Predicting Student Outcomes

Virgilia Antonucci

Overview/Business & Data Understanding

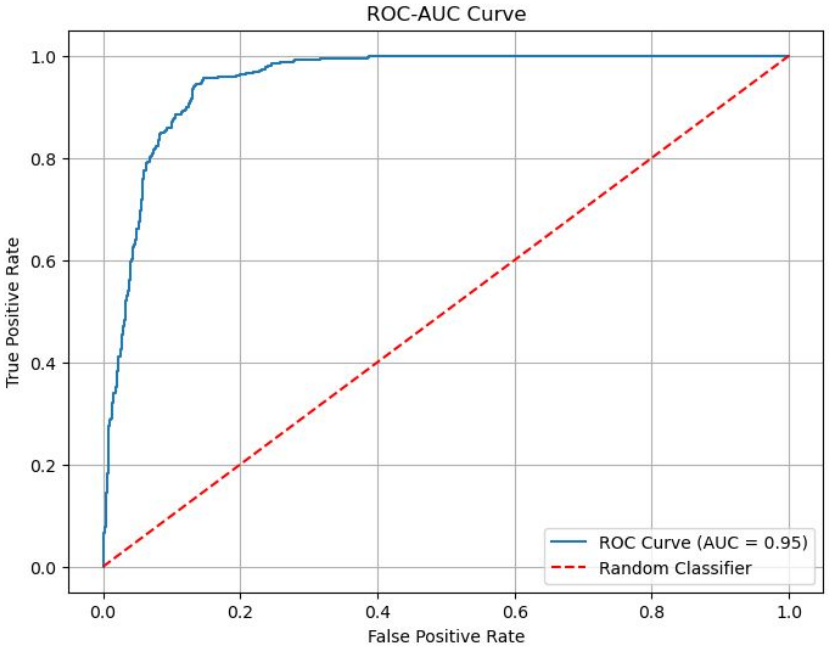
Business Problem:

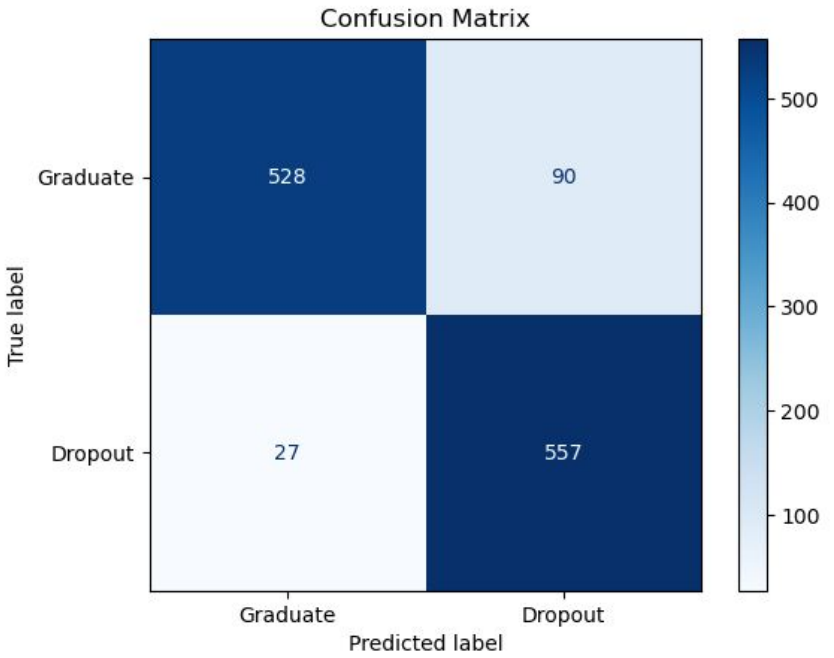
I have been tasked with evaluating students based on their academic performance and background to predict whether or not they will drop out of the university.

Data:

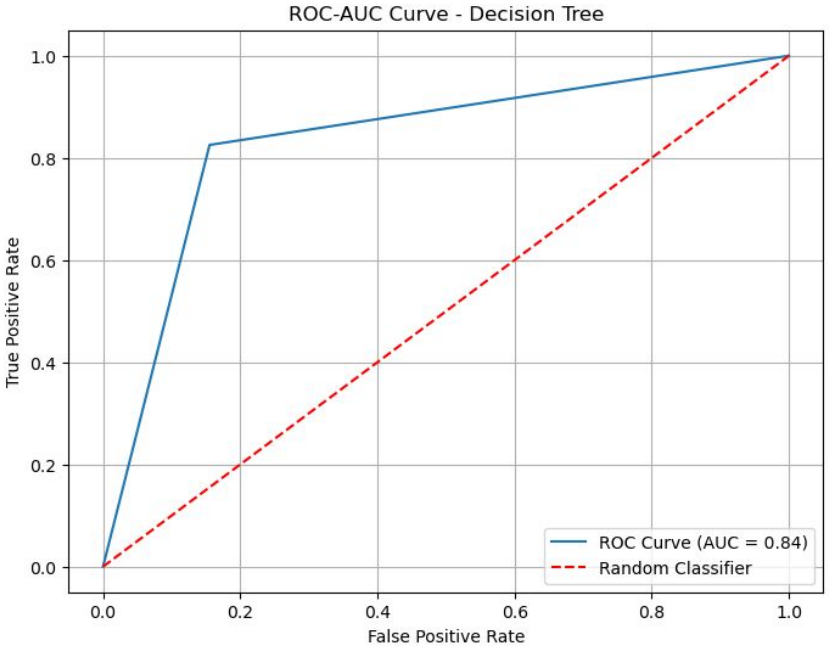
- I have been provided a dataset of approx. 4500 rows and 40 features.
- Each row is a student.
- The information provided for each student includes marital status, nationality, curriculum units for each semester, the amount of education achieved by each student's parents, etc.

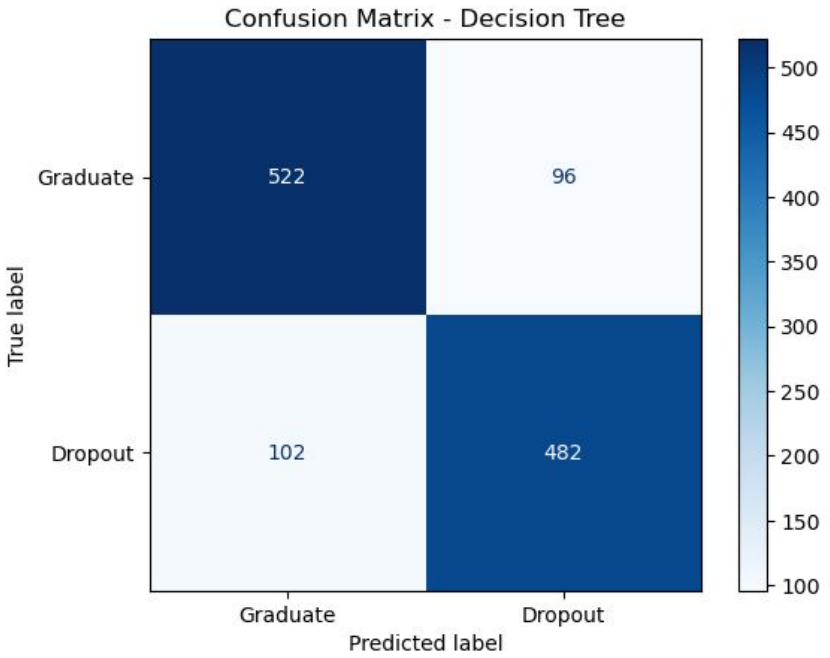
Modeling - Logistic Regression



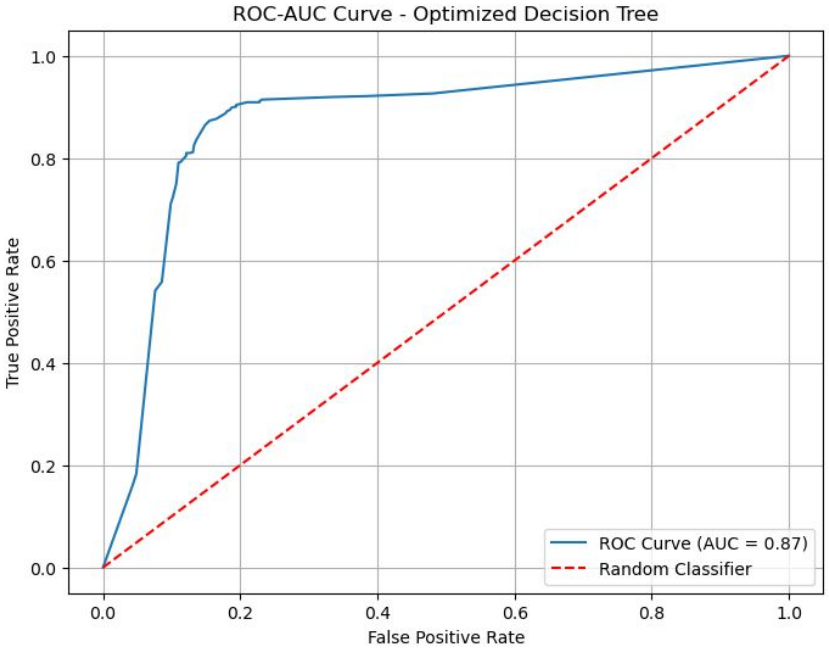


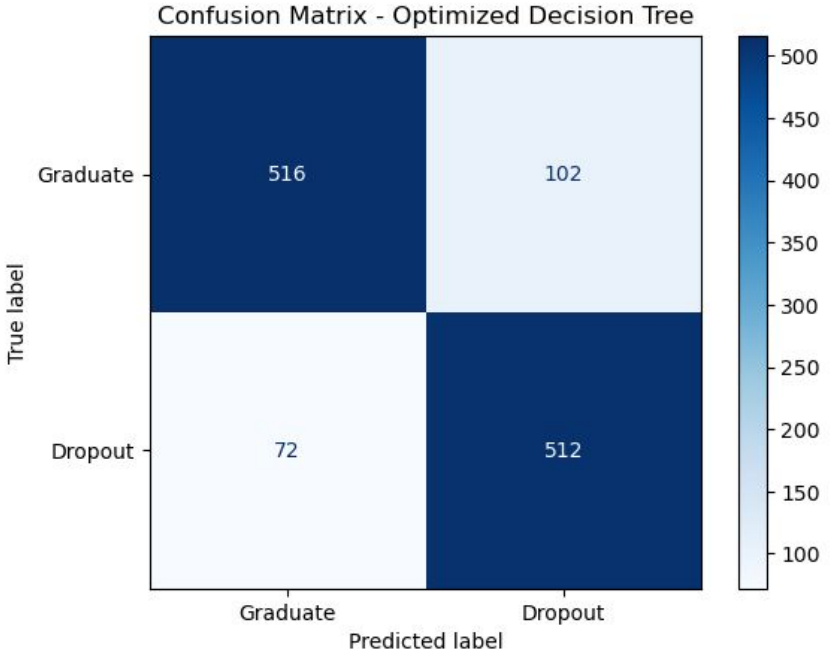
Modeling - Decision Tree



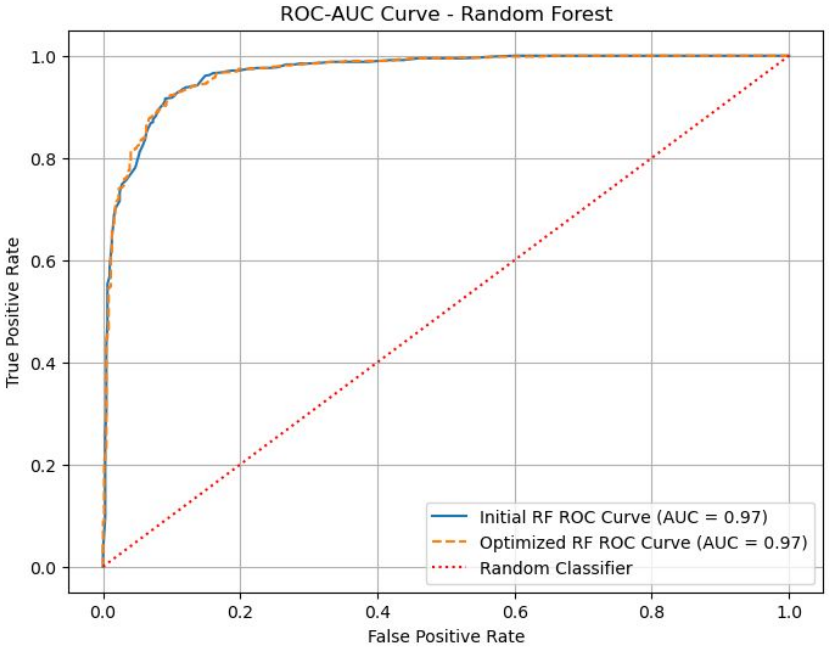


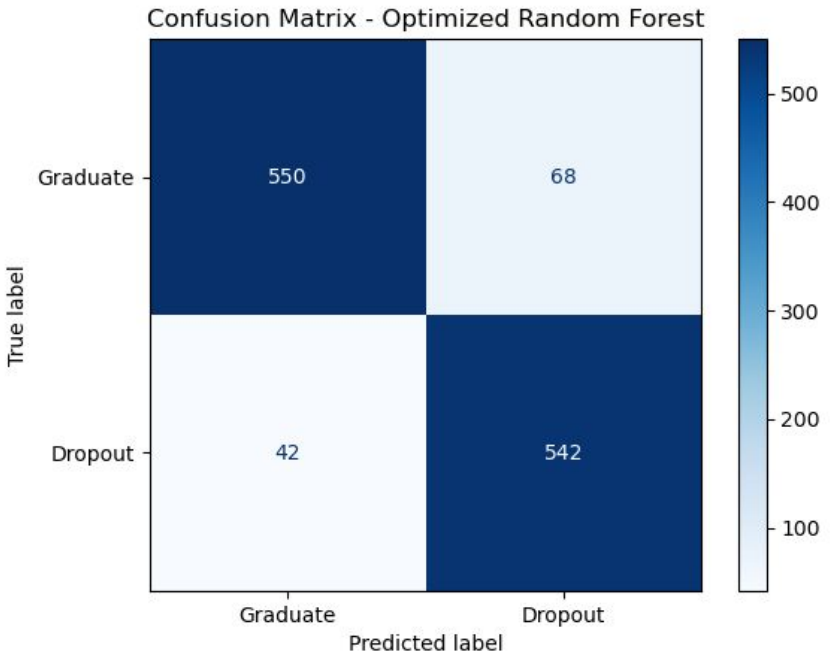
Modeling - Dec Tree Hyperparameter Tuned



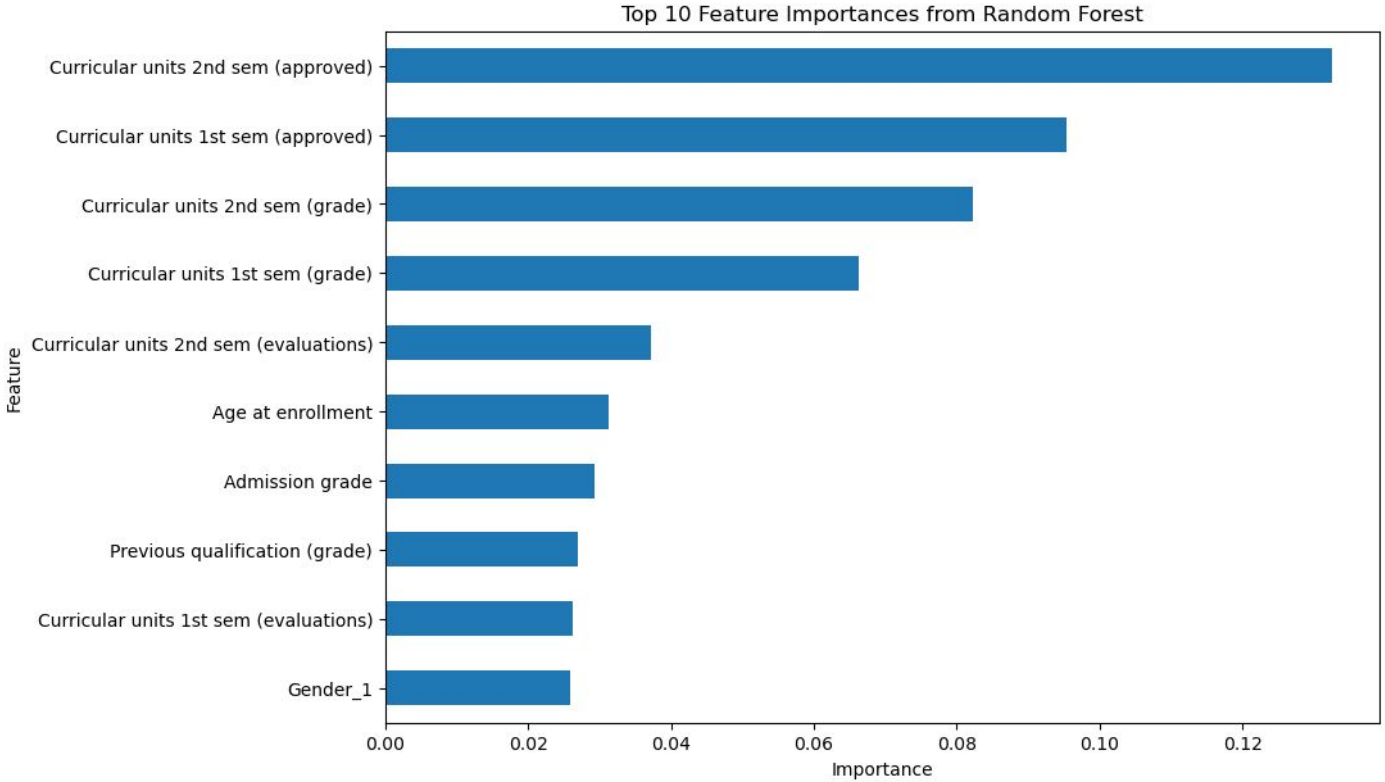


Modeling - Random Forest





Features


Top 10 Feature Importances from Random Forest
Curricular units 2nd sem (approved)
Curricular units 1st sem (approved)
Curricular units 2nd sem (grade)
Curricular units 1st sem (grade)
Curricular units 2nd sem (evaluations)
Age at enrollment
Admission grade
Previous qualification (grade)
Curricular units 1st sem (evaluations)
Gender_1
Feature
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Importance

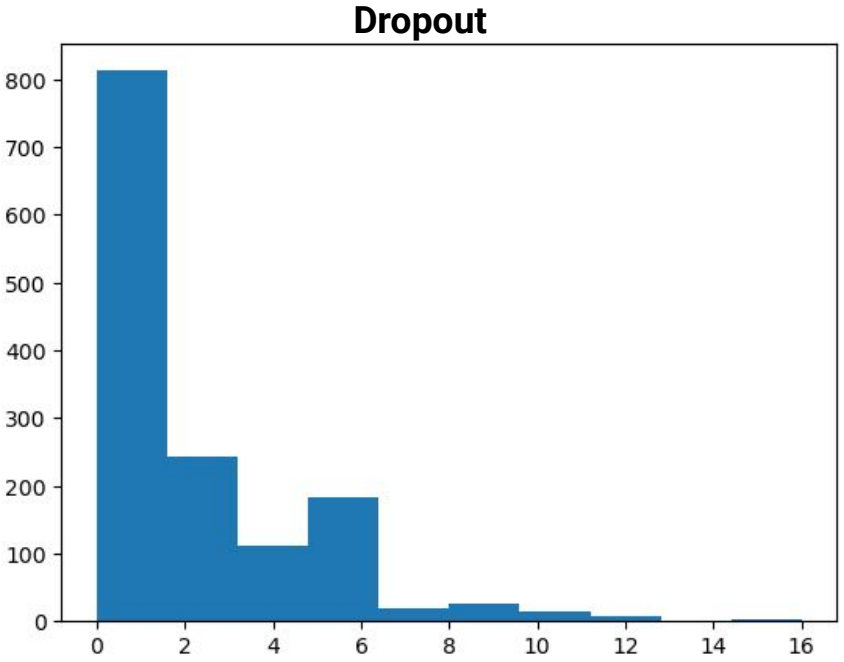
Evaluation

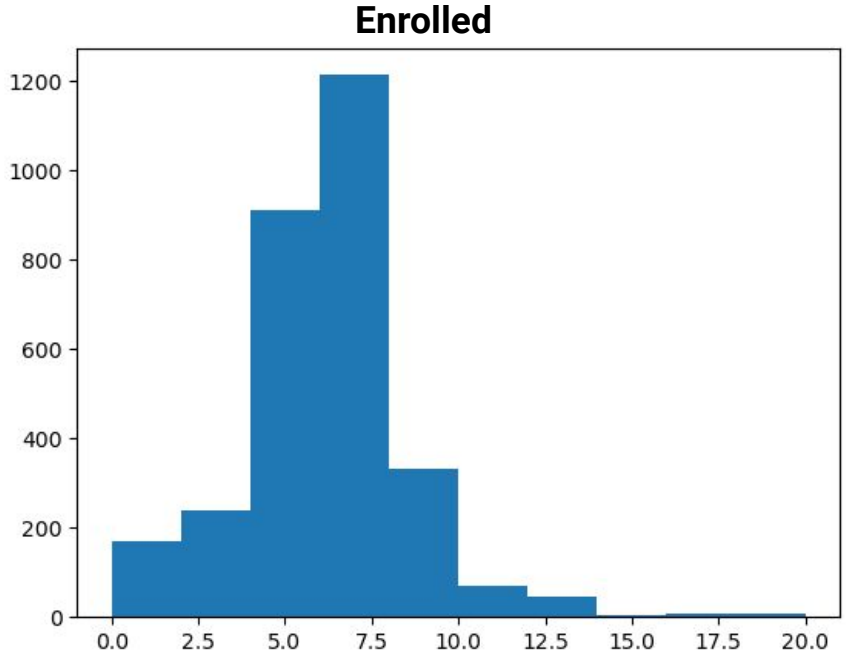
- **Tuned Random Forest** model performed the **best**
  - Random forest is an ensemble of decision trees
- **Hyperparameter tuning** was performed on the Random Forest model to find the optimal parameters that **yield the best predictive performance**
- The model's feature importance revealed **key features** that **influence student dropout**

Recommendations & Next Steps

- If certain variables are not affectable, then the university should make sure to be aware of them and help the variables that are within their control.
- The feature importances show us that curriculum units, tuition, and mother’s occupation are the biggest predictors for whether or not someone will drop out.
- Going forward, the school should closely monitor these values and make sure these values are reported for each student.

Recommendations





Approved 2nd Semester Credits

Thank You

GitHub

LinkedIn