VenkataRamana Ravuri
Email: venkataramanaravuri800@gmail.com
Phone: +1(972) 428-3834

PROFFESSIONAL SUMMARY:

- Over 8 years of IT Experience in Designing, Development, Administration, Implementation of Data Warehouse & Data marts with Tools like Azure Data Factory, Informatica Power Center, OLTP and OLAP using Data Extraction, Transformation, Loading and Analysis in various domains like Banking, healthcare and Insurance.
- Extensive experience in working with business stakeholders in gathering business requirements and prepare Functional Specification Documents (FSD), Business Requirement Documents (BRD), logical and Physical data models.
- Experienced in Extract, Transform and Load (ETL) data from different Sources Systems to Azure Data Lake Storage (ADLS) using a combination of Azure Data Factory (ADF), Spark SQL and U-SQL Azure Data Lake Analytics and processing the data in In Azure Databricks.
- Expertise in various Hadoop distributions like Cloudera, Hortonworks distributions and Azure.
- Experience in creating Hive Internal/External tables, loading with data and troubleshoot with Hive jobs.
- Experience in writing complex Hive Queries for various business use cases. Extended Hive and Pig core functionality by writing Custom UDFs.
- Extensively worked in developing Stored Procedures, User Defined Functions (UDF), Views and Triggers on SQL Server, Oracle, DB2 and Teradata databases.
- Experienced in providing Lights On support for ongoing production issues and bug fixes for reporting applications on Data Warehouse.
- Experience in Writing Producers/consumers and creating messaging centric applications using Apache Kafka.
- Good understanding of Data concepts, principles, approaches, Data mart, Star Schema, Snowflake Schema & Galaxy Schema, Fact/Dimension tables and Normalization/De normalization of data.
- Good understanding of Informatica Cloud, Informatica Data Quality, Analyst and Power exchange tools.
- Experience in using Automation Scheduling tools like Autosys, Tivoli job scheduler, Workload Manager and Control – M.
- Extensively worked on projects in the SDLC, Incremental and Agile approach.
- Prepared new and modified existing dashboards and reports on Tableau Desktop and SAP BO.

TECHNICAL SKILLS:

BI (Reporting & Dashboard) Tools:	Tableau 9.3/10, Excel, PowerBI
Databases:	Oracle 10g/11g, MS SQL Server, HIVE, MySQL
Data warehouse / ETL Tools:	Informatica, Talend, Azure Data Factory
Languages:	Java, Python, Scala, PL/SQL, T-SQL, Java, UNIX Shell Scripting, R
Big Data/Hadoop framework:	Map Reduce, Spark
Job Scheduling:	Airflow, IBM Tivoli Manager 5.0, Autosys, Control-M, Mainframes.
Data Modeling:	ERWIN, ERStudio 7.1/6.1
Cloud Technologies:	Microsoft Azure, AWS, Databricks

Client: Ambercrombie & fitch   **Jan 2023 – till date**
Role: Data Engineer

Responsibilities:
- Build Data pipelines using Azure Data Factory (V2) to Extract Load and Transform data from Sources like Oracle, Teradata, SQL Server, SFTP and Amazon S3 into Azure Blob storage and Azure Data Lake (ADLS)
- Automated the process of data gathering, cleaning and analysis of organization data which saved 4M$.

- Designed data pipelines for analytics to provide efficient insights into performance improvement, key business performance metrics and product optimization.
- Played integral role in increasing data visibility and building dashboards to drive 2B$ in revenue which helped business leaders make informed decisions.
- Utilized SnowSQL (CLI Client) to connect to the data warehouse and performed bulk loading from external stage to internal stage to snowflake cloud using COPY command.
- Performed Data transformation using Data Frames on Azure Data bricks.
- Scheduled ADF pipelines to execute automatically using the Triggers.
- Developed new Mappings, Mapplets and Workflows to transform and migrate the data coming from various sources using Informatica Cloud and load the data into Azure SQL Server.
- Used control flow and data transformations like Copy Data activity, Get Metadata Activity, lookup, Sort, Join, Store procedure, for each etc. To build the Pipeline activities.
- Performed data profiling on Informatica Analyst.
- Wrote ad-hoc SQL queries to override the ETL logic in mappings during UAT testing.
- Scheduled new Informatica workflows using Control – M by making use of in-build calendar.
- Prepared new reports on Power BI using the summary data from Azure SQL Data warehouse.
- Prepared Unit test cases to test the new creating ETL mappings for data quality.
- Working Knowledge on Polybase, Azure VM, Logic Apps and Azure Dev-Ops (Repository, CICD).

Environment: Informatica Power Center 10, Oracle 10g, DB2, Flat Files, Mainframe, PL/SQL, SQL Server, Azure Data Factory (ADF), Azure Data warehouse, Azure Data Lake V2 (ADLS), Power BI, SQL Server, Snowflake, Python.

Client: Dematic   **may2020 – Dec 2022**
Role: Data Engineer

Responsibilities:
- Retrieved relevant data by extracting raw data from disparate sources to develop strategic solutions for unbounded (streaming) and bounded(batch) with Spark, Kafka, HBase.
- Involved in designing dimensional modelling and implemented STAR flake, SNOWflake schemas.
- Assisted in designing Logical/Physical Data Models, forward/reverse engineering.
- Performed data migration and validation testing on Source and Target data sets for all the workflows.
- Improved accessibility and security by initiating a Data Lake strategy ; ensured optimum service by deploying Kafka and Flume to ingest data , using Apache Spark to transform data and loading downstream applications for reporting.
- Enhanced productivity by creating distributed data processing solutions for MiFiD regulatory compliance.
- Designed and implemented a Data Lake strategy to improve accessibility and security.
- Refined efficiency by migrating existing workloads to use AVRO and ORC format instead of XML and JSON which reduced the storage space by 80% and running time by 40%.
- Created executive dashboards on Tableau desktop 9.3 and deployed them on to Tableau Server.

- Wrote Stored Procedures in PL/SQL and implemented them through the Stored Procedure transformation.
- Improved effectiveness by enhancing present Spark jobs to meet the SLA of 15 minutes which minimized processing time by 60%.
- Provided bug fixes to fix the production issues.
- Developed UNIX shell scripts for event waits, wait for files and pre-session tasks.
- Performed Unit testing of new ETL changes to test the data in source and target tables.

Environment: Hortonworks Sandbox, UNIX, Kafka, Spark, Hive, HBase Putty, Flume, Tableau, SQL, PL/SQL and UNIX Shell Scripting.

Client: Amex   **feb2017– April2020**
Role: Data Engineer

Responsibilities:
- Tuned and optimized the performance of existing SQL queries by creating necessary indexes and keys on database tables.
- Optimized technical performance by interpreting data needs into technical system requirements ; also wrote scripts and employed various analytical tools to solve complex problems.
- Efficiently utilized Hive and Pig to model pre-processing and transform data to meet business requirements ; also collaborated with team members to ensure best performance.
- Created dynamic dashboards on Tableau by extracting the data.
- Developed Map Reduce, Spark jobs and Hive queries to process 500GB to 1TB of data comprising of multiple file formats.
- Created Command tasks to run scripts to truncate the Work Tables, Landing Zone Tables, Staging Area tables etc.
- Facilitated the execution of projects by analyzing extensive data stores , correcting errors and proposing solutions suitable for each project.
- Involved in testing the SQL Scripts for report development, Tableau reports, Dashboards, Scorecards and handled the performance issues effectively.
- SQL Queries and Stored procedures were used to extract data and perform tasks like Aggregations on Customer data.
- Provided support to resolve unplanned environmental outages that caused batch issues.
- Prepared project documentation and support documents to train the business users.

Environment: Spark, Kafka, Hive, Map Reduce, Java, AWS, UNIX, Putty, Tableau, SQL, PL/SQL,UNIX Shell Scripting.

Client : VignaSoft Tech, India   **May 2014– June 2016**
Role: Data Engineer

Responsibilities:
- Worked with business users to gather, define business requirements and analyze possible technical solutions.

- Used Cloudera distribution for data transformation and data preparation.
- Validated and recommended on Hadoop Infrastructure and data center planning considering data growth.
- Transformed data to and from the cluster , using Sqoop and various storage media such as flat files.
- Developed Map Reduce programs and Hive queries to analyze patterns and customer satisfaction index over the data present in various relational database tables.
- Developed Pig scripts in the areas where extensive coding needs to be reduced.
- Involved in Agile methodology for the entire project.

Environment: Cloudera, Pig, Hive, Map Reduce , Sqoop.