**Ravindra Kumar Velidandi**

Tampa, FL | +1 656 200 5481 | ravindravelidandi99@gmail.com| [linkedin.com/in/ravindravelidandi/](https://www.linkedin.com/in/ravindravelidandi/)

**SUMMARY**

Data Engineer with 4+ years of experience in designing scalable data pipelines and architectures, leveraging Azure and AWS Cloud Technologies. Proven track record in optimizing ETL processes and transforming raw data into actionable insights using Python, SQL, Spark and Data visualization

**EDUCATION**

**University of South Florida** Tampa, FL

Master of Science (MS) in Business Analytics and Information Systems.Aug 2022 - May 2024

**SKILLS**

* **Programming Languages:** Python, R, Pyspark, SQL, HTML, CSS, Java, Spark
* **Database Systems**: Oracle, MS-SQL, MySQL, NoSQL, Snowflake, Data Warehouse, Non-relational databases
* **Data Engineering:** SQL, Python, ETL, Databricks, Data Pipelines, Airflow Data Integration, Data Modelling, Data Wrangling, Machine Learning, Deep Learning, Pandas, NumPy.
* **Cloud Technologies:** Azure (Data Factory, Synapse Analytics, Databricks, Event Hubs, Data Lake Storage) SSIS, Azure Blob Storage, Business Intelligence, Azure HDInsight, Git, Airflow, Azure Blob Storage
* **Data Visualization and Reporting:** Tableau, Power BI, Qlik, looker

**Certifications**

AWS Data Engineer Associate, AWS Solutions Architect Associate, Tableau, Azure Fundamentals,

**WORK EXPERIENCE**

**Infosys Ltd, Client: Microsoft Ltd** **Nov 2020 – Jul 2022**

**System Engineer (Data Engineer)**

* Spearheaded the management and processing of **multi-terabyte data** from diverse vendors, ensuring data integrity, consistency, and reliability to enhance Microsoft's **customer data analytics** capabilities and drive informed decision-making.
* Integrated and transformed **customer data** of over **1 TB** from various sources, including SQL databases, NoSQL databases, cloud storage**, APIs**, XML, and JSON, using **Azure Data Factory** and **Azure Databricks**, ensuring comprehensive and accurate analytics for Microsoft customers
* Managed **MS-SQL** Server databases, ensuring high availability and performance tuning, which improved query execution speed by 50% and reduced downtime by 60%.
* Leveraged **Azure Synapse Analytics for scalable data warehousing and analytics**, optimizing performance and query execution, resulting in a 45% improvement in query times
* Designed and implemented modern **ETL pipelines with Apache Airflow**, resulting in a 50% reduction in data processing times and improved data accuracy.
* Engineered data pipelines using **REST APIs** to automate the ingestion of external data sources, enhancing data processing workflows and increasing data update frequency by 30%.
* Designed and implemented Spark applications for advanced **data wrangling and modeling**, leveraging Python, PySpark, Hive, and Spark-SQL to optimize data processing workflows and enhance analytical capabilities.
* Utilized **Git** for version control and implemented **CI/CD** pipelines, decreasing deployment time by 40% and enhancing team collaboration on code changes.

**Digital Stalk Private Limited May 2019-Nov 2020**

**Business Intelligence Analyst**

* Engineered Azure-based data pipelines using **Azure Data Factory**, **Databricks**, **Data Lake Storage**, and **Synapse Analytics**, driving a 25% increase in data processing efficiency
* Conducted **Power BI** deployments and optimizations, enhancing dashboard performance and usability, leading to a 20% increase in user satisfaction scores.
* Orchestrated **ETL** processes using **dbt**, transforming **raw data into structured formats**, resulting in a 30% improvement in data accuracy and reliability.

**PROJECTS**

**Bing News Data Analytics: Jan 2024-May 2024**

* Built an ETL pipeline using Azure Data Factory and **Service Fabric** to ingest and process real-time data from **Bing API**, optimizing data flow and storage in Azure SQL Database and Cosmos DB, achieving a 25% improvement in query performance.
* Built and optimized interactive dashboards with **Power BI** and **Azure Synapse Analytics**, providing actionable insights and enhancing real-time data analysis, reducing data latency through efficient data integration and processing.

**Flight Status Prediction and Delay Analysis (Pandas | Scikit learn | Machine Learning) Aug 2022- Dec 2022**

* Analyzed flight status for the top 25 busiest domestic routes in the US, using Bureau of Transportation Statistics data to predict delays and cancellations, incorporating factors like holidays, seasons, and historical trends, improving predictive accuracy by 15%.
* Implemented machine learning models such as Decision Tree, Random Forest, and Gradient Boost, achieving an F1 score of 84.35%, and focused on reducing false positives and negatives by 20% to enhance airline operational efficiency and customer satisfaction.

**Kickstarter Campaign Prediction (R | Machine learning | Statistics) Aug 2023- Dec 2023**

* Conducted a comprehensive analysis of over 1,000 Kickstarter campaigns, applying **web scraping, social media analysis**, and advanced **text analytics**. Developed a predictive model with over 85% accuracy, identifying key success factors such as project descriptions, funding targets, and creator credibility.