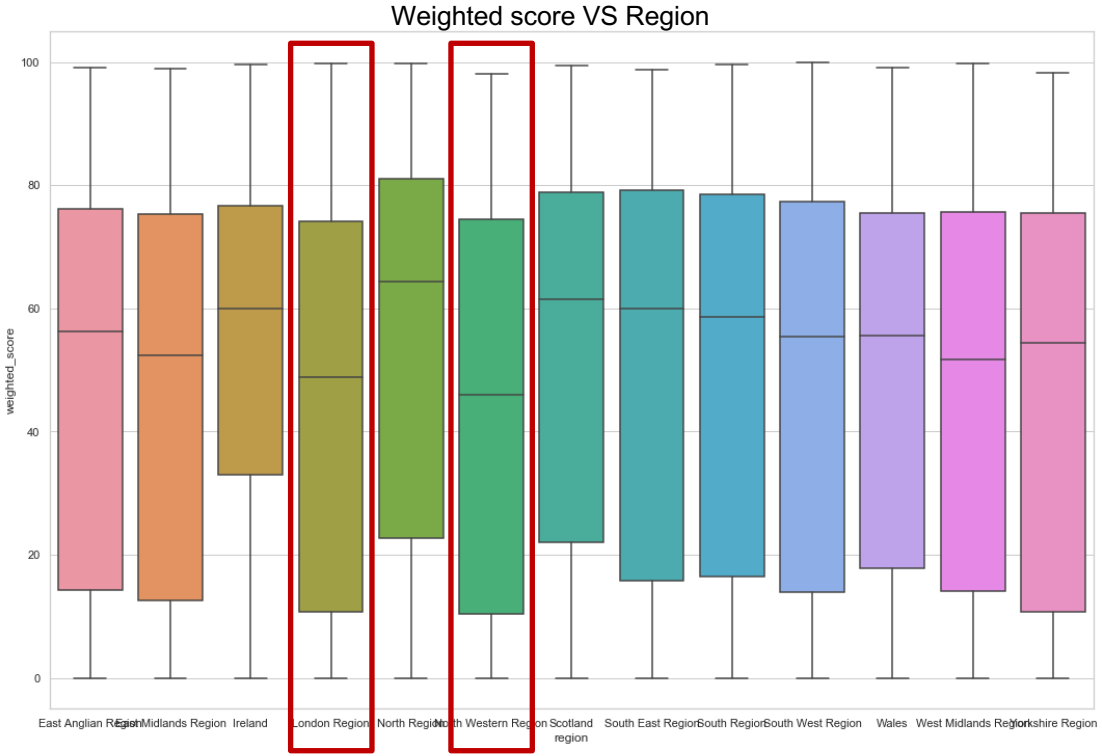
Measurement

- How they impact the final result?
- How to deal with them?

- How to prove the strategies we take will work?

Demographical Features - Region

The weighted score in **London Region** & **North Western Region** is lower than others.

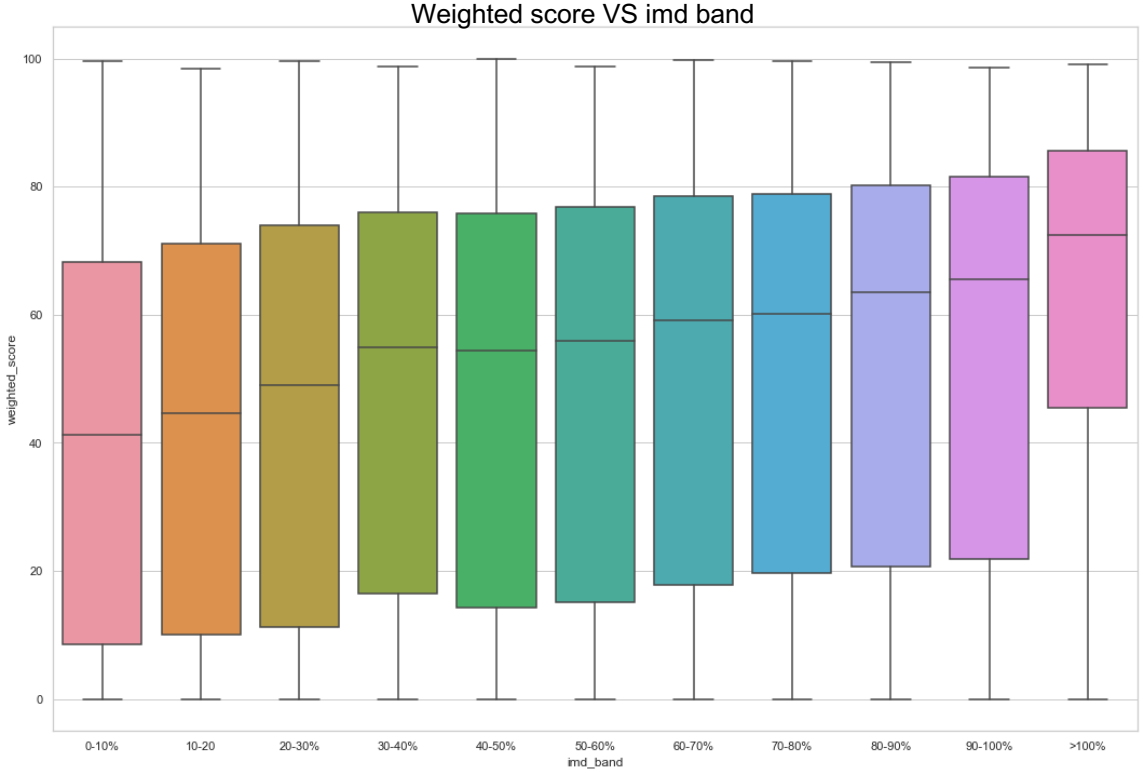


Solutions:

Further investigate the policy, culture, GDP, overall education level... in these 2 regions to figure out why that happen.

Demographical Features – imd band

The **less deprived** the areas where the students live, the **more likely** they are to **succeed**



Solutions:

Offer free basic level educational resource in the deprived regions.

Demographical Features – code presentation

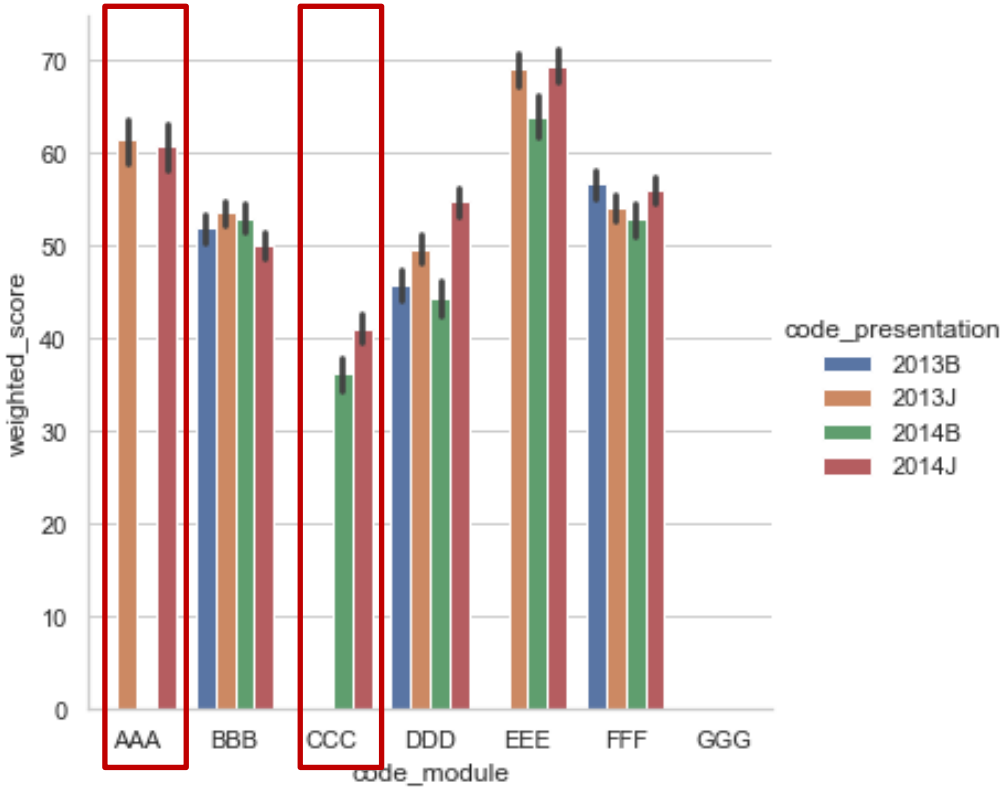
Student didn't perform as good as other presentations in the 2014B

That looks wired!

Let's **dive deeper!**

Demographical Features – code presentation

- Module ‘AAA’ (higher average score) didn’t hold in the 2014B
- Module ‘CCC’ (lower average score) started from 2014



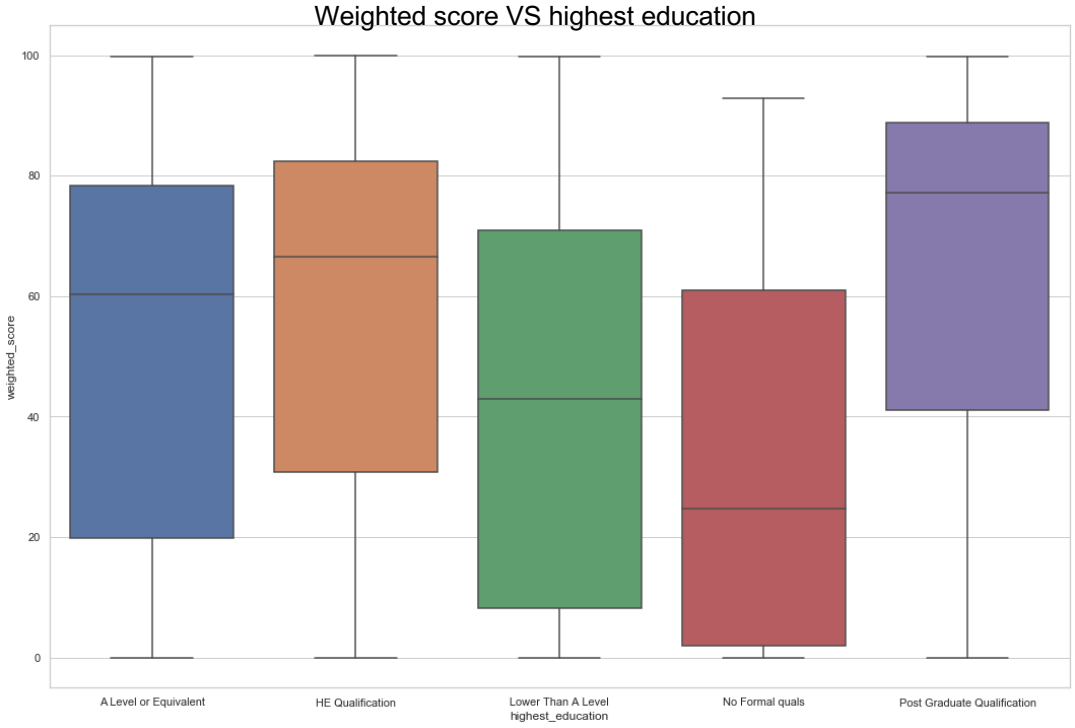
the code_presentation itself doesn't affect student result, it affects the weighted score simply because of the **course schedule**.

Demographical Features – highest education

The **higher education level**, the **more likely** they are to **succeed**

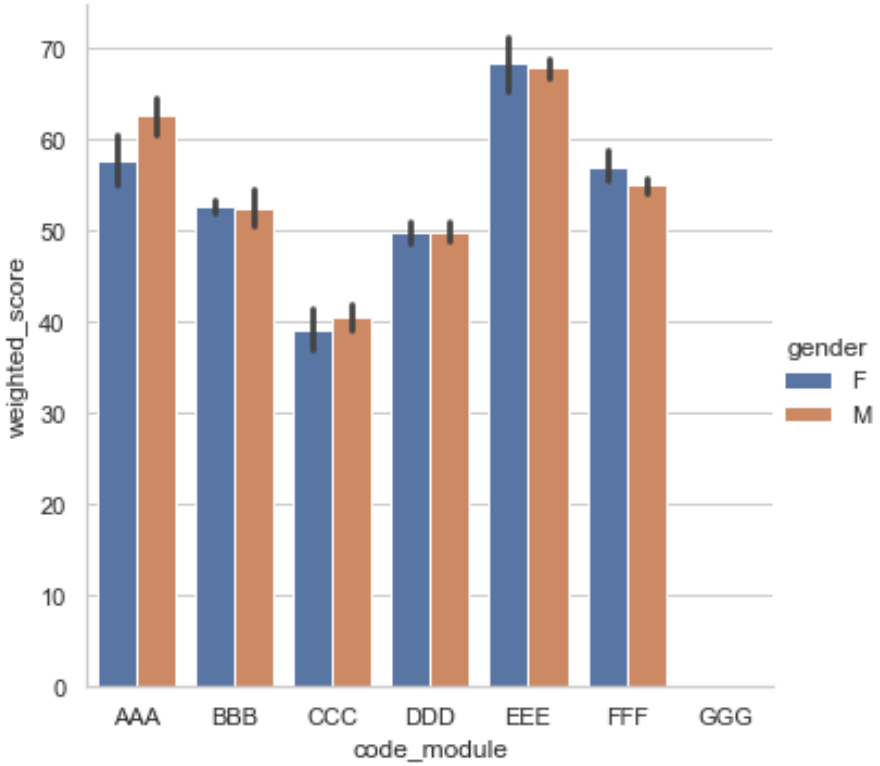
Solutions:

- **Direct survey** on students with lower education level.
- **Recommend** more basic level courses to them.
- Give **promotions** to encourage them take the courses of suitable difficulty.
- Allocate **tutors** to them

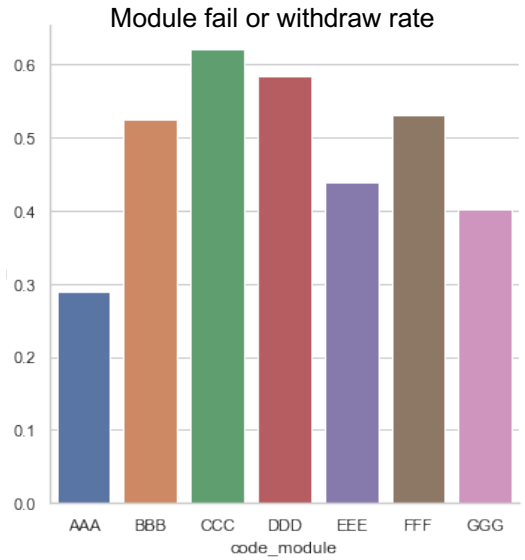


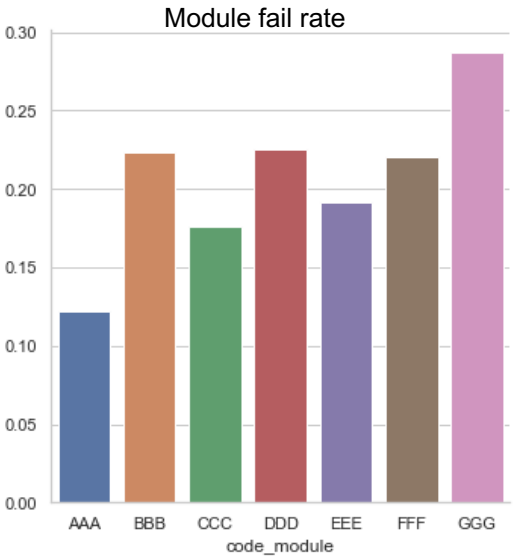
Demographical Features – Gender

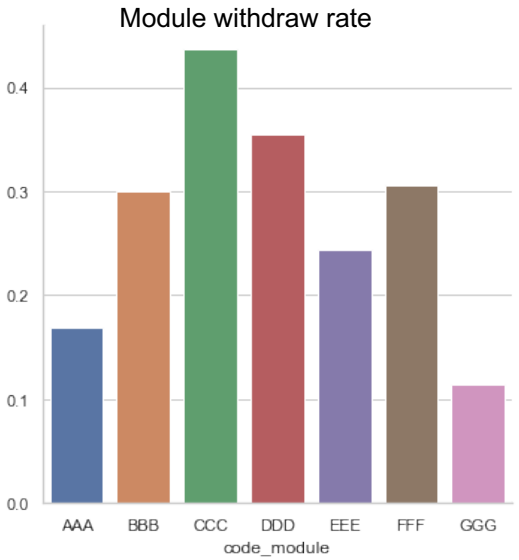
Gender does not seem to have a significant impact.



Demographical Features – Module

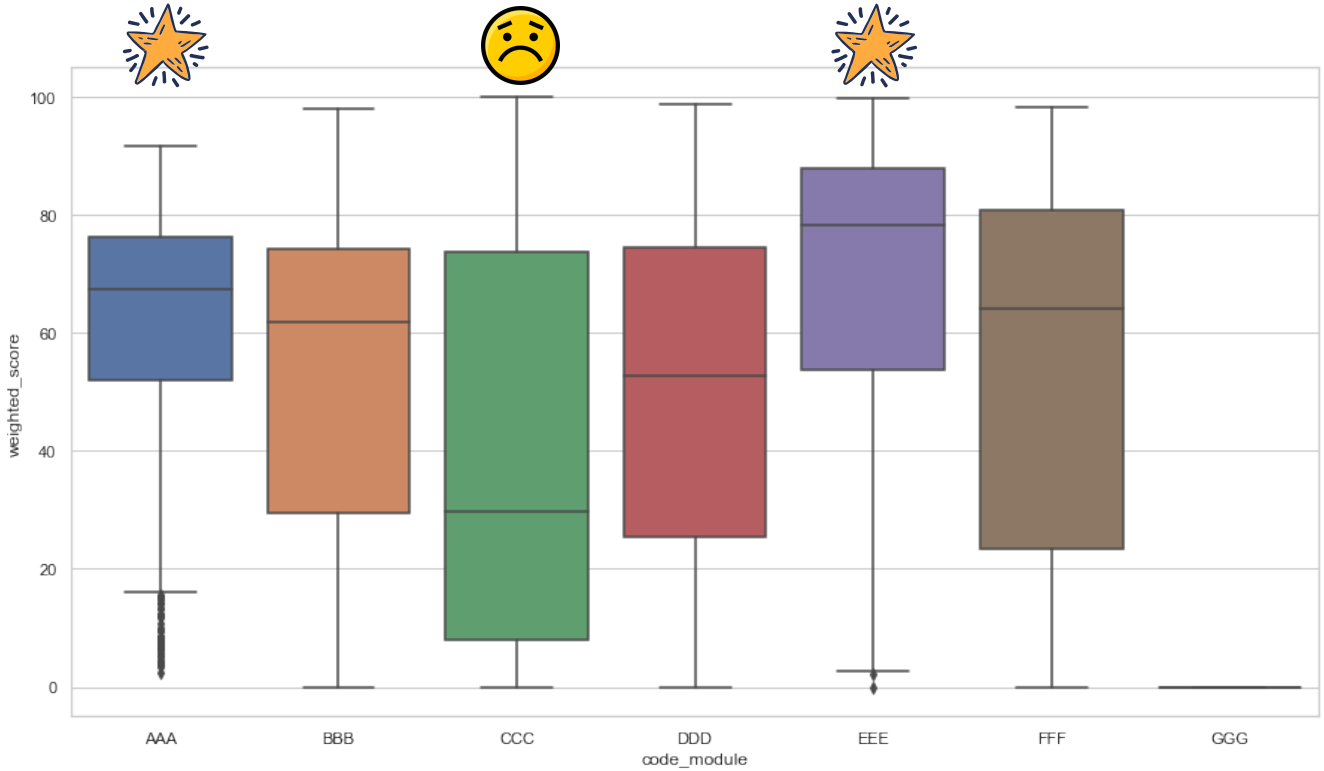






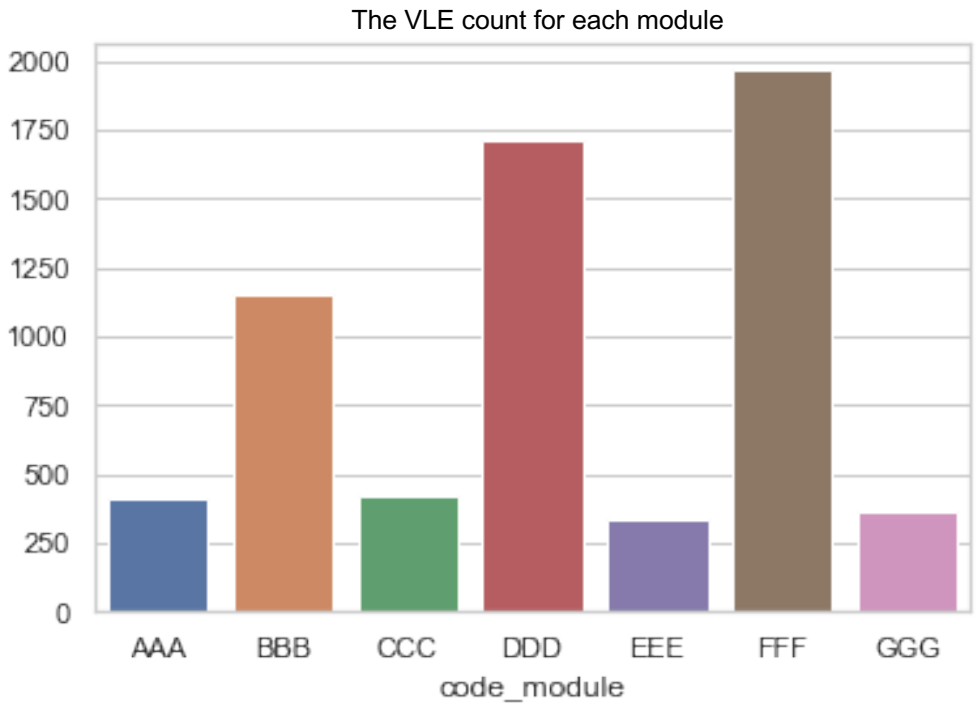
- Module ‘GGG’ may be both compulsory and difficult, for its low withdraw rate and high fail rate.
- Module ‘CCC’ may be a hard course or have a low quality, for the withdraw rate is the highest.
- The fail or withdraw rate of module ‘BBB’, ‘DDD’ and ‘FFF’ is also high.

Demographical Features – Module



- How did the module 'GGG' be evaluated? Look at the data, both assessment score and exam score is 0.
- The result of module 'CCC' is the worst!
- The result of module 'BBB', 'DDD' and 'EEE' is not as good as module 'AAA' and 'EEE'.

Demographical Features – Module



The workload of module 'BBB', 'DDD' and 'FFF' is **bigger** than 'AAA' and 'EEE'

Solutions:

Split each of the module 'BBB', 'DDD' and 'FFF' into two modules(eg. module 'DDD1' module 'DDD2'), and each new module still take one presentation. Try to **reduce the workload** in each module and improve result.

Demographical Features – Module

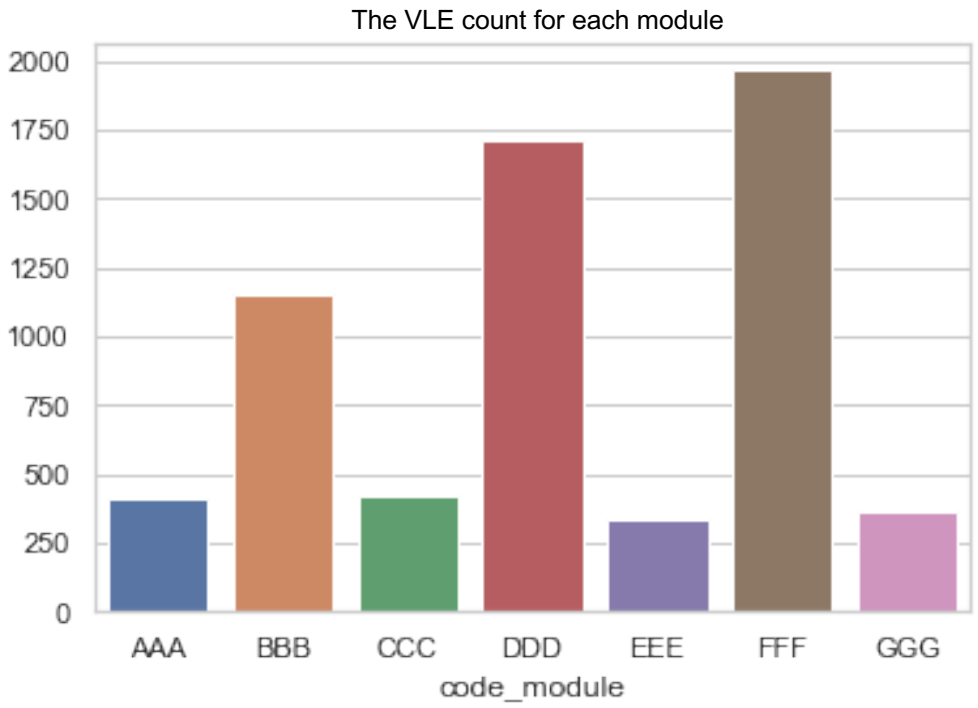
Module 'DDD'

Module 'DDD1' — 2021B

Module 'DDD2' — 2021J

Reduce the workload

Demographical Features – Module

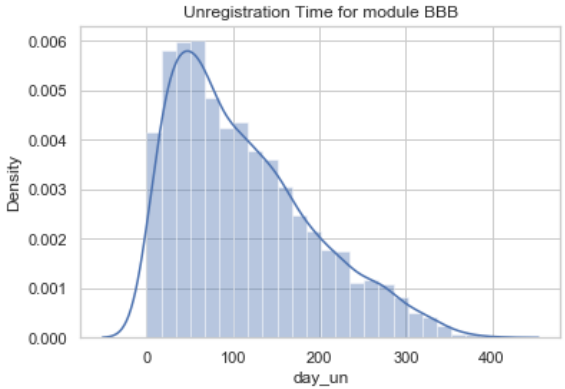


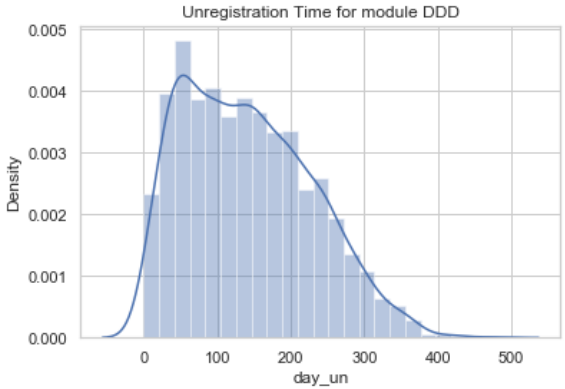
The workload of module 'CCC' and 'GGG' is **almost the same** with 'AAA' and 'EEE'

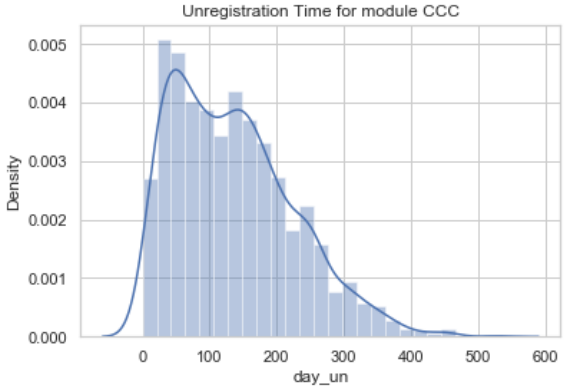
Solutions:

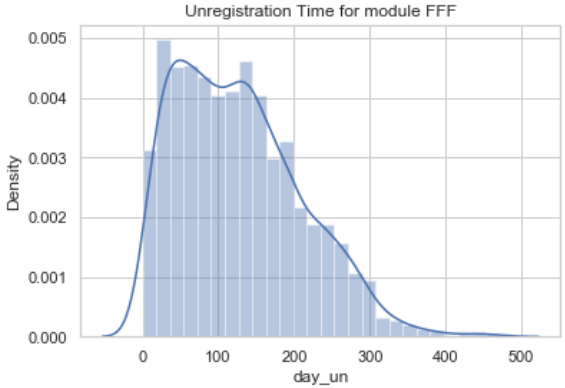
- Collaborate with the module providers to investigate what the problems are. Too difficult? The way of teaching? Quality of material? Mentorship?
- Come up with possible strategies according to the problem. (eg. Hierarchical teaching)

Behavioral Features – Unregistration Time







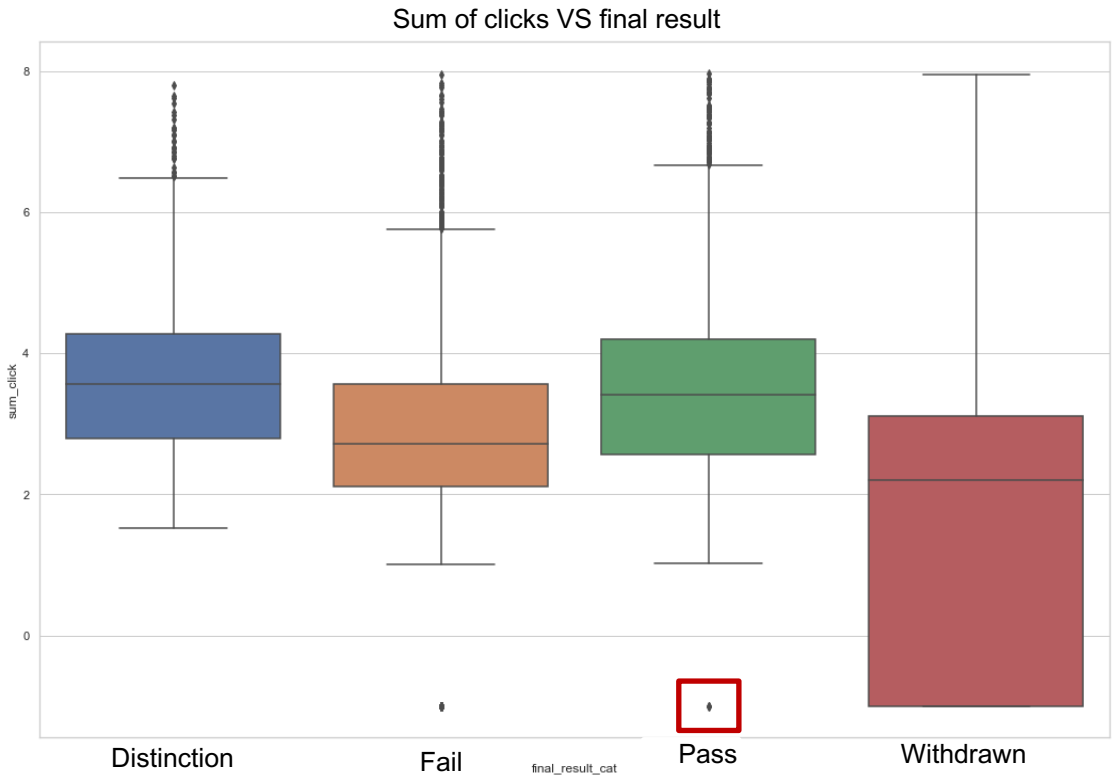


In most case, student **drops** a module **after 30 – 50 days** of registration.

Solutions:

- Direct Survey: Ask students why withdraw once they click the withdraw button.
- Give them incentives to continue the course when they decide to withdraw. (eg. More credits if they chose to continue)

Behavioral Features – Sum of Clicks



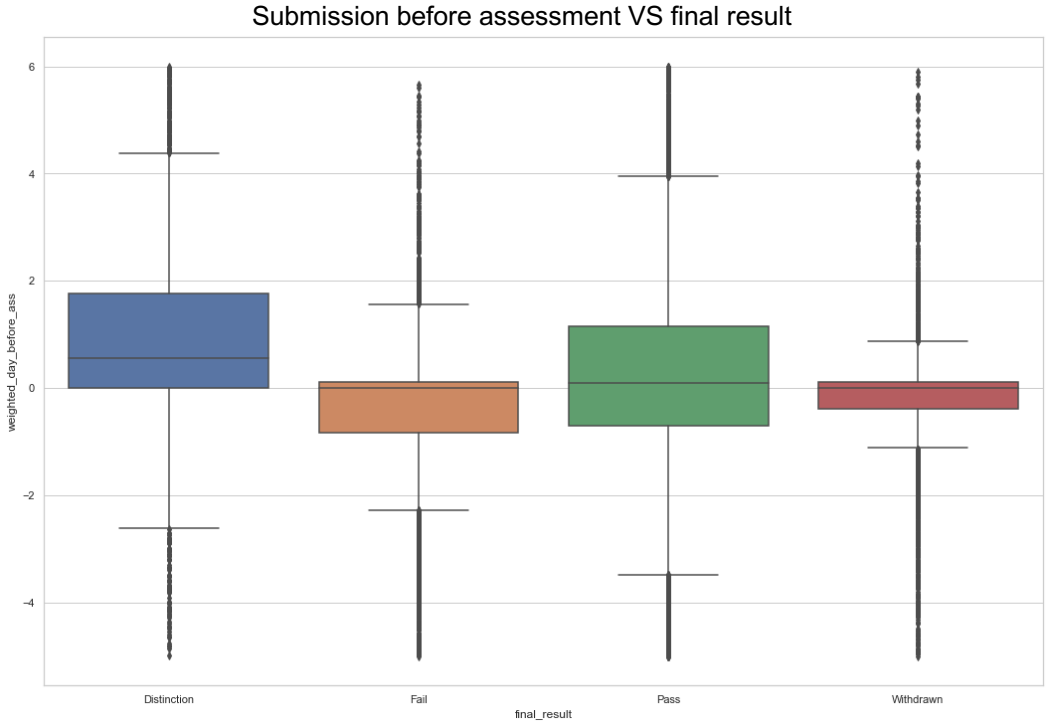
More clicks on the material in the VLE, **better result**.

Students pass but also no clicks

Solutions:

Incentive students to let them go to the VLE more and use the material more (eg. Give them credits once they finished 2 hours learning in the VLE per week).

Behavioral Features – Submission Before Assessment Due Date



Earlier submission, better result.

Solutions:

Incentive students to finish the assessment and submit earlier. (Give them credits as the same amount of their submission day before assessment due date.)

Behavioral Features – Measurement

The overall idea:

1. Implement the strategies.
2. Track metrics we care about.
3. See if the metrics go where we want.

- Cannot use the **final result** as the metric, because we need to wait too long time to get it.
- Instead, use the **weighted assessment score, submission before assessment** and **sum of clicks** as proxies of the final result.

Conclusion

4

Summary

Defined relevant **metrics** that could impact the student final result.

A

B

Found **potential causes** of low performance from both demographical and behavioral perspective.

C

D

Made corresponding **solutions** and **measurement methods**.

Need more data and varied features to **predict future result**.

Limitation & Future Work

The dataset is too old

Need more varied features

Model selection (parameters, supervised VS unsupervised)

Thanks

Do you have any questions?

Xiaoran.Li@ur.Rochester.edu
+1 5857641431
https://www.linkedin.com/in/xiaoransharonli/