- Team members:
  - **Aidan Gillespie**
  - **Kashaina Nucum**
- Topic:
  - **Student Retention Prediction with Feature Selection Techniques**
- Project Description:
  - Understanding factors that influence student retention rate and predicting potential dropouts may the university to implement methods and services, such as financial support or tutoring, to help students who are predicted likely to drop out stay at the university longer.
  - Well-working model(s) can be used at other real universities.
- Objectives:
  - To implement and compare different supervised machine learning models, such as random forest, support vector machine, and neural network to determine which model yields the highest performance measures, such as accuracy, precision, F1-score, and recall.
  - To explore various feature selection techniques, which may include filter and/or wrapper methods to improve performance measures and speed.
  - Overall goal: to determine which demographics, academic metrics, and socio-economic factors best affect student retention rates.
  - Overall goal: to predict student retention with a satisfactory level of accuracy, as excessively high accuracies could indicate overfitting, especially with the complexity of student behavior and outcomes.
- Dataset information:
  - Source: https://www.kaggle.com/datasets/thedevastator/higher-education-predictors-of-student-retention
  - 4000+ rows
  - 35 columns:
    - Marital status
    - Application mode
    - Application order
    - Course
    - Daytime/evening attendance
    - Previous qualification
    - Nationality
    - Mother's qualification
    - Father's qualification
    - Mother's occupation
    - Father's occupation
    - Displaced
    - Educational special needs
    - Debtor
    - Tuition fees up to date
    - Gender
    - Scholarship holder
    - Age at enrollment
    - International
    - Curricular units 1st sem (credited)
    - Curricular units 1st sem (enrolled)
    - Curricular units 1st sem (evaluations)
    - Curricular units 1st sem (approved)
    - Curricular units 1st sem (grade)
    - Curricular units 1st sem (without evaluation)
    - Curricular units 2nd sem (credited
    - Curricular units 2nd sem (enrolled)
    - Curricular units 2nd sem (evaluations)
    - Curricular units 2nd sem (approved)
    - Curricular units 2nd sem (grade)
    - Curricular units 2nd sem (without evaluation)
    - Unemployment rate
    - Inflation rate
    - GDP
    - Target
  - Roughly 35% of the dataset dropped out