VISHAL VALLAPU

940-758-4780 | vishalvallapu@gmail.com |Linkedin | GitHub| Portfolio

PROFESSIONAL SUMMARY

Data enthusiast with around 4 years of experience in data processing, analysis and visualization. Proficient in Python, SQL, Tableau, Power BI, with a strong background in machine learning and big data technologies. Aiming to leverage analytical skills and technical expertise to drive data-driven decision making. I am a promising team player who believes in continuous growth and learning

EDUCATION

Masters in Science, Advanced Data Analytics-University of North Texas - 4 GPA   *Aug 2022-Dec 23*
Bachelors in Technology ,Computer Science - Jawaharlal Nehru Technological University, India   *Aug   2014-May 2018*

EXPERIENCE

Zonestra Technologies *Software Engineer*   **Dec'23 - Present**
- .Built a robust data ingestion system using Python and Cloud Functions to fetch data from APIs, CSVs, and databases, reducing manual data handling .
- Implemented SQL queries and data transformations to correct missing data, handle duplicates, and create KPIs such as sales growth and churn rate, leading to a 20% increase in reporting accuracy.
- Scheduled daily and weekly reports on sales performance, inventory management, and customer behavior, cutting the report preparation time by 75%.

University of North Texas *Student Graduate Assistant*   **Sept'22 - Dec'23**

- Assisted professor in grading assignments and exams with attention to detail, ensuring fair and consistent evaluation of student performance.
- Provided timely feedback to students on coursework, clarifying expectations and offering constructive guidance for improvement.
- Facilitated effective communication between the professor and students, addressing queries and concerns via email and during office hours.

Ministry of Railways, Government of India *Software Engineer*   **July'20- July'22**
- Developed and implemented data pipelines to stream continuous data from sensors, ensuring real-time data collection and processing.
- Analyzed sensor data to predict wear and tear of machinery, improving maintenance scheduling and reducing downtime by 20%.
- Designed and scheduled data processing workflows to handle large volumes of streaming data efficiently.
- Stored processed data in a relational database, ensuring data integrity and accessibility for analysis.
- Created Power BI dashboards to visualize data insights, enabling stakeholders to make informed decisions.
- Collaborated with cross-functional teams to ensure seamless integration of the predictive maintenance system with existing infrastructure.

SKILLS

Programming and Query Languages: Python, Spark,Java, PySpark, Scikit, ML, Deep Learning (DL) and Natural LanguageProcessing(NLP),MachineLearning,(Supervised-Regression,Classification,Unsupervised-Clustering,Dimensionality Reduction)
Databases: MySQL, Postgres, Cosmos DB, Cassandra DB,Gremlin,MongoDB,
Cloud Technologies and Tools: Microsoft Azure Cloud,Google Cloud PlatformGCP,Amazon web servicesAWS.
Visualization, Analytics and Big Data Tools: Tableau, Microsoft Power BI, MS Excel, Hadoop, Spark, Amazon EMR, AWS Lakeformation,Redshift,Kinesis,AWS Glue,Apache Kafka,Quicksight,Athena,Open Search service, matplotlib,seaborn,SSIS ,SSRS, SSMS,Airflow,Rapid miner, Nifi, Git,Agile-Kanban, CI/CD Devops,Data Modelling
Soft Skills:Critical thinking, attention to detail, adaptability, and a strong work ethic.,Problem Solving, Highly organized with the ability to manage projects efficiently and meet deadlines consistently, Communication, Team player, Leadership, Time-Management, Willingness to learn and adapt to any new environment/tool.

TECHNICAL CERTIFICATIONS

Certified by IBM in CE-Essentials of Big data with Hadoop.
Certified by Udemy: SQL-MySQL for Data Analytics and Business Intelligence. Python for Data Science and Machine Learning Bootcamp.
Certifications by AWS : Fundamentals of Analytics on AWS ,Amazon EMR ,Amazon Kinesis , Data Analyst Quest Essentials on AWS , Planning on   large data migration to AWS
Certifications by Google : Introduction to Generative AI , Large Language Model ,Responsible AI by Google

KEY PROJECTS- http://github.com/vishalvallapu

Virtual Internships:

1. PWC in Power BI(Analyzing Customer behavior, Customer Analytics, Customer Churn analysis)
Completed a job simulation where I strengthened my PowerBI skills to better understand clients and their data visualization needs.

Demonstrated expertise in data visualization through the creation of Power BI dashboards that effectively conveyed KPIs, showcasing the ability to respond to client requests with well-designed solutions.

2. **Accenture in Data Analytics and Visualization**(Data preprocessing,analysis and visualization) by Forage.
3. **British airways in Data Science**(Web scraping and customer behavior analysis using machine learning models)

Azure end to end Data engineering project : The use case for this project is building an end to end solution by ingesting the tables from on-premise SQL Server database using Azure Data Factory ADF and then store the data in Azure Data Lake ADLS. Then Azure databricks is used to transform the RAW data to the most cleanest form of data and then Azure Synapse Analytics to load the clean data and finally using Microsoft Power BI to integrate with Azure synapse analytics to build an interactive dashboard. Azure Active Directory (AAD) and Azure Key Vault for the monitoring and governance purpose.

Azure Datafactory pipeline challenge: Azure Data Factory ADF Pipelines to prepare log data for analysis in Azure Synapse Analytics. The log data must be available in a specific storage account, blob container, and folder for Azure Synapse Analytics.

Baggage Handling System Optimization for American Airlines
• Aimed to address the challenges faced by an airline company with a shortage of personnel and equipment in its BHS.
• Predict and analyze the number of baggage items that will arrive for any given outbound flight in Python.

Netflix Data analysis: Created an interactive dashboard using Netflix dataset, analyzing top genres, descriptions, duration, release year, and various other variables in Tableau.

Smart Contract in solidity: Wrote a smart contract in Solidity using Remix IDE, Truffle-Ganache to mint coins, check balance, send and receive coins.

Crypto MVRV, Realized value calculation: The MVRV ratio and MVRV ratio z-score are calculated using the market value and realized value of Bitcoin and Solana, based on data collected over a 24-hour period.