

JASH PALEJA

☎ +1 716-936-5704 ✉ jashmanishpaleja@gmail.com [in LinkedIn](#) [GitHub](#)

Experience

Healium

San Mateo, CA, USA

Data Scientist Intern

May 2023 - Aug 2023

- Executed an ETL (Extract, Transform, Load) pipeline on electronic health record (EHR) data, optimizing its structure and format for further analysis.
- Developed functionality to automate mapping medical transcript entities to codes using Python, Chromadb, and instructor-large model. Resulting in reduced time by 40% for manual code assignment.
- Implemented search suggestions of EHR codes & and representations using Python, and Whoosh API improving search efficiency for medical professionals by 80%.

Dotminds LLP

Mumbai, Maharashtra, India

Software Engineer Intern

Aug 2020 - April 2022

- Built React Native app to streamline interior design project management, cutting processing time by 60%.
- Created web app with dashboard & admin panel using ReactJS for political services and info requests.
- Managed a team of 5 to architect and develop apps using agile methodology.
- Added new features like state management with Redux and push notifications, enhancing speed & performance by 50%.

Education

University at Buffalo

Buffalo, New York, USA

Masters in Data Science

Aug 2022 - Jan 2024

- * **Selected Coursework:** Probability, Numerical Mathematics for Computing, Statistical Learning and Data Mining, Programming and Database Fundamentals, Reinforcement Learning, Machine Learning, Data Models Query Language, Data Visualization.

University of Mumbai

Mumbai, Maharashtra, India

Bachelors in Computer Science

Aug 2018 - May 2022

- * **Selected Coursework:** Data Structures, Machine Learning, Database Management System, Artificial Intelligence, Big Data Analysis, Applied Data Science, Data Structure & Algorithm.

Projects

Portfolio Optimization using Reinforcement Learning - [LINK](#)

May 2023

- * Extracted live data from Yahoo Finance for real-time market information upto 20 years of data for about 10 companies.
- * Developed a customized environment using OpenAI Gym and constructed an Advantage Actor-Critic Model from scratch and implemented stable baseline models (A2C, PPO) to optimize portfolio management.
- * Conducted a comparative analysis with the Uniform Portfolio benchmark, visualizing agent's performance through reward per episode graphs, demonstrating its adaptability when an agent was proportional to a uniform weight's portfolio.

Survival Analysis on Employee Attrition - [LINK](#)

October 2023

- * Analyzed various factors influencing attrition and presented findings through Kaplan Meier Survival Curve.
- * Implemented Cox Proportional Hazards Model to delve into individual survival probabilities until employee attrition.
- * Articulated insights on statistical data interpretation and assessed goodness of fit through Cox Proportional Hazards graphs.

Exploratory Data Analysis on Layoff Dataset - [LINK](#)

June 2023

- * Identified key patterns. Demonstrated proficiency in data analysis and trend recognition.
- * Utilized Tableau for data visualization and created an interactive dashboard.

Skills

Machine Learning: OpenCV, Scikit-learn, PyTorch, Regression (Linear, Logistic, Multilinear), Classification, Unstructured Data, Chromadb, Whoosh Api, OpenAI Gym.

Programming Languages: Python, SQL, R, JavaScript, Java, C, HTML, CSS.

Data Visualization: Tableau, Power BI, Microsoft Excel, Matplotlib, Plotly.

Technologies/Frameworks: Git, PySpark, Apache Spark, Pandas, Numpy, PostgreSQL, SQLite, MySQL, PyTorch, JIRA, FireBase, React Js, React Native, Redux, Bootstrap.

Technical & Soft Skills: Management, Database Management, Data Visualization, Data Scraping, Data Wrangling / Data Cleaning, Data Analysis, Descriptive Analysis, Statistical Modelling, Decision Tree, Hypothesis Testing, Exploratory Data Analysis.

Certifications

- Google - Foundations: Data, Data, Everywhere | Coursera | 2021
- Google - Ask Questions to Make Data-Driven Decisions | Coursera | 2022