2019

Data Science and AI

Module 0

Capstone Project

Capstone Project

- You are required to **define, implement and present** a **Data Science project** towards the end of the course.
- Project milestones:
  - June 3: present **3 ideas for the project**
  - June 10: decide on one option
  - June 17: collect data
  - June 24: present initial findings
  - July 8: present an update
  - July 22: Dry run of final presentation
  - July 29: **Present final report**

What to present

- **Business perspective**
  - Business insights uncovered
  - Business scenarios for how the project can be deployed and used
  - Approach for estimating business value
- **Technical perspective**
  - Techniques used
  - Pipeline
  - Model validation results

Mini Project III

- The objectives of this mini-project is to:
  - Apply all what you have learned so far
  - Prepare you for the Capstone project
    - You can use this mini-project to qualify ideas for the Capstone project
  - Understand your **learning gaps**
- Choose a **dataset** from Google Public Datasets, scikit-Learn or anywhere else; or select a **question** and identify and gather the data you need

- Please prepare present on Saturday 8, June on:
  - **Business aspects** of the project
  - **Outcomes**
  - Potential **additional insights** that can be obtained from available data
  - **Data analysis** and **Machine Learning** techniques used
  - **Gaps** in your skills

Mini Project III

- The **objectives** of this mini-project is to:
  - Apply all what you have learned so far
  - Prepare you for the **Capstone project**
  - Understand your **learning gaps**
- Please present as if it is **your job interview take-away test** on:
  - **Business aspects** of the project
  - **Outcomes**
  - Potential **additional insights** that can be obtained from available data
  - **Data analysis** and **Machine Learning** techniques used
  - **Gaps** in your skills
- Please **listen actively** to your fellow classmate's presentation and **ask questions** as if you are the **hiring manager**

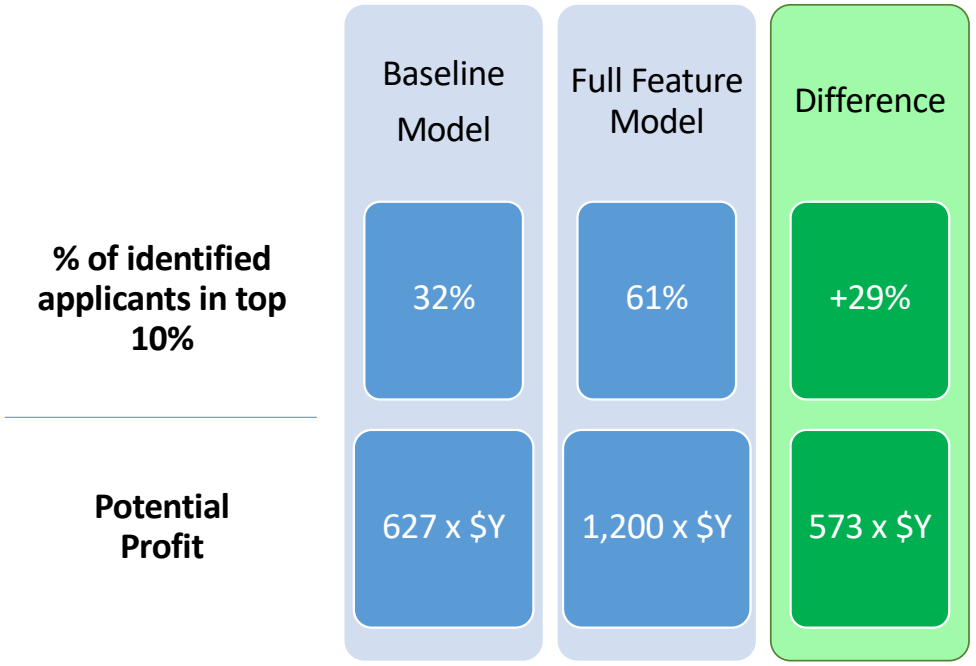
Questions?

Appendices

Results comparison and business case overview

Applying the model for Banking can lead to potential annual **revenue twice as big** as the current model.

Results overview



Business case overview based on the Final Model

Assumptions:

- Customer Value/year is $**1000**
- Customer base = 1.4 million
- Top 10% = 140,000 customers
- 2.8% applicants over 2 years ie. 1.4% annually
- 1.4% applicants in top 10% = 1,960
- 61% identified to target = 1,200
- 32% identified to target = 627

Potential profit = 573 x 1000 ≃ $500,000/ year

≃ $1.5m over 3 years

End of presentation