

C SHREYA

+1(408) 892-4441 ♦ San Jose, CA

cshreya955@gmail.com ♦ [Portfolio](#) ♦ [Linkedin](#) ♦ [GitHub](#)

SUMMARY

Data Science professional with approximately 2 years of experience in applying advanced analytics and machine learning techniques to address complex business challenges. Possesses a solid foundation in statistical analysis, predictive modeling, and data mining, with proficiency in Python, SQL, and various data visualization tools. Strong communicator with a proven ability to translate technical findings into strategic recommendations. Eager to leverage expertise in data science to contribute to a dynamic and collaborative team environment.

EDUCATION

Master of Science in Data Science , San Jose State University	Expected 2024
Master of Technology in Information Technology , JNTU of Hyderabad	2018 - 2020
Bachelor of Engineering in Computer Science Engineering , Osmania University	2014 - 2018

SKILLS

Core Skills: C, C++, Python, SQL, PyTorch, TensorFlow, Keras, scikit-learn, pandas, NumPy, Apache Spark, Hadoop, Kafka

Natural Language Processing (NLP): Text classification, sentiment analysis, language modeling

Computer Vision: Image classification, object detection, deep learning frameworks

Data Visualization: Matplotlib, Seaborn, Plotly, Tableau, Power BI, Streamlit, Flask

Cloud and DevOps: AWS, Azure, Docker, Git, Kubernetes

Databases: MySQL, PostgreSQL, MongoDB ; Data Platforms: Data Warehouses, Data Lakes, Snowflake

Collaboration: JIRA ; Web Development: TypeScript, React

EXPERIENCE

Graduate Research Assistant San Jose State University, San Jose, CA, Client: CalTrains	August 2023 - Till Date
--	-------------------------

As a Graduate Research Assistant, I focused on developing machine learning models for pre-construction cost estimation. This involved conducting statistical analysis, leveraging NLP techniques to extract information from unstructured textual data, and applying text classification, sentiment analysis, and language modeling to assess project sentiment and enhance data augmentation techniques.

Data Analyst Accenture, Hyderabad, India	March 2020 - February 2022
--	----------------------------

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.

Collaborated on Uber's NLP-based projects, enhancing user experiences and operational efficiency. Improved vehicle location accuracy using NLP techniques, enhancing location services for users. Utilized Ludwig for NLP tasks like text classification, generation, and named entity recognition. Utilized C and C++ in various projects, demonstrating versatility in programming languages and enhancing the overall efficiency and performance of the implemented solutions. Contributed to data accuracy and efficiency by extracting information from driver licenses and identifying points of interest in conversations

Data 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 deep learning model for 5G network slicing and achieved 92 % accuracy on large dataset. Demonstrated expertise in deep learning, computer vision, and 5G network technologies.

Heart Disease Prediction: Created ML model for predicting heart disease risk, and achieved 85 % accuracy. Developed Power BI dashboards for effective communication of insights and showcased skills in machine learning, data analysis, and data visualization.

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data: Designed a personalized movie recommendation system using Apache Spark and Hadoop and Created interactive Tableau visualizations for showcasing recommendations. Demonstrated expertise in big data technologies, data engineering, and data visualization.

SentimentScope - Analyzing Customer Sentiment in Product Reviews: Developed sentiment analysis system for product reviews with 85 % accuracy using NLP and implemented real-time analysis in a web app with interactive Plotly visualizations, showcasing skills in natural language processing, sentiment analysis, and web application development.