C SHREYA

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

SUMMARY

Data Science professional over 3+ years of experience utilizing advanced analytics and machine learning to solve intricate business problems. Well-versed in statistical analysis, predictive modeling, and data mining, with strong skills in Python, SQL, and a variety of data visualization tools. Excellent communicator adept at converting technical insights into actionable strategic recommendations. Enthusiastic about applying data science expertise in a dynamic and collaborative team setting.

EDUCATION

Master of Science in Data Science, San Jose State University — January 2023 - December 2024
Master of Technology in Information Technology, JNTU of Hyderabad — January 2018 - August 2020
Bachelor of Engineering in Computer Science Engineering, Osmania University — August 2014 - May 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

Application Engineer — August 2023 - Till Date
San Jose State University, San Jose, CA, Client: CalTrains

- Developed machine learning models for pre-construction cost estimation as an Application Engineer.
- Conducted statistical analysis to support the model development process.
- Leveraged natural language processing (NLP) techniques to extract information from unstructured textual data.
- Applied text classification, sentiment analysis, and language modeling to assess project sentiment and enhance data augmentation techniques.

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

- 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.
- Demonstrated versatility in programming languages by utilizing C and C++ in various projects, 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 January 2017 - July 2017
AppCloud Software Solutions, Hyderabad, India

- Employed data mining techniques and tools, including Python, R, MySQL, and AWS services, to collect, store, and analyze historical project data.
- Established key performance indicators (KPIs) based on project success criteria.
- Utilized data visualization tools like Power BI for reporting and communicating insights.

Undergraduate Research Assistant April 2015 - March 2016
Stanley College of Engineering, India

- Conducted data cleaning and preprocessing using MySQL to ensure data integrity and reliability.
- Developed a text classification system for categorizing news articles, customer reviews, and tweets, showcasing natural language processing expertise.
- Collaborated with cross-functional teams to define data requirements, gather business requirements, and translate them into actionable data solutions.
- Leveraged cloud-based data storage and processing platforms, such as AWS services, to enhance the scalability and performance of data-driven initiatives.
- Demonstrated proficiency in programming languages like Python and R, utilizing them for data manipulation, modeling, and predictive analytics.

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.