Ayman Khan

5793 Tyndall Ave, Bronx, NY 10471 | 332-215-2663 | aymankhaninfo@gmail.com | www.linkedin.com/in/aymanawsafkhan | https://aymankhan2003.github.io/aymansportfolio.github.io/

Education

Middlebury College — **Middlebury, VT**
Candidate for Bachelor of Arts, Major in Computer Science — *September 2021 - May 2025*

Relevant Coursework: Data Structures, Machine Learning, Computer Graphics, Regression, Delving into Data, Statistical Learning, Algorithms and Complexity, Software Development, Natural Language Processing

Languages: Python, R, C++, C, Java, JavaScript, HTML, CSS, Stata, SQL, Excel

Frameworks and Libraries: React, NextJS, Flask, Material-UI, Sci-kit learn, pandas, NumPy, Tesseract, p5.js, HuggingFace, PyTorch, Transformers, Jupyter, MatplotLib, Node, Docker, Typescript, Git, Jira

Experience

Green Mountain Care Board — **Montpelier, VT**
Data Analyst Intern — *June 2024 - August 2024*

- Evaluated provider network adequacy across multiple insurance product subtypes, using patient-provider ratios to identify gaps in healthcare access.
- Analyzed datasets covering over 200,000 healthcare providers and 2 million patients to assess network adequacy, provider availability, and access across various insurance plans.
- Created detailed visualizations and reports, highlighting trends and disparities in provider networks, which informed policy recommendations to improve Vermonters' healthcare access.
- Utilized R to process large datasets and generate actionable insights, enhancing the organization's data-driven policy-making approach.

Fairbanks Museum & Planeterium — **Hybrid, St. Johnsbury, VT**
Data Analyst Intern — *June 2023 - August 2023*

- Used the Tesseract machine learning library in Python to convert over 1000 handwritten flashcard data to Excel.
- Employed R libraries (dplyr, tidyverse, ggplot) for data analysis and visualization of flower bloom and fruit dates trends.
- Managed an online Zooniverse data dashboard for publishing visualizations and data.

Middlebury Consulting Group — **Middlebury, VT**
Student Data Analyst - Data Team — *September 2024 - Present*

- Collaborated with a team of 5 to analyze datasets containing over 100,000 entries, improving model accuracy by 15%
- Researched and implemented innovative approaches to enhance predictive models, leading to measurable improvements in data-driven decision-making.
- Developed validation frameworks to assess and compare model performance, ensuring the reliability of new data estimates.

Projects

Investigating the Attributes that Influence Player Ratings in FC24

- Utilized Support Vector Machine (SVM) to predict player ratings in FC24 , achieving an R^2 of 0.908, indicating high model accuracy.
- Applied hierarchical clustering and principal component analysis to assess attribute impacts on player ratings.
- Conducted data analysis and visualizations using R, identifying key player attributes that influence ratings with minimal error (MSE = 4.44).
- https://aymankhan2003.github.io/aymansportfolio.github.io/assets/FinalStats.html

Market Value Predictor

- Implemented machine learning models (random forest regression (RFR) and LASSO) to predict soccer player market values, achieving 95% accuracy with RFR and 75% accuracy with LASSO.
- Conducted data manipulation and extensive analysis using R, including web scraping.
- Performed feature engineering in Python with scikit-learn, pandas, and numpy.
- Developed a UI using Flask (backend) and HTML (frontend) to interact with the prediction models.
- https://aymankhan2003.github.io/posts/project/ProjectBlogPost.html