Dharun Suryaa Nagarajan

669-243-7230 suryadharun4772@gmail.com linkedin.com/in/dharun4772 github.com/dharun4772

Education

Master of Science in Data Science Dec 2025
Northeastern University, Boston, MA **GPA: 4.0/4.0**
Relevant Coursework: Supervised and Unsupervised Machine Learning, Natural Language Processing, Big Data Systems, Deep Learning, Database Management Systems, Time Series Analysis.

BTech in Civil Engineering May 2020
Amrita Schoole of Engineering, Coimbatore, India **GPA: 3.6/4.0**

Skills

Languages : Python, SQL, Big Query, C++
Statistical Analysis: Hypothesis testing, Univariate and multivariate statistics, Bayesian Optimization, factor analysis, Time Series Analysis
Machine Learning: NLP, Deep Learning, Classification, Regression, Clustering, Recommendation System
Data Science Pipeline: Data Wrangling, Feature Engineering, EDA, Predictive Modeling, Model Evaluation, MS PowerBI
Other Tools: Git version control, Jira, Docker, Cloud Run, GCE, Snowflake, GCP DevOps
Libraries: PySpark, PyTorch, LangChain, spaCy, SciPy, JAX, Numpy, Sklearn, Optuna, nltk, transformers, huggingface

Professional Experience

Data Science Intern **Boston, MA**
Bevi *July 2024 - Present*

- Used Multi-Query Enhancement, Few-Shot classifier, ensemble retrieval technique to develop RAG based QnA Bot on machine diagnostics and live order tracking. Tools used: Gemini, LangChain, Vertex AI, GCP Cloud Run, Pub/Sub, Snowflake, Slack API.
- Debugged the impact of demand pump failure from probable feature set using significance test (Chi-square, Pison-Biserial Correlation, ANOVA). This helped hardware quickly rectify the uptick in demand pump failures.

Data Science Analyst **Bangalore, India**
Swiggy *May 2022 - July 2023*

- Utilized segmented hierarchical bandits and Thompson sampling for estimating CTR of various PNs, leading to better decision making and an 8% increase in new user traffic, 10% reduction in un-subscription rate.
- Developed a Referral Product fraud report identifying 9 real-time causal features using ANOVA ,Logistic Regression, chi-square statistics which contributed to reduced operational cost of $1.5M across 85+ cities.
- Drove a 5pp increase in new user sign-up rate through targeted A/B testing of Onboarding pages and widgets and deployed an end-to-end experiment tracker on PowerBI using Snowflake as data source for 250K+ daily traffic.

Data Analyst **Bangalore, India**
TATA Technologies *Oct 2020 - Apr 2022*

- Optimized 25+ GCP SQL pipelines to Google Big Query Warehouse reducing run time from 1hr to under 25 mins.
- Designed 4+ looker dashboards to monitor MRR, ARR. Conducted 3 knowledge sessions on Looker for the Partner teams.
- Lead RCAs for all partner/sales metrics for data with 25M+ customer traffic/week. Collaborated with several different teams on reasoning the discrepancies, exploring the hypothesis and finally presented the findings to the leadership team.

Projects

Airline Aspect Analysis: | ***Northeastern University*** **April 2024**

- Performed sentiment analysis on 20,000 online customer reviews of major airlines using LSTM, PyTorch, GloVe embedding to identify sentiments, achieving an accuracy of 82% and an F1 score of 0.82.
- Implemented aspect-based sentiment analysis (ABSA) using the pyABSA framework to extract specific aspects or features mentioned in the reviews and their associated sentiments, providing insights into areas of improvement for airlines
- Performed tokenization, lemmatization, and TF-IDF vectorization, to clean and prepare the dataset for analysis, and explored techniques like topic modeling and word cloud visualization for exploratory data analysis.

Pairs Trading Strategy: | ***Northeastern University*** **April 2024**

- Developed an unsupervised learning-based pairs trading strategy using k-means and DBSCAN clustering algorithms to identify correlated stock prices for statistical arbitrage.
- Performed data preprocessing, PCA and Autoencoder for parameter reduction and cointegration to select optimal pairs.
- DBSCAN performed better with 0.45 Silhouette Score and 0.78 DBI and gave 20+ cointegrated stock pairs to target