Abhinav Bandaru

Philadelphia, PA | abhinow@seas.upenn.edu

linkedin.com/in/abhinav31bandaru | abhinav-bandaru.github.io

MACHINE LEARNING EXPERIENCE

National Aeronautics and Space Administration (JSC, NASA), Machine Learning Intern, Houston **June 2024 – Present**

• Utilized RFID data from the ISS and leveraged nearest neighbor optimization techniques (PyNNDescent, Numba) & SQL to solve RFID localization as a classification problem (**80% accuracy**); deployed model in production via **Bash** script and cron job

• Reduced the training time on **~ 20 million data** points by **30 times** and created **dashboards** on Grafana to evaluate inferences

• Designed a **multitask learning** neural network architecture using hard parameter sharing in **PyTorch**, achieving 70% accuracy in distinguishing RFID fingerprints across antennas and objects.

• Removed outliers using **clustering-feature engineering** strategy across various dimensions using HDBScan and K-Prototypes and dimensionality reduction techniques like **UMap**, **PCA** to filter the corrupted data from the training pipeline.

Fox Entertainment, Data Science Intern, Philadelphia, PA **January 2024 – May 2024**

- Spearheaded a team of 6 as the **Technical Lead** to identify behavioral and demographic segments of Fox show viewers.
- Conducted extensive data analysis and research on user viewing patterns to uncover actionable insights and previously unknown skewed patterns, leveraging **pandas**, **seaborn**, **matplotlib** and **Tableau**.
- Executed comprehensive feature engineering on key metrics including recency and frequency. Applied various clustering algorithms including **DBSCAN**, **HDBSCAN**, **Hierarchical Clustering**, K-means to categorize viewer segments.
- Completed the project with detailed profiling of **customer segments**. Evaluated dissimilarity within and among segments using the **Silhouette Score** and Davies-Bouldin Index to ensure meaningful segmentation and recommendations.

Institute of Electrical and Electronics Engineers, IEEE, Data Science Intern, Philadelphia, PA **August 2023 – December 2023**

- Worked with Aria Khademi (Sr. Data Scientist, IEEE) and Andrew Sproul (Data Scientist, IEEE) to build an AI web scraper for IEEE, capable of asynchronous web scraping using advanced **NLP** techniques in Python.
- Leveraged **Amazon EC2** instances for accessing proprietary IEEE data from a **Redshift database**.
- Employed **spacy** & **nltk** for text analytics and processing, **SQLAlchemy** and **PostgreSQL** for querying and **git** for version control to develop a cross-platform AI model. Expertly authored parameterized SQL queries, incorporating multiple joins and CTEs, to extract and transform data for seamless integration into data pipelines.

EDUCATION

University of Pennsylvania, School of Engineering and Applied Science, Philadelphia, PA **August 2023 – May 2025**

Candidate for Master of Science in Engineering: Data Science

Cumulative GPA: 4.00/4.00

Vellore Institute of Technology, School of Computer Science and Engineering, Vellore, India **June 2019 – May 2023**

Bachelor of Technology: Computer Science Engineering & Data Science

Cumulative GPA: 9.55/10.00

MACHINE LEARNING PROJECTS

Cardiovascular Disease Prediction, *Dr. Zachary Ives* **October 2023 – December 2023**

• Utilized BRFSS survey datasets to design a Classifier capable of predicting the risk of heart disease with **90% accuracy**.

• Addressed **class imbalance** using SMOTE, ADASYN and representative metrics: **F1 score**, precision, recall and **AUC-ROC**.

• Performed extensive preprocessing: addressing Null values, One Hot Encoding, Dimensionality Reduction, Feature Engineering, Resampling, Standardization and Exploratory Data Analysis using NumPy, Pandas, Matplotlib and Seaborn.

• Trained, validated, and tested models based on Neural networks, SVM, Decision trees and Ensemble learning techniques such as **Xgboost**, **Adaboost** and **RandomForest** using Scikit-learn, **PyTorch** and **TensorFlow**.

DATA SCIENCE SKILLS

Programming Languages: Python, R, SQL, Java, PySpark, C++, C

Technologies: Data Science, Applied ML, Data Analysis, Large Language Models LLM, Transformers, Retrieval Augmented Generation RAG, Software Engineering, Predictive Modeling, Model Training & Deployment, Probability, Statistics,

Frameworks: PyTorch, Tensorflow, Keras, sklearn, pandas, Spark, Excel, Tableau, Power BI, Git, AWS, A/B Testing

Essential Skills: Creativity, Problem solving, Communication, Team Player, Adaptability, Critical Thinking