C SHREYA

+1(408) 892-4441 ◇ San Jose, CA
cshreya955@gmail.com ◇ Portfolio ◇ Linkedin ◇ GitHub

EDUCATION

Master of Science in Data Science, San Jose State University — Expected 2024 Grade: 3.5 / 4
Master of Technology in Computer Science Engineering, JNTU of Hyderabad — 2018 - 2020 Grade: 3.8 / 4
Bachelor of Engineering in Information Technology, Osmania University — 2014 - 2018 Grade: 3.9 / 4

SKILLS

Core Skills: Python, SQL, PyTorch, TensorFlow, Keras, scikit-learn, pandas, NumPy, Apache Spark, Hadoop
Data Visualization: Matplotlib, Seaborn, Plotly, Tableau, Power BI, Streamlit, Flask
Cloud and DevOps: AWS, Azure, Docker, Git
Databases: MySQL, PostgreSQL, MongoDB
Collaboration: JIRA

EXPERIENCE

Graduate Research Assistant
San Jose State University, San Jose, CA — August 2023 - February 2024
Developed machine learning models for pre-construction cost estimation using Python, R, and MySQL. Conducted statistical analysis and data exploration to discern trends and patterns in construction cost data, informing data-driven decision-making.

QA Engineer - Data Integration
Accenture, Hyderabad, India — March 2020 - February 2021
Led data quality assurance efforts for ETL projects, implementing data validation routines and ETL testing frameworks using Apache Spark and Talend. Proficiently executed SQL queries and scripts to validate data transformations and assess data completeness.

Software Engineer Intern
AppCloud Software Solutions, Hyderabad, India — January 2017 - July 2017
Employed data mining techniques and tools such as Python, R, MySQL, and AWS services to collect, store, and analyze historical project data. Established key performance indicators (KPIs) based on project success criteria and utilized data visualization tools like Power BI for reporting.

Undergraduate Research Assistant
Stanley College of Engineering, India — April 2015 - March 2016
Conducted data cleaning and preprocessing using MySQL for data integrity and reliability. Developed a text classification system for categorizing news articles, customer reviews, and tweets, showcasing natural language processing expertise.

PROJECTS

Network Slicing Recognition: Developed a deep learning-based model to recognize network slicing configurations in 5G networks. The model achieved an accuracy of 92% on a large-scale dataset. (GitHub repository)
Heart Disease Prediction: Built a machine learning model to predict the risk of heart disease using various clinical and demographic features. The model achieved an accuracy of 85% on a public dataset. (GitHub repository)
Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data: Designed and implemented a movie recommendation system that incorporates demographic information to provide personalized book suggestions. The system was built using big data technologies like Apache Spark and Hadoop. (GitHub repository)