JASH PALEJA

7169365704 | jashmanishpaleja@gmail.com | jashpaleja.github.io | in/jashpaleja | github/jashpaleja

EDUCATION

UNIVERSITY AT BUFFALO, SUNY **08/2022 – 12/2023**
Master of Science, Data Science

UNIVERSITY OF MUMBAI **08/2018 – 05/2022**
Bachelor of Technology, Computer Engineering

SKILLS

Programming Languages: Python, SQL, R, JavaScript, Java, C.
Data Visualization: Tableau, Amazon QuickSight, Power BI, Matplotlib, Plotly, ChartsJS
Technologies/Frameworks: Databricks, DBT (Data Build Tool), Git, PySpark, PyTorch, Pandas, NumPy, Microsoft Excel (VLOOKUP, pivot tables), Google BigQuery, Scikit-learn, PostgreSQL, SQLite, MySQL, JIRA, Slack.
Technical and Soft Skills: Data Wrangling / Data Cleaning, Data Analysis, Descriptive Analysis, Statistical Modelling, Decision Tree, Hypothesis Testing, A/B Testing, AI (Artificial Intelligence), ML (Machine Learning), Generative AI

EXPERIENCE

Data Analyst, Lightforce Orthodontics: Burlington, MA, USA **04/2024 – Present**
- Identified and integrated stakeholder requirements to create a sales dashboard that provide actionable insights into customer interactions & optimize travel decisions, boosting efficiency by 70%.
- Partnered with senior and executive leadership in Finance and Sales to execute Sales Quotas and analysis for multiple SPIFs.
- Assessed the impact of promotions on accounts, alongside Sales and Marketing Directors, to guide strategic promotional planning.
- Automated daily Excel reporting and HTML email distribution using Python, reducing manual effort by 99% and delivering personalized performance insights to 100+ ANZ customers.

Data Science Intern, Healiom: San Mateo, CA, USA **05/2023 – 08/2023**
- Converted unstructured Electronic Health Records (EHR) into a standardized, analysis ready format by architecting an ETL pipeline, boosting data loading speeds by 4x, and reducing computation costs by 60%.
- Devised functionality to automate mapping medical procedures, diagnosis and medications of patients to EHR codes using Python, Chroma DB, and instructor-large model (LLM). Reduced time by 50% for manual code assignment.
- Integrated a fuzzy matching algorithm, using Python and the Whoosh API allowing medical professionals to more easily find codes, and increased efficiency by 80%.

Full Stack Developer, Dotminds LLP: Mumbai, MH, India **08/2020 – 04/2022**
- Designed and optimized a normalized database for a 20-location restaurant chain, improving query performance by 60% and supporting 40K+ monthly customers.
- Developed data-centric applications, including a ReactJS dashboard for political service tracking and a React Native app that cut interior design workflow time by 60%.
- Led and mentored a team of 5 through agile sprints while collaborating with stakeholders to deliver full-stack products with robust data flow integration between frontend and backend systems.

Academic Projects

Portfolio Optimization using Reinforcement Learning
- Extracted up to 20 years of live market data from Yahoo Finance and developed a custom OpenAI Gym environment, implementing an Advantage Actor-Critic (A2C) model with PyTorch and stable-baselines3 to improve portfolio management by 50%.
- Benchmarked agent performance against a Uniform Weighted Portfolio using reward-per-episode analysis, with trends showing strong alignment to the benchmark.

Survival Analysis on Employee Attrition
- Analyzed factors that are most influencing attrition through Random Forest Classifier and presented the impact through Kaplan Meier Survival Curve. Conducted log-rank tests to compare survival curves for understanding attrition dynamics.
- Employed Cox Proportional Hazard Model to quantify individual survival probabilities until attrition and predict survival function, evaluated model performance, and fit through summary and visualization.