Siddharth Ganesh

srganesh@andrew.cmu.edu

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
Bachelor of Science in Business Administration — GPA 3.96 — May 2023
Concentration in Business Analytics and Technologies

Relevant Completed Coursework
• Applied Data Science for Business • Modern Data Management • End-to-End Business Analytics • Data Mining & Business Analytics • Optimization for Business • Reasoning with Data • Multivariate Analysis • Methods of Statistics and Data Science

TECHNICAL SKILLS

Programming:	Python, Java
Data Analysis:	Python (Pandas, Skikit, Pytorch), R, SQL, Tableau, Tableau Prep, Power BI
Excel:	Open Solver, VLookup, Pivot Tables, Macros, Financial Formulas

RELEVANT EXPERIENCE

Ingredion — Chicago, IL
Data Intern — Summer 2022

- Worked on a project aimed to reduce bad batches of a certain food product by deploying a Machine Learning model with the goal of identifying irregularities.
    - Merged historical data from a data lake and real-time incoming data from IOT edge devices at the plant. Used **PySpark** and **Pandas**.
- Created **Power Apps** for operators to enter production data and goals, that was then stored in **SQL Server** tables I created.
- Brought together data from different **SQL Server** tables to create a **Power BI** report that helped plant managers keep track of their aggregate production numbers and their relation to the production goals.

Tredence Consulting — Seattle, WA
Data Science Intern — Summer 2021

- Worked with an AT&T dataset to evaluate which of their retail stores performed well, understand why those stores performed well, then created a machine learning model that predicted what a certain store's performance should be given the store's characteristics.
- Worked with a senior consultant daily to make sure my solution was viable and best met my client's needs.
- Used **SQL** to run the analysis.
- Used **R** to create the machine learning model.
- Presented my model and results to executives at the company.

Carnegie Mellon University Pittsburgh, PA
Project Work Aug 2019 to Jun 2022

Applied Data Science for Business

- Worked with the CMU Endowment Office to understand their work and needs.
- Worked with a group of students to come up with a data-based solution and implementation.
- Researched different data sources and variables that aligned with the client's needs.
- Created a dashboard in **Power BI** that allowed the CMU Endowment Office to compare different countries and evaluate the economic landscapes of those countries.

Business Technology Group

- Worked with a group of students to develop an app, Pocket Closet: an app that lets you track items in your closet and the outfits you have worn.
- Worked closely with software developers and designers to develop a database schema for the app.
- Implemented database in **Dynamo DB** and **AWS** to store clothing items and images.

Carnegie Mellon University
Course Work

Machine Learning for Business (Python)

- Regularized linear regression models for a sports analytics assignment using **lasso and regression**.
- Used **forward selection and backward selection** to select features for linear regressions.
- Used **decision trees and logistic regressions** with different regularizations to predict heart disease. Compared the models to see which model was more accurate.
- Reduced the dimensionality of a gene pool data set using **PCA**, and then used models like **logistic regression and random forest** to predict which people have high probability of cancer.
- Predicted listening behavior on a music platform using **K-nearest-neighbors**.
- Classified images using a **neural network.**

End to End Business Analytics (R)

- Created a Machine Learning model using **R** to recommend an investment strategy for a Lending Club user.
- Analyzed the features of the dataset to see which features correlated the greatest to a loan defaulting.
- Derived loan grades by doing cluster analysis.
- Ran decision trees to understand the top rules for a loan defaulting and utilized logistic regression to predict the chance of loan defaulting.
- Optimized portfolio investment with **Excel Open Solver** by maximizing expected profit calculated from ML regression model based on investment amount and quantity constraints.

Modern Data Management

- Developed **SQL** queries on relational database to answer business queries on data of HealthyRide, a bike renting business in Pittsburgh and PGH Parking, a parking business in Pittsburgh.
- Modeled a relational database for an outdoor activities company to keep track of people signing up, payments, schedule etc.,
- Wrote **Javascript** queries (**Mongo DB**) to analyze data from Rotten Tomatoes, IMDB and NBA.
- Used **Tableau Prep** to clean and format raw data about New York City Real Estate, downloaded from the state government's website. Created various types of visualizations like price of real estate based on property age, location etc., using **Tableau**.