**VISHAL VALLAPU**

940-758-4780 |[Email](mailto:vishalvallapu@gmail.com) |[Linkedin](https://www.linkedin.com/in/vishal-v-1a6632b6/) | [Github](https://github.com/VishalVallapu) | [Portfolio](https://vishalvallapu.github.io/)

# SUMMARY

Data Analyst with almost 4 years of experience in data processing, analysis, and visualization. Proficient in Python, SQL, Tableau, and 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.

# EDUCATION

|  |  |
| --- | --- |
| University of North Texas |  |
| *Masters in Science, Advanced Data Analytics-* | Aug 2022 - Dec 2023 |

|  |  |
| --- | --- |
| Jawaharlal Nehru Technological University |  |
| *Bachelors in Technology - Computer Science* | Aug 2014 - May 2018 |

# WORK EXPERIENCE

|  |  |
| --- | --- |
| Winfocus Solutions Pvt ltd |  |
| *Data Analyst* | Oct 2018 - Jul 2022 |

* Worked on the data processing of the SLAM (Software for Locomotive Asset Management) project which improved locomotive shed operations.
* Developed relational database solutions using SQL, improving data processing efficiency.
* Collected raw data from sheds, converted it into template forms, and loaded it into local servers, achieving a high data accuracy rate.
* Automated web services using provided API documentation, reducing manual intervention.
* Attended daily client calls, provided on-site demos, and trained end-users, resulting in increased user satisfaction.
* Identified gaps in received files and generated reports, leading to a reduction in data inconsistencies by 25%.
* Mailed weekly progress updates to relevant sheds, ensuring timely communication and project transparency.
* Leveraged data analysis to drive business decisions and improve operational efficiency.
* Identified and resolved data gaps and inconsistencies, ensuring data integrity.
* Worked closely with various teams, including client representatives and end-users, to ensure successful project implementation.
* Transformed complex data analyses into actionable insights, facilitating informed decision-making.

|  |  |
| --- | --- |
| Ministry of Railways, Government of India (contract) |  |
| *Software Engineer* | Jul 2020 - Jul 2022 |

* 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.
* Identified and resolved data gaps and inconsistencies, ensuring data integrity.
* Worked closely with various teams, including client representatives and end-users, to ensure successful project implementation.

# SKILLS

●**Programming and Query Languages:** Python, Spark,Java, PySpark, Scikit, ML, Deep Learning (DL) and Natural Language Processing (NLP), Machine Learning (Supervised-Regression,Classification,Unsupervised-Clustering,Dimensionality Reduction)

●**Databases:** MS Access,MySQL, SQL Developer, NoSQL(MongoDB, Dynamo DB,Cosmos DB).

**Data Modelling:**  Dimensional Modelling,Star schema,snowflake schema,Relational data modelling, Data Vault,Data warehouse methodologies:Kimball,Inmon

●**Cloud Technologies and Tools:** Microsoft Azure Cloud, Amazon Web services AWS, Google cloud platform GCP.

●**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.

●**Soft Skills:**Critical thinking, attention to detail, adaptability, and a strong work ethic. 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.

**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**.
* Virtual Internship by **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 visualisation 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.
* Strong communication skills reflected in the concise and informative email communication with engagement partners, delivering valuable insights and actionable suggestions based on data analysis.
* Leveraged analytical problem-solving skills to examine HR data, particularly focusing on gender-related KPIs, and identified root causes for gender balance issues at the executive management level, highlighting a commitment to data-driven decision-making.

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

# PROJECTS

|  |  |
| --- | --- |
| *Baggage Handling Optimization for American Airlines* | - |

* Aimed to address the challenges faced by an airline company with a shortage of personnel and equipment in its BHS.
* Predicted and analyzed the number of baggage items that will arrive for any given outbound flight in Python using Xgboost,decision tree algorithms

|  |  |
| --- | --- |
| *Netflix Data Analysis* | - |

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

*Data Engineering project on Crytpocurrency on GCP:*

* Collected cryto data using API to Bigquery, automated the ingestion using Cloud schdeuler , performed few calculations , created a function using Cloud Functions and presented MVRV ration in Looker Studio.