Pranab Islam

M: (917) 574-0680
pfi203@nyu.edu
github.com/pranabislam
linkedin.com/in/islampranab

Education

Sep 2020 – Dec 2022 **New York University** *New York, NY*
Master of Science in Data Science | GPA: 3.73
Coursework: Deep Learning, Natural Language Processing, Machine Learning, Big Data, Probability & Statistics, Optimization, Linear Algebra, A/B Testing (Google Online Course)

Sep 2015 – Jun 2019 **The University of Chicago** *Chicago, IL*
Bachelor of Arts in Economics | GPA: 3.81

Professional Experience

Jun 2021 – Nov 2021 **Algorand** *Boston, MA*
Data Science Intern, Product Team
- Created the first weekly clustering system that grouped 10,000+ decentralized apps on Algorand's blockchain using set similarity on segments of the apps' code. This allowed Algorand to use common cluster characteristics to collaborate more systematically with app developers
- Constructed 10+ metrics from existing data and designed the first iteration of event tracking metrics with seven separate event types for Algorand's developer documentation website. These metrics enable better data-driven design decision making based on the discovered user behaviors
  - Tools: Google Analytics API, Google Tag Manager, Google Cloud Platform, Python, Pandas, SQL
- Led weekly product analytics meetings by creating presentations of various data visualizations and in-depth analyses of the Algorand blockchain network and user activity for 15 separate weeks

Feb 2021 – May 2021 **Global Association of Risk Professionals** *Jersey City, NJ*
Data Analyst Intern
- Developed a pipeline in Python to parse large volumes of non-standardized financial pdf files, extract relevant tabular data, and output the results in a shared repository for the data visualizations team

Jul 2019 – Jul 2020 **Mizuho Americas** *New York, NY*
Investment Banking Analyst, Financial Sponsors Coverage Group
- Conducted data-driven financial analyses including credit risk analysis, regression analysis, 7-year financial modeling, and data visualizations in order to execute or evaluate 20+ leveraged buyout and debt funding transactions for multibillion-dollar companies
- Automated Mizuho's weekly leveraged finance market report to translate raw web data into a seven-page, client-ready presentation using Excel VBA, saving two man-hours of labor per week

Technical Projects

Nov 2020 – Dec 2020 **NBA Underdog Matchup and Betting Analysis**
Using Python, analyzed how one could create profitable betting positions by focusing on significant underdog matchups; project detail and results: https://tinyurl.com/nbaUnderdogs
- Trained random forest and gradient boosted tree classification models to make positive expected value bet decisions that outperformed the baseline models (logistic regression, support vector machine, and random guessing). Final classifier was an ensemble of the two tree-based classifiers and generated a 9% return on investment
- Extracted, transformed, and loaded the data into shared repositories; scraped 200,000+ lines of data with BeautifulSoup; reconfigured a basketball-reference.com API to pull statistics; cleaned, merged, and transformed data to create a finalized set of features to analyze

Apr 2021 – May 2021 **Music Recommender System**
Implemented a recommender system with the Million Song Dataset using collaborative filtering via alternating least squares in PySpark; project detail and results: https://tinyurl.com/musicRecSys
- Final model earned 0.08 mean average precision (30x improvement vs. popularity-based model)
- Created low-dimensional visualizations of user factor embeddings with t-SNE and UMAP techniques
- Produced data processing and hyperparameter tuning pipeline via PySpark, Bash scripts, and HDFS

Skills

Python (NumPy, Pandas, scikit-learn, PySpark, PyTorch, Matplotlib), SQL, MapReduce, Hadoop, Dask, Tableau, Google Cloud Platform, Google Analytics API, Linux, Bash, Git, Financial Modeling