Kaushik Tummalapalli

Available: May 2024 for Full Time

(+1) 9179944435 kaushiktummala55@gmail.com kaushik-tummalapalli github.com/kaushik-42/

Education

New York University **Sep'22 – (Expected) May'24**

Masters of Science in Computer Engineering, 4.0/4.0 GPA *Brooklyn, New York*

Coursework: Distributed Big Data systems, Machine Learning For Cyber Security, Data Science, Deep Learning.

BV Raju Institute of Technology, Narsapur **Aug'18 – May'22**

Bachelors of Technology in Computer Science and Engineering, 8.38/10 GPA *Hyderabad, India*

Coursework: Analytics, Data Structures and Algorithms, Statistics, Object Oriented Design, Data Base Systems.

Skills

Languages/Domains: Python, Java, SQL[MySQL], Machine Learning, Generative AI, Large Language Models, NLP
Developer Tools: Azure, MLFlow, AWS, CI/CD, Rest API, Streamlit, Kubernetes, Git, VS Code, Snowflake, Airflow
Libraries/Frameworks: PyTorch, SciKit Learn, Spark MLLib, Jupyter, Apache Kafka, Apache Spark, Langchain, Pandas

Experiences

CVS Health **May'23 – Aug'23**

Machine Learning Engineer Intern (Data Engineering, Machine Learning)

- Contributed to CVS Retail's Analytics Engineering (Platform Optimization team) by addressing cloud spend granularity and budgeting challenges. Introduced an accurate forecasting system with pipelines performing data transformations and utilizing historical data, enabling teams to make informed decisions for resource allocation and budget planning.
- Successfully deployed a Prophet model using Snowpark and Snowflake worksheets, ensuring consistent and accurate budget forecasting with an average MAPE of 20 Percent across all applications and maintaining an 80 Percent accuracy.
- Conducted comprehensive experimentation with a range of time series models, including Prophet, ARIMA, SARIMA, and XGBoost, to forecast budgets for cloud compute and storage usage across various applications within CVS Health.
- Transformed the ML results into actionable insights by storing them in Snowflake tables, enabling seamless integration with the reporting layer for comprehensive data visualization and analysis.
- Created a new data model for Azure Cloud costs to enhance forecasting and reporting capabilities.
- Developed a Streamlit application integrating application-level machine learning models, resulting in improved financial control and operational efficiency.

Zemoso Technologies **Dec'21 – May'22**

Software Developer Intern

- Spearheaded the development of a Blinkist-style web application, achieving an exceptional code coverage of over 85 Percent, which significantly reduced system vulnerabilities and increased reliability.
- Worked on the development of service-oriented architecture for large-scale web applications.
- Enhanced user engagement by implementing a responsive and intuitive interface using React and JavaScript, leading to a 50 Percent increase in average session duration within the first three months post-launch.
- Fostered a cross-functional team environment, applying Agile methodologies to reduce time-to-market by 20 Percent and increase team productivity by 35 Percent with project debugging.

Carnegie Mellon University (CMKL - Thailand) **Jun'21 – Aug'21**

Research Intern(Machine Learning, Data Collection)

- Improved existing data pipeline results as a part of the masking task to 90 Percent Accuracy for model development.
- Extracted the images from the video source using FFmpeg, created annotations using labelImg for two different warehouses for the raw data to the model, and used MATLAB to mask various warehouses.

Omdena **Aug'20 – Nov'20**

Machine Learning Engineer and Volunteer

- Designed custom Machine Learning and Deep learning models for Improving the Lives of Cancer Patients by Identifying Non-Cancer Generic Drugs and Modeling Economic Well-Being through Satellite Imagery.
- Utilized LabelBox to gather a diverse range of labels as part of an NLP project on extracting crucial information.

Projects

Starbucks Recommendation System **Project Link**

- Implemented FUNKSVD algorithm to recommend the current user about best selling offers and also recommend new customers[Mean Squared error of 0.003823 for 15 latent features.]
- This project helps us to understand how people make purchases from the various offers that Starbucks offers around its different locations Worldwide.

Cab Cancellation Prediction Using Machine Learning **Project Link**

- Estimates the cancellation rate of a cab at the current time of booking based on historical data by EDA(Exploratory Data Analysis) with 85 Percent Accuracy by using various ML Algorithms like Decision Trees, Logistic Regression, and SVM.