RAHUL VEMURI

Address: **Malvern | PA** Phone: **(484) 645-1984** Email : **rbv5178@psu.edu**
Website: *https://rahul-vemuri.onrender.com* LinkedIn: *https://www.linkedin.com/in/rahul-vemuri-a97278184/*
Github: *https://github.com/rahul9115* Leetcode: *https://leetcode.com/rv_9115/*

PROFILE

A result-oriented professional with two and half years of experience in AIOPs (Artificial Intelligence in Operations). An enthusiastic problem solver who can think creatively, implementing state-of-the-art technologies. Successful in implementing AIOps solutions to an eclectic set of problem statements. Actively pursuing potential internship opportunities in the realm of Data analytics, Data science, and AI/ML. Zealous to apply my refined skills to contribute effectively to your company's triumphs.

EXPERIENCE

PQ Corporation, Malvern, PA Mar 2024 – Dec 2024
Data Scientist Intern

- Implemented an end-to-end AI automated critical Demand Forecasting Solution integrating external market indicators to predict net weight sales. Successfully delivered a Proof of Concept, designed, developed, tested, and deployed a full stack web solution for forecasting Demand within the internal infrastructure. Proficient in AI/ML algorithms such as Nbeats, NHits LSTM, and DeepAR, with hands-on experience using tools like Amazon Forecast and Amazon Sagemaker Canvas. Notably achieved project completion ahead of schedule with comprehensive research, providing a solution surpassing existing options, including those from Amazon Team.
- Led the implementation of a Customer Geocoding Solution leveraging Amazon Location Services to validate customer addresses. Integrated Claude v3 Sonnet model in Amazon Bedrock as an additional validation step to check whether the customer is mapped to the system's address. Implemented an automated pipeline for weekly validation of new customer addresses, with results stored in the company's production S3 bucket.
- Implemented a Chatbot using RAG (Retrieval-Augmented Generation) architecture for querying product literature, integrating Pinecone vector database and Amazon Knowledge base with Titan Embeddings. Developed the frontend of the Chatbot using React.
- Developed a Streamlit application utilizing a pre-trained Claude V3 Sonnet model within Amazon Bedrock. The application was designed to evaluate and prioritize potential companies based on their application of the company's products in their business operations.
- Acquired proficiency in developing AI agents on the LangChain framework, integrating ChatGPT 4.0 with the internet to automate tasks.

Penn State Great Valley, Malvern, PA Oct 2023 - Mar 2023
Research Assistant

- Harnessing text analytics solutions across diverse datasets, entails applying advanced techniques to analyze and extract valuable insights from text-based information. The implementation of these solutions involves utilizing NLP for analysis, and pandas and selenium for efficient data collection and cleaning. Moreover, the position demands expertise in hyperparameter tuning of Deep Neural networks, ensuring the optimal selection of optimizers, and the implementation of effective loss functions to elevate model performance. This comprehensive approach guarantees the efficient utilization of cutting-edge technologies for extracting meaningful information from various textual sources.

LTIMindtree, Mumbai, Maharashtra July 2021 – August 2023
Senior Software Engineer

- Led end-to-end implementation of a Capacity Forecasting solution, optimizing virtual machine resource consumption. Executed design, development, testing, and deployment within LTIMindtree's infrastructure from the data monitored in Zabbix. Employed Flask, TensorFlow (single/multi-variate time series forecasting using LSTMs), GUI-less Linux, and MySQL to execute the project successfully. Implemented transfer learning using TensorFlow to enhance the model's overall performance, showcasing a commitment to continuous improvement and innovation. Praised by the infrastructure team for proactive decision-making and optimizing resource usage.
- During a three-day hackathon organized within the Business Unit, spearheaded the creation of an efficient system to categorize and detect anomalies within application support ticket trends rapidly. The solution effectively utilized advanced algorithms such as Pynomaly for anomaly detection, a Non-Negative Matrix transformation ML algorithm for clustering, and a multi-output logistic regressor for ticket classification. This initiative led us to achieve the Runner Up 1 position, successfully fulfilling the challenge and showcasing our innovative prowess.

- Under the directive of the Vice President, led the execution of a Sentiment Analysis project on company reviews sourced from Glassdoor. Developed a comprehensive tool to extract valuable insights from Glassdoor reviews, ultimately facilitating a deeper understanding of employee sentiments and feedback. The delivery was appreciated for its ease of use and in-depth analysis rendered on an Excel file.
- Engaged in an AIOPs project focused on automating the promotion lifecycle through Artificial Intelligence for learning and development purposes. Designed the solution to give a holistic view of automating the promotion life cycle, integrating various internal modules managed across the company

TECHNICAL SKILLS

Data Science: Python, Pandas, NumPy, Scikit-learn, NLTK, TensorFlow, PyTorch
Programming Languages: Python, Javascript, Java, Powershell, Bash
Development: Flask, Azure, Git, Django, React, Node, Express, MongoDB, Streamlit, LangChain
Others: Amazon Bedrock, Retrieval Augmented Generation, Demand Forecasting, Amazon S3, Amazon Sagemaker Canvas, Amazon Glue, Excel, Power BI, Problem Solving, Computer Information Systems, ML Algorithms, SQL, Hadoop, Time-Series Forecasting, NLP, AWS, Time Series Forecasting, Recommendation Systems like Content-based and Collaborative filtering, Big Data, Neural Networks and Deep Learning, TensorFlow, Git Flow, Hyper-parameters Tuning in DNN, Hypothesis Testing, Data Cleaning, SDLC, Azure Cognitive & Applied AI services, Solution Design, WMI, IT Foundations, ITOPs Automation, Data Structure & Algorithms, Data Modeling, Docker, Generative AI, PostgreSQL, Selenium, Agile, Tableau, Large Language Models Sentence Transformers, Image Processing Algorithms, LLAMA2, Langchain, Bootstrap, Convolutional Neural Networks, MVC, Automation, Data Models, Data Quality, Mathematics, Statistics, Big Data, DevOps, Intelligent Automation, Business Process Improvement, Technical Innovation, Technology Integration, Google Analytics

RELATED PROJECTS

- At the NASA Space Apps Hackathon(2023), collaborated with team members to tackle the mission of leveraging Artificial Intelligence to revolutionize NASA technical standards. The project aimed to bring about a significant transformation in the field of NASA Technical Standards through the innovative use of AI tools like GPT-2 summarizer & large language models like sentence transformer. This project expanded on the technical standards given in NASA STD, delivering a clear and comprehensive interpretation of these standards. The project was well acclaimed & was shortlisted in the top 40 globally.
- At the NASA Space Apps Hackathon(2022), collaborated with team members to tackle the mission of leveraging Artificial Intelligence to preserve Science Legacy. The solution involved using AI to summarize extensive PDF documents for NASA's NTRS (NASA Technical Reports Server) database. By enhancing accessibility and manageability, we aimed to make scientific knowledge more available to researchers and the public. Our goal is to ensure valuable research remains alive for the future. We developed our solution on Python rendered on the Flask web framework.
- Researching on fine-tuning of the LLAMA 2 open-source LLM model for implementing sentiment analysis on financial and economic information. Leveraging Large Language Models (LLMs) provides a powerful means to analyze sentiment in news articles. This research is highly relevant for gaining valuable market insights, ranging from risk management to investment decisions, by gauging the sentiments of stakeholders, investors, and the public.
- Integrating Time Series Forecasting predictions with the insights from LLAMA 2 for stock market analysis. The goal is to leverage the predictive power of Time Series Forecasting while enhancing it with sentiment analysis from LLAMA 2. This combined approach aims to provide more accurate and insightful predictions for stock market trends and movements.
- Analyzing Boston crime trends, predominant types of crime, their concentration and frequency. Gaining insights on city's crime landscape, which could aid in law enforcement strategies, through examining the crime incident reports provided by Boston Police Department (BPD). The solution was developed using Python, Tableau & R. The project incorporated Large Language Model sentence transformer to group the crime using offense description to humanly understand the crime patterns.

EDUCATION

The Pennsylvania State University, Malvern, PA — Anticipated December 2024
Master of Professional Studies in Data Analytics
GPA: 3.87/4.0

AWARDS

Team of the Year Award ~ LTIMindtree, Global Finalist at NASA Space Apps Challenge Global Hackathon 2023, Global Nominee at NASA Space Apps Challenge Global Hackathon 2022, Runner Up 1 in Neurohack'22 ~ NAUT BU at LTIMindtree, Best Manager Award in training academy ~ LTIMindtree. 3rd in Presentation Competition 15 minutes to Fame

CERTIFICATIONS

TensorFlow Developer Certificate, Microsoft Associate AI Engineer, Microsoft AI-900 & AZ-900 Fundamentals, LTIMindtree AIOPs Training Certification

PUBLICATIONS

Real & Automated Meter Reading using Image Processing Reducing Human Error – IRJET (Volume 09 issue 11)