RISHITHA GOLLA

Los Angeles, CA | 213-691-4690 | gollarishitha129@gmail.com | https://www.linkedin.com/in/gollarishitha/

EDUCATION

University of Southern California — Los Angeles, CA
Master of Science in Applied Data Science — August 2022-May 2024
Coursework: Machine Learning, Natural Language Processing, Data Management, Data Mining — 3.5/4

Kakatiya Institute of Technology and Science — Warangal, India
Bachelor of Technology in Electronics and Communication Engineering — August 2018-May 2022
Coursework: C, Object-Oriented Programming in Java, Data Structures and Algorithms, Internet of Things — 9.18/10

EXPERIENCE

BIOMED — Los Angeles, CA
Data Engineer Intern — June 2023-October 2023

- Leading development and spearheading implementation of a cloud-based data management and AI platform on AWS, harnessing power from AWS SageMaker and AWS S3 services, resulting in a remarkable 25% reduction in data retrieval time and storage costs.
- Implementing data cleaning pipelines using AWS SageMaker Data Wrangler, resulting in a 95% improvement in data accuracy and consistency.
- Conducting in-depth data analysis on CBD assessment data which resulting in a remarkable 20% increase in operational efficiency.

PROJECTS

Government Benefit Chatbot – BenefitWhiz — October 2023-December 2023

- Developed "BenefitWhiz", a user-centric chatbot for simplifying access to U.S. government benefits with two other teammates.
- Integrated Llama2, a state-of-the-art large language model and LlamaIndex for enhanced data accuracy and added features like voice recognition, text-to-speech, bilingual support in both English and Spanish.
- Successfully enhanced the chatbot's accessibility and user-friendliness, significantly aiding users with visual impairments and Spanish speakers seeking government benefit information.

Financial Text Analysis: Using NLP to Forecast Stock Prices — February 2022-April 2023

- Collaborated with three teammates to extract financial data from quarterly, current, and annual reports of top 10 companies over past five years.
- Conducted sentiment analysis with FinBERT, TextBlob, VADER, Flair, and technical analysis with stock price history.
- Utilized LSTM and GRU models to predict stock prices based on sentiment analysis and technical analysis, achieving an average RMSE value of 6.

An Intelligent Approach to Detect Fraud in Mobile Money — September 2022-December 2022

- Worked alongside three team members to uncover a fraud in mobile money transactions.
- Analyzed data containing 6,362,620 transactions and 11 columns. Forecasted fraud with five algorithms including Logistic Regression, K-nearest Neighbor, Naive Bayesian, XGBoost and Random Forest.
- XGBoost showed good performance in finding fraud with 99% accuracy, least False negative cases among all other algorithms.

An emulation-based system for distributed file storage and parallel computation — September 2022-November 2022

- Built an emulated distributed file system (EDFS) using Firebase and stored partitions of data file in Firebase.
- Implemented partition-based map and reduce on data accumulated on EDFS for 4 search and analytics functions.
- Deducted number of computations for Reduce function by adding combine function after each map function.

Real Time Weather Prediction System Using IOT and ML — October 2021-December 2021

- Built a weather prediction system to predict weather conditions of agricultural fields for smart irrigation with a team of two.
- Collected data from temperature and humidity sensor and sent sensor data to Thingspeak IOT server by WIFI module.
- Performed random forest, linear regression, and XGBoost-random forest to forecast temperature and humidity in real time environment from stored data and got an accuracy of 75%, 79%, 82% respectively.

TECHNICAL SKILLS

Languages and ML libraries: Python, R, Java, C, NumPy, Pandas, Matplotlib, Scikit-learn, PyTorch, TensorFlow
Data Processing Frameworks/Database: Spark, MapReduce, HTML, MySQL, MongoDB, JavaScript, DynamoDB
Software and Version Control: AWS Sagemaker, AWS S3, AWS EC2, Firebase, Hadoop, Matlab, Git, Linux