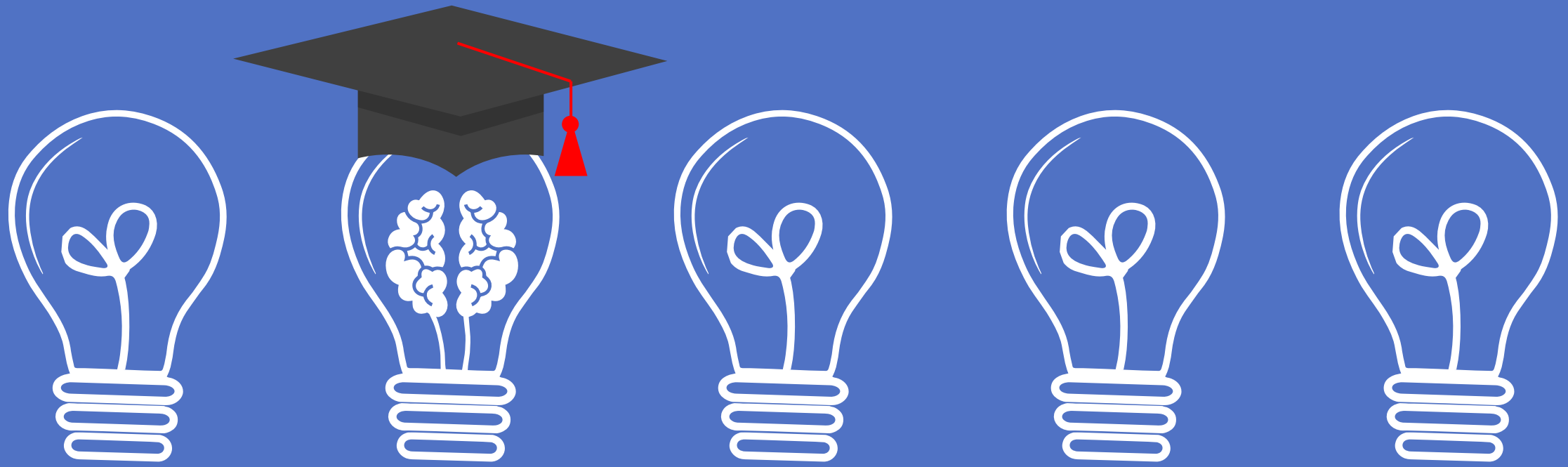


Predict Students' Dropout and Academic Success

IRONHACK Final Bootcamp Project

Mojgan Shabani

About Project

The Data Source for this project has been provided from SATDAP - Capacitação da Administração Pública under grant POCI-05-5762-FSE-000191, Portugal

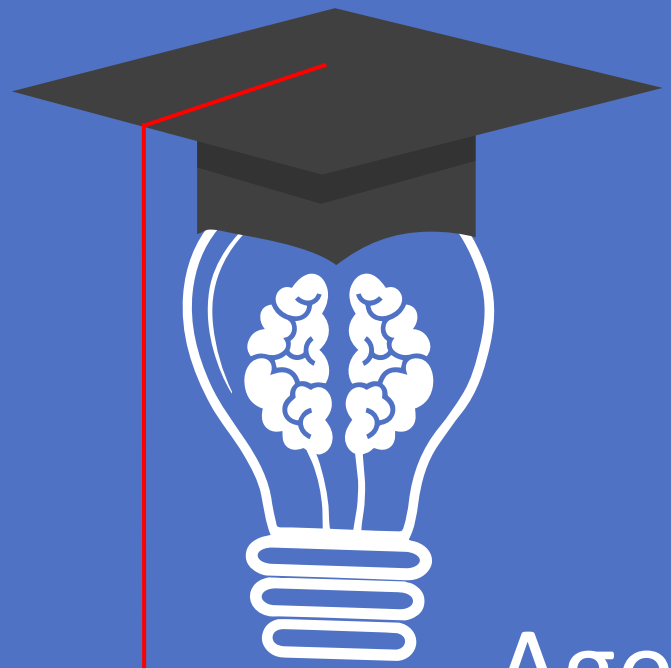
Overview

The dataset was created in a project that aims to contribute to the reduction of academic dropout and failure in higher education, by using machine learning techniques to identify students at risk at an early stage of their academic path, so that strategies to support them can be put into place.

Goal

Predict students' dropout and academic success

https://www.kaggle.com/datasets/ankanhore545/dropout-or-academic-success?select=Dropout_Academic+Success+-+Sheet1.csv



Agenda

01 Data Exploration

- ✓ Demographic Information
- ✓ Course popularity.
- ✓ Dropout & Graduate

02 Prediction

- ✓ Dropout Prediction.
- ✓ High score Features in predictions

03 More insights

- ✓ Comparing High scores Features in predictions
- ✓ Dropout in different Majors

04 conclusion

- ✓ Conclusion & solutions

Prediction Process

01 Cleaning Data

We have around 4442 samples with 40 Features. 18 Categorical and 22 Numerical.

02 Removed Multicollinearity

With VIF method remove high multicollinearities

03 Normalized data

Normalized data with OneHotEncoder and MinMaxScaler

04 Up Sampling

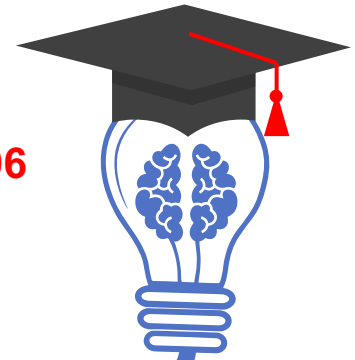
We had class imbalanced, so up sampling for prediction have been performed.

05 Prediction

Preformed 4 different types of prediction's Models.

06 Data Visualization

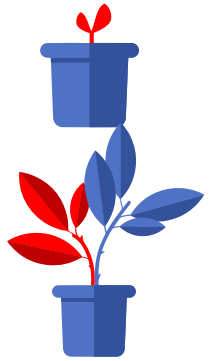
Using Tableau for Data visualization and gathering insights.



Demographic Information

65%

Female



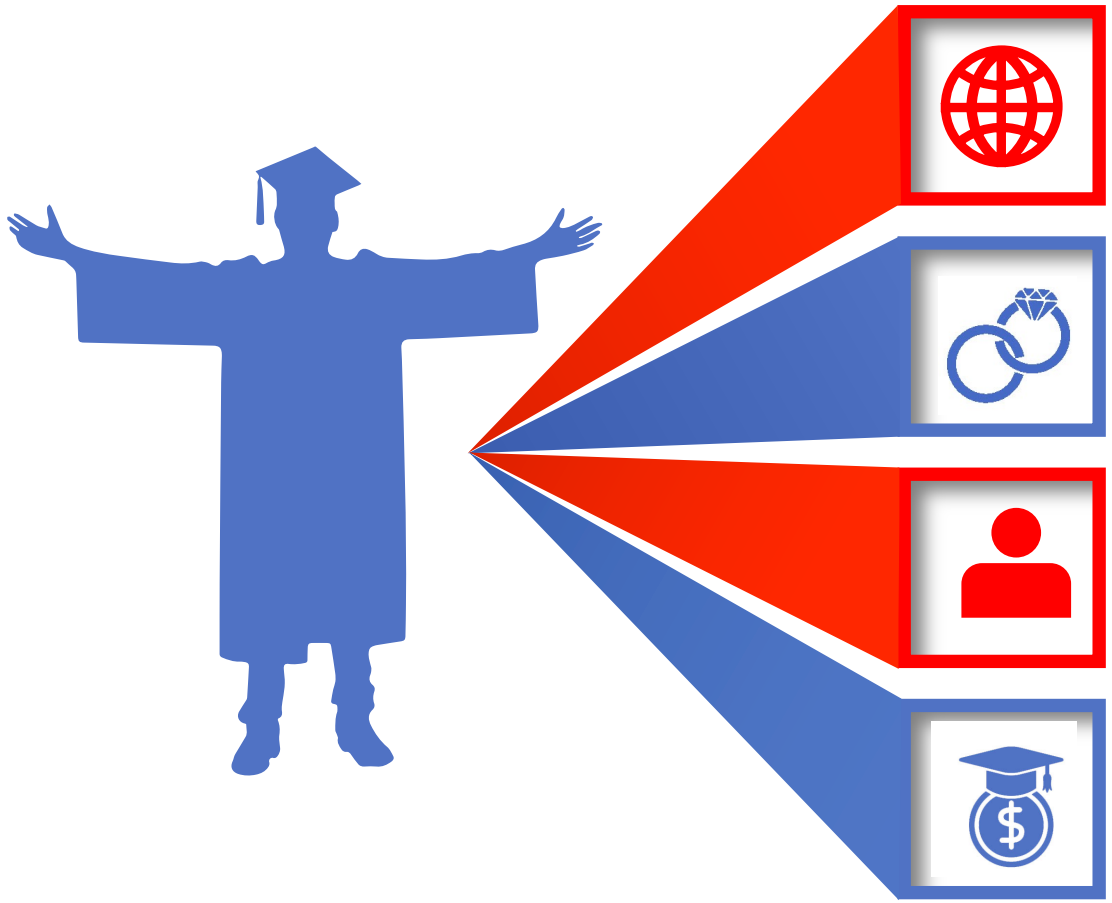
Youngest age is 17

Oldest age is 70

Male

35%

Demographic Information



Nationality

98% of students are from Portugal

Marital status

89% Of student are single

Age

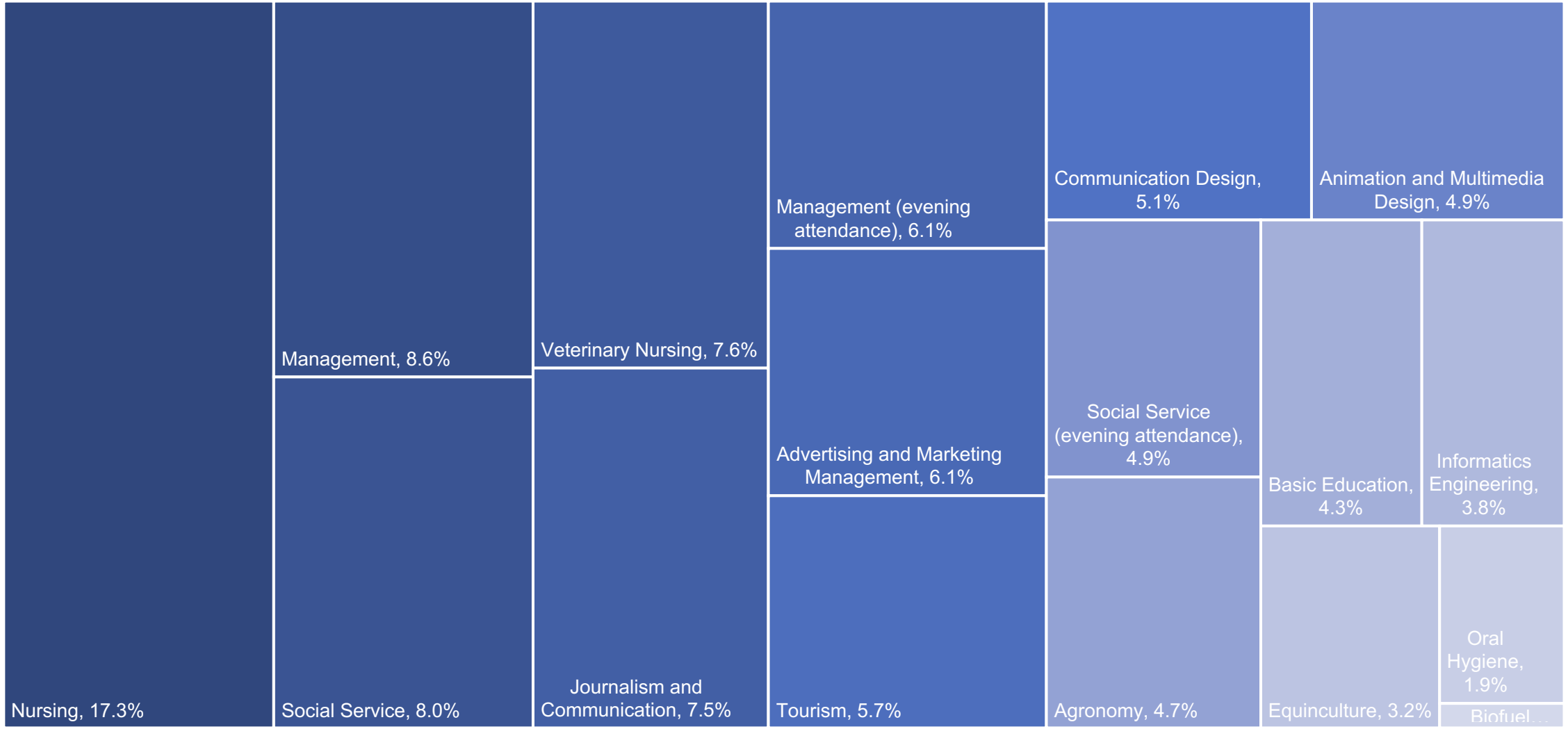
68% of students are younger Than 22

1% of them are older than 50

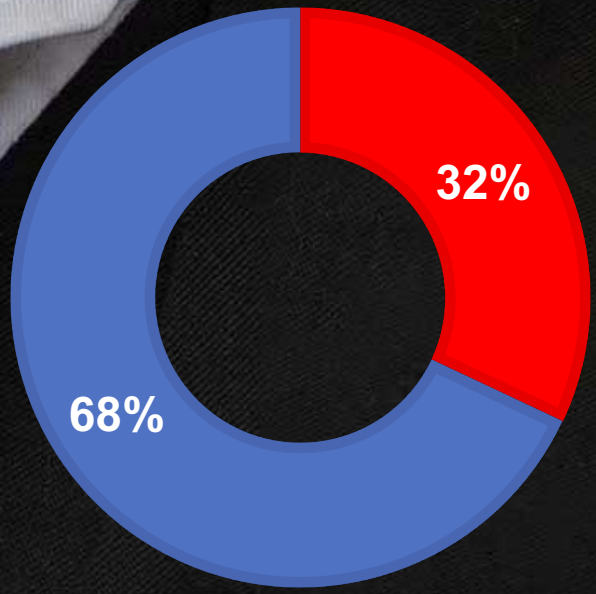
Scholarship

25% of students have scholarship

Major's Popularity

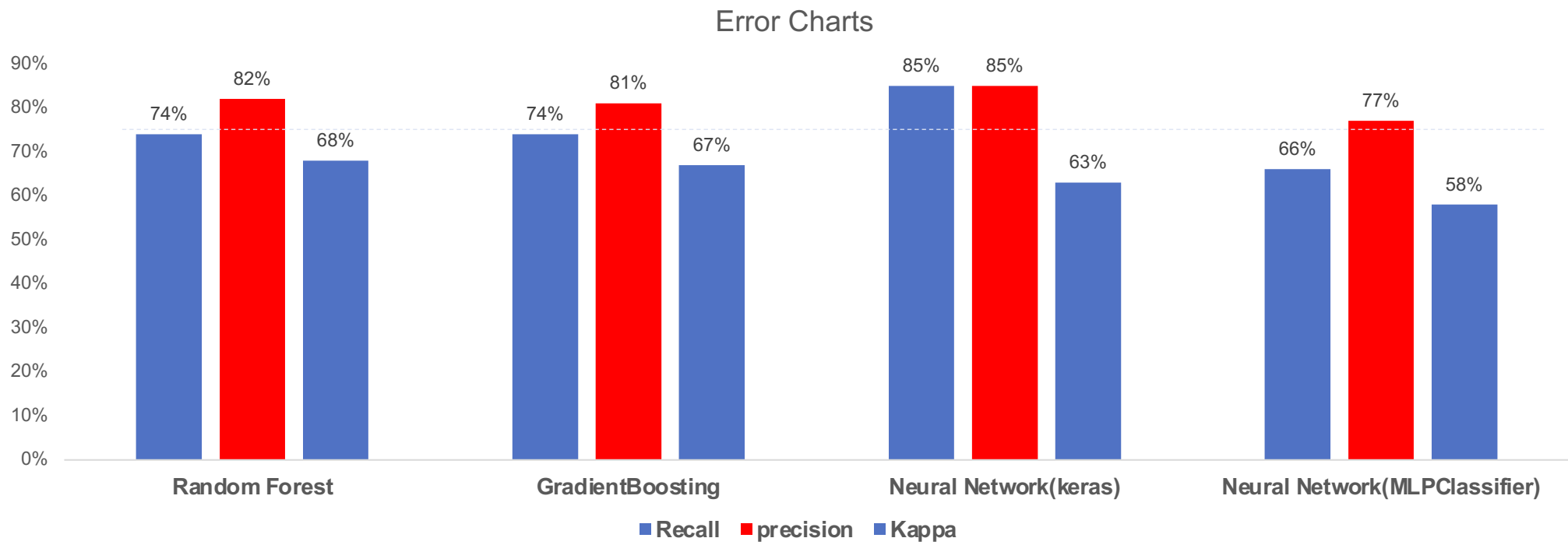


Graduate /Dropout



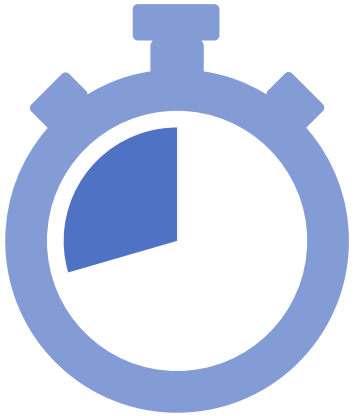
Prediction

For this project, four different type of prediction models have been performed and Based On the Error Metrics the best models is **Random Forest**.
Based on this Model, Dropout Can be predicted by 74% of success



High score Features in predictions

Grade average in the 2nd semester



17%

Average Grade in 2nd semester for dropout student is half of Graduated ones

Dropout Grade(Avg): **6**

Graduated Grade(Avg): **12**

Grade average in the 1st semester

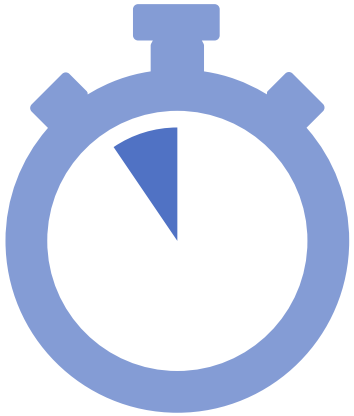
10%

Average Grade fin 1st semester is dropout student is half of Graduated ones

Dropout Grade(Avg):**7**

Graduated Grade(Avg): **12**

Paid Tuition fees



8%

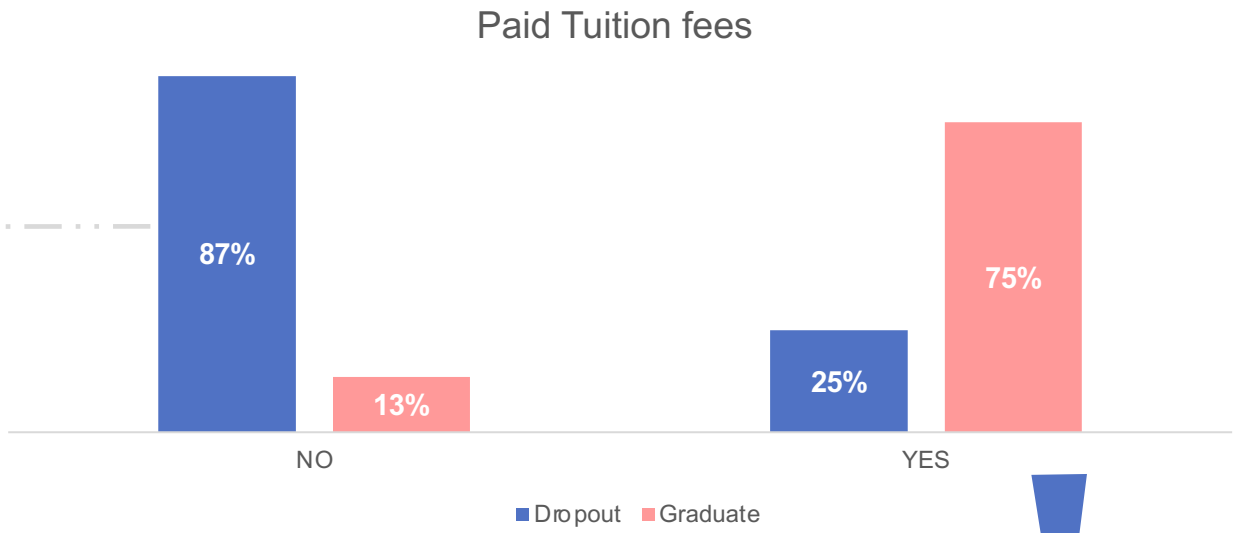
Just **32%** of Dropout students has paid tuition but for graduated ones is **98%** ,we will see the related graph in next page

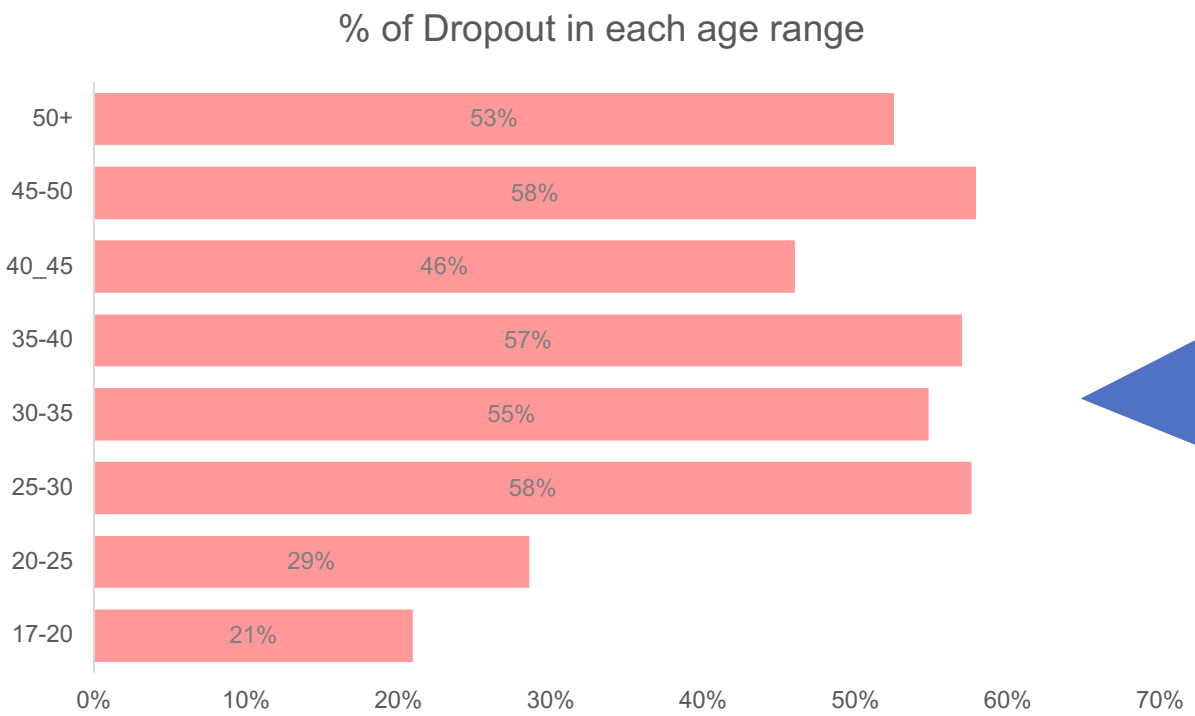
.

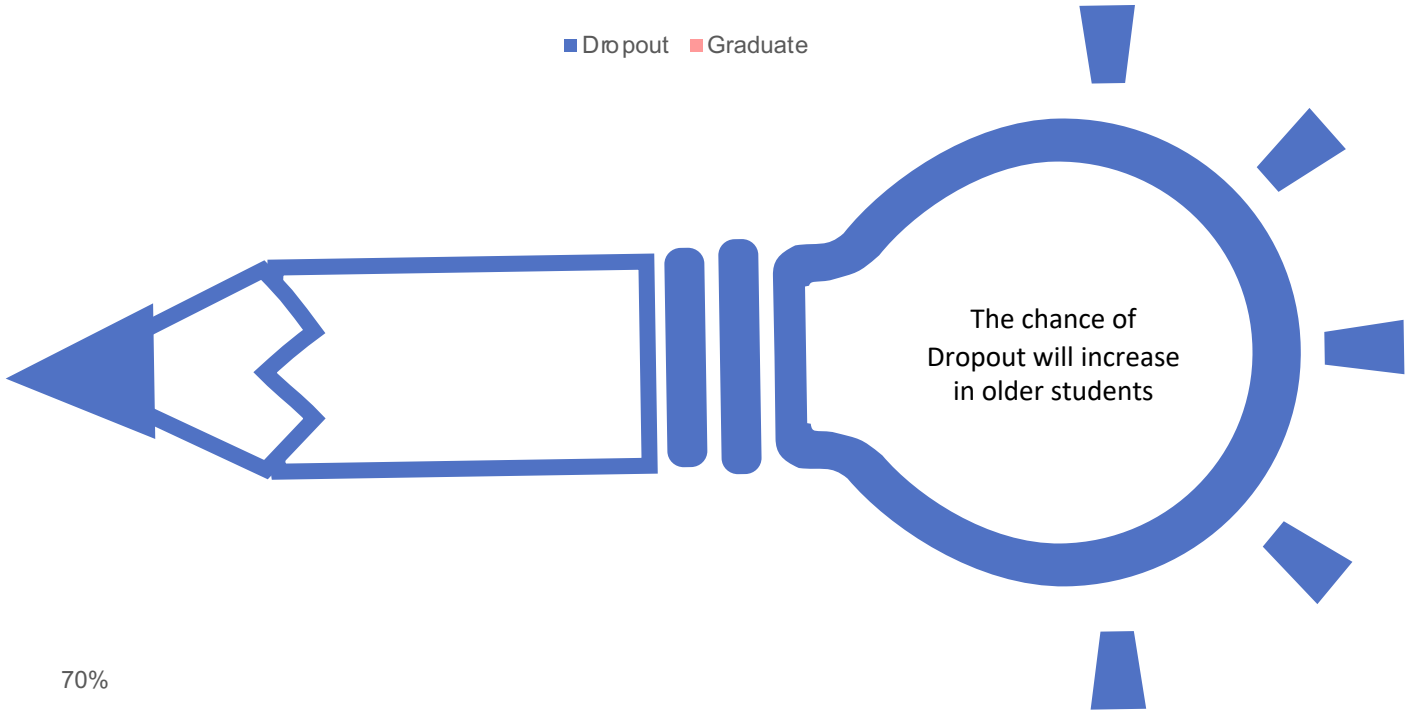
The average of Previous qualification grade((Dropout:131,Graduate:133) and admission grade(Dropout:125,Graduate:128) is almost the same for Both Groups so something else must has been the main driving factor.

Comparing Other Features

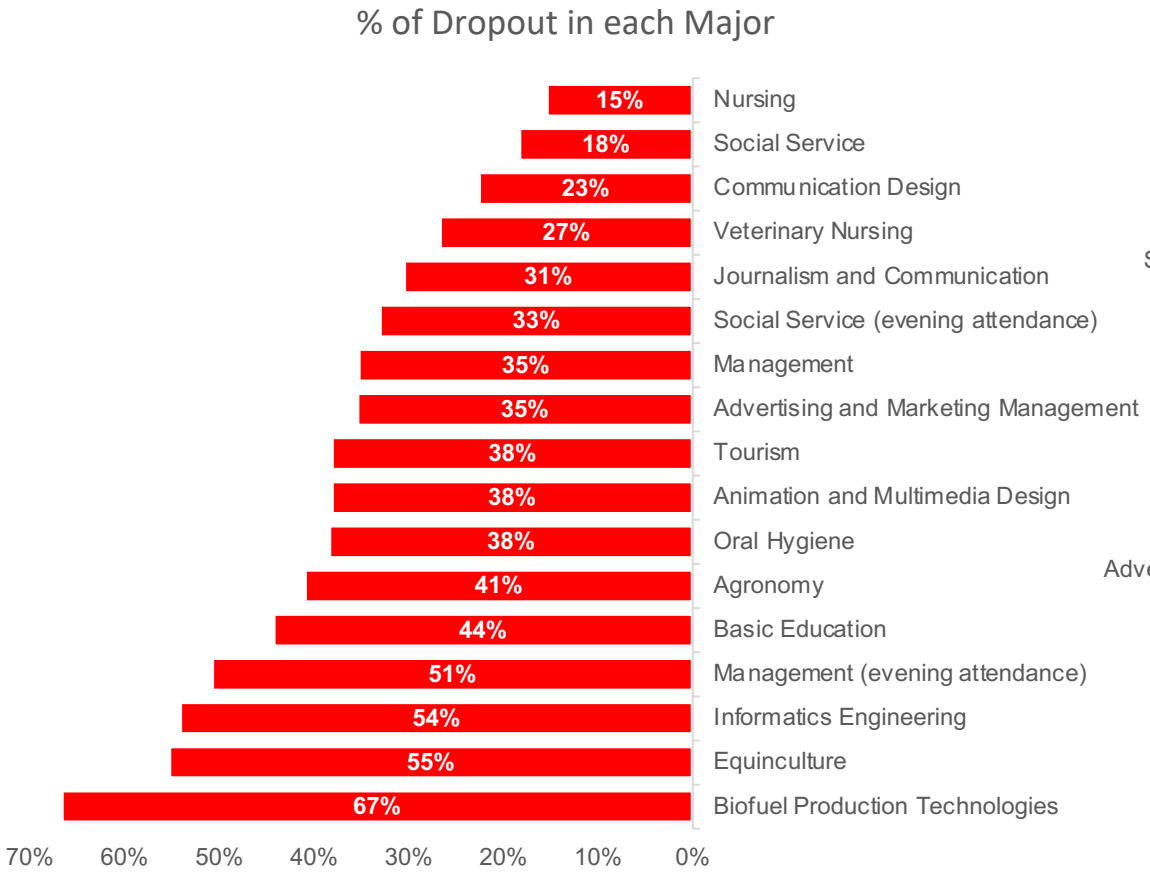
It shows that 87% of students hadn't paid their tuition fees, so it seems they had doubt about finishing university

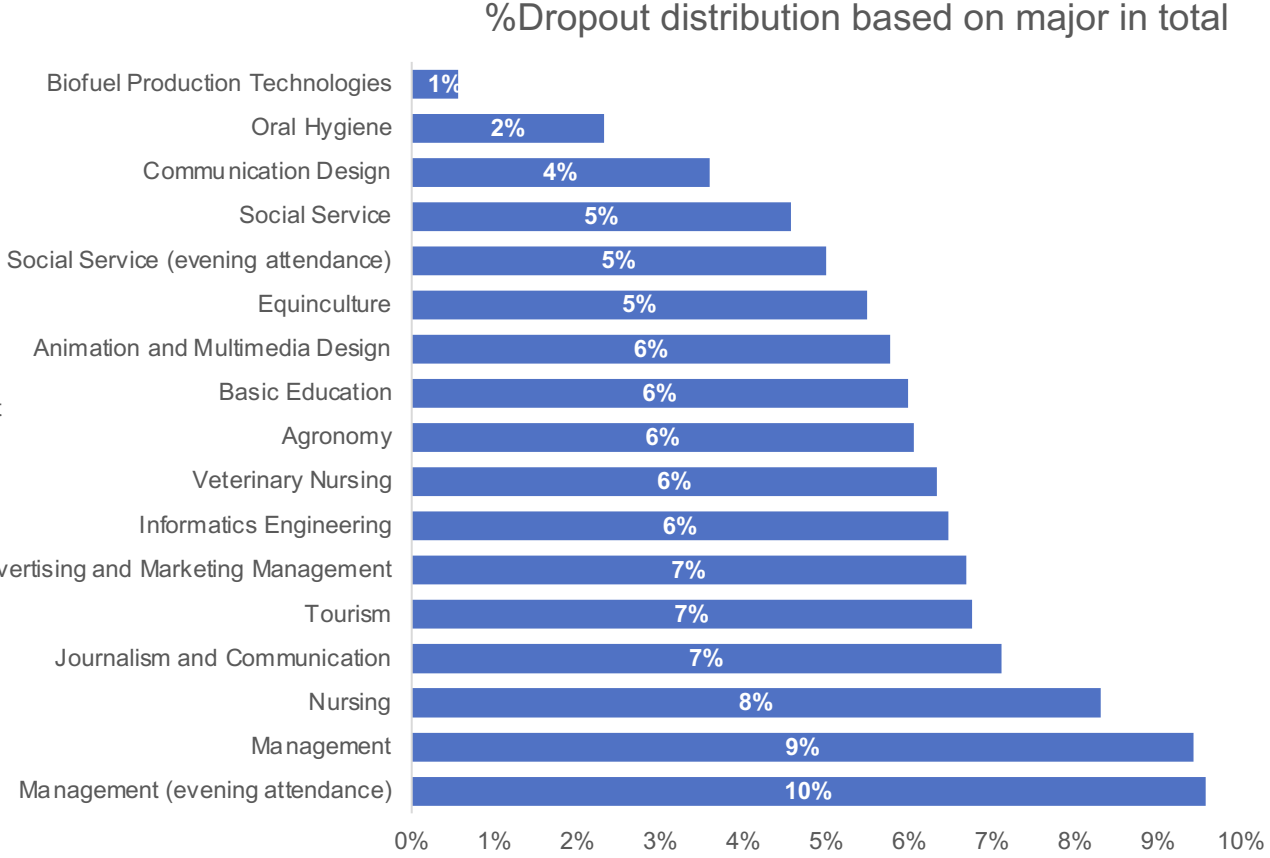




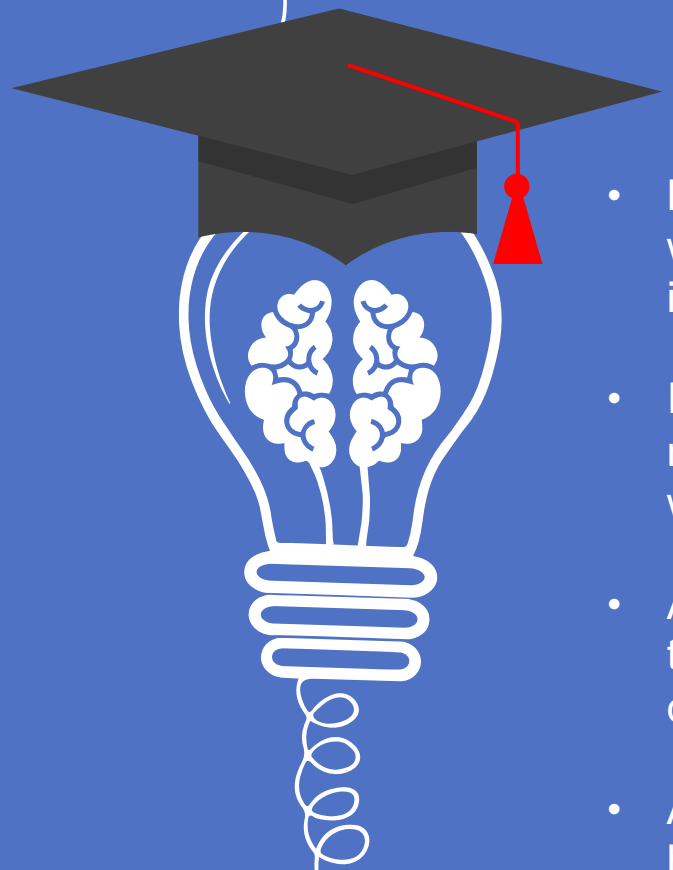


Dropout in different Majors



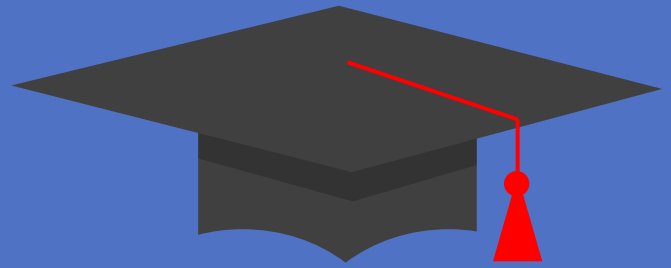


- Managements(evening attendance) has highest proportion between majors in total dropout of university.
- However significant number of Biofuel Production Technologies tends to drop out with 67% dropout rate!!!



Conclusion

- Based on This Sample data , Students who preform well in 1st and 2nd semester will be unlikely to Dropout from the university ,so further study and data collection is required to identify drivers for lack of performance in the initial semesters.
- In the mean time universities can help reinforce studies and potentially offload the number of course in the 2nd semester for those students who doesn't performed well in the 1st semester.
- As tuition payment correlates with the increase in commitment and a reduction in the dropout, the universities can further study the reason for non-payment or drivers that could facilitate tuition payment.
- As highlighted in the findings, older students would be more likely to drop out; hence, the universities can create coaching/guiding programs to navigate through academic semesters for this demographics.
- For those majors which has high rate of dropout another research is recommended to identify missing link the reasoning.



THANK YOU