

Yashvi Thakkar

yaashviiphthakkarr@gmail.com | [GitHub](#) | [LinkedIn](#) | [Website](#) | 412-618-9420

EDUCATION

CARNEGIE MELLON UNIVERSITY, Pittsburgh, PA

Jan'21-Dec'22

Master of Information Systems Management, Business Intelligence & Data Analytics

GPA: 3.88

Coursework: Machine Learning, Unstructured Data Analytics, Statistics, A/B Testing, Marketing Analytics, Database Management

Leadership: Heinz College Ambassador, Graduate Student Assembly – Socials Committee, Girls Who Code

VELLORE INSTITUTE OF TECHNOLOGY, India

Jul'16-May'20

Bachelor of Technology – Computer Science and Engineering

GPA 3.97

Leadership: Marketing head for two clubs, Primary Education Program Coordinator for an NGO

WORK EXPERIENCE

HERSHEY

Pittsburgh, PA

Student Data Science Consultant, Hershey Enterprise Data Science Team

Aug'22-Dec'22

- **Shelf Space Optimization:** Created a strategic tool for Hershey to assess quality of retail execution by leveraging POS, market, shelf-space, and SKU data to **optimize** product spacing by shelf, store, and occasion using **machine learning models** and **brute force optimization**.
- Designed **A/B** testing experiment to identify average treatment effect of recommended strategy, resulting in 10% lift in revenues.

FEDERATED HERMES

Pittsburgh, PA

Data Science Intern, Advanced Analytics and Business Intelligence

May'22-Aug'22

- Analyzed the behavior of financial bonds from 2014 to 2022, based on Morning Star categories and forecasted future values using **Multivariate ML Models** to identify high-performing ETFs for clients. Built visualizations and interactive dashboards using Power BI.
- Increased Sales Team's competitiveness by identifying financial advisors & banks with high **propensity** to recommend Federated ETFs to investors, resulting in **10%** quarter-over-quarter **lift** in customer acquisition.
- Won the "**Hackathon MVP Intern**" award among 150 interns for outstanding contribution to Client Engagement Centre by developing **ML-driven apps** to automate performance management for sales reps.

LINDE ENGINEERING INDIA

Gujarat, India

Junior Data Scientist

Nov'20-Apr'21

- Devised a **real-time analytical dashboard** using Python and Tableau, drawing insights for client server load management, and eliminated need for searching and merging the data to fetch current license users.
- Enhanced employee productivity by 10+ hours per day, saving \$1000 per month.

BOTREE TECHNOLOGIES

Gujarat, India

Data Analyst

Jun'20-Oct'20

- Constructed 10+ time-series forecasting models including Exponential Smoothing, ARIMA and Prophet to predict windmill parts **demand** for **inventory optimization**, enhancing status-quo forecasting accuracy by 12%.
- Uncovered key drivers of demand fluctuation using **explainable AI** methods like LIME, generating **\$100k** in annual savings for SCM teams.

SALESFORCE.COM

Gurgaon, India

Solution Engineer Intern

Dec'19-Apr'20

- Developed an XGBoost classification model to identify high-risk consumer based on transactional, demographic & financial data, thereby ascertaining consumers' repayment abilities & **reduced quarterly spending by ~\$300k** for Consumer Risk Management team.
- Acquired 10 Salesforce badges in the Business Analyst and Einstein Analytics trails.

ACADEMIC PROJECTS

Ford Ka Clustering Analysis [python | exploratory analysis | segmentation | marketing strategy]

Spring'22

- Implemented a **segmentation** scheme using **k-means clustering** on psychographic and demographic data to develop a small-car marketing strategy and advertising message.

Location Intelligence for Music Concerts [python | clustering | data visualization]

Spring'21

- Created a **Python application** to assist music artists with campaign launch strategy by **recommending** most profitable concert locations based on regional audiences' genre preferences, population density, purchasing power etc. using clustering algorithms.
- **Web scraped** Ticketmaster (concert details), Numbeo (US states), and Spotify (music data) and generated **dynamic visualizations** using seaborn, matplotlib and geo-mapped all 50 US states according to filters.

SKILLS

Languages/Programming: Python(numpy, pandas, nltk, spacy, scikit-learn, pytorch, TensorFlow, scipy), R, SQL, Java, C/C++, Swift, HTML, PHP, CSS, JavaScript, MATLAB

Data Analysis Tools: Databricks, PySpark, Tableau, RStudio, Stata, MS SQL Server, Oracle SQL, Power BI, Procreate, Advanced MS Excel

Functional: Predictive & Prescriptive Modelling, Regression, Classification & Clustering, A/B Testing, Statistical Modelling, Data Visualization