SREERAJ R M

Data Scientist / GenAI

Address : Trivandrum, Kerala, India
Phone : +91 8281243125
Email : sreerajfeb400@gmail.com
Linkedin : https://www.linkedin.com/in/sreeraj-rm-91302a1b4

Work Experience

**Data Scientist , EY GDS - 09/2022 to Present - Lead Data Scientist
for Global EY Capacity Management Team**

● Architected an AI chatbot using Azure Functions (Python) as the backend, ensuring data warehouse information was not exposed to the LLM.
● Leveraged Azure OpenAI Service with models like
GPT-4 for language generation and utilized Azure Cognitive Search as the vector store to enable Retrieval Augmented Generation (RAG).
● Integrated the chatbot with Microsoft 365 and deployed it to Teams for user access.
● Improved execution time by 50% through parallelization and implemented RBAC using metadata from the vector store.
● Created Random Forrest regression model to Optimize Infrastructures of Application and Servers.
● Created a combination of Statistical and Regression model to find the end capacity date for filesystem drives to reduce downtime of application across EY.
● Created numerous Power BI reports for Server and Application infra optimization.
● Worked in creating data that is consumed by Power BI using Databricks Pyspark
● Connected to Various data sources(databases,API's) and brought the data to ADLS.

**Jr Data Scientist , Sentient Scripts- 12/2020 to 07/2022 Client:
Dalmia Cements, Prologis**

- Developed a financial forecasting model for a real estate firm to project the next 10 years, integrating a what-if analysis feature. The solution leveraged Azure Functions for backend processing, with PowerApps seamlessly integrated into Power BI for interactive data visualization.

- Built an SVM model to accurately classify support tickets, ensuring they are routed to the correct queues, improving efficiency in ticket management.

- Designed and developed multiple Power BI reports for a construction conglomerate, covering areas such as daily sales, key performance indicators (KPIs), warehouse operations, and transportation.

- Developed Databricks notebooks to create the Gold layer of data for efficient consumption and reporting in Power BI.

Skills

- Programming languages: { Python : [Pyspark, Pandas]}
- SQL : [Window Functions, CTE, Indexing]
- Distributed Compute : [Spark]
- Data Warehouse : [Databricks SQL Warehouse]
- Cloud [Azure, AWS]
- Database : [SQL Server, PostgreSQL]
- Vector DB [Pinecone]
- RAG, Prompt Engineering
- Power Platform : [Power BI, Power Apps] (Basics)
- Machine Learning : [Regression, Classification]

Education

College of Engineering Trivandrum - **08/2016 to 08/2020**
Bachelor of Technology
Mechanical Engineering 7.77
CGPA

Kerala State Board - **08/2014 to 08/2016**
ST Mary's Pattom
97.25