GATTU NAVANEETH RAO

Navaneeth.dwh@gmail.com +91 7306389971

LinkedIn**:** linkedin.com/in/navaneeth-rao-3a2616237

Professional Summary:

- 5.5 years of IT experience in data-driven application design and development.
- 4.4 years of experience in Azure data services and Talend.
- Proficient in Azure technologies such as ADF, ADB, Azure Synapse Analytics, and Azure Storage.
- Hands-on experience with Unity catalog, Autoloader, and Delta Live Tables (DLT).
- Experience with Apache Hadoop frameworks and Microsoft Azure Cloud.
- Proficient in Big data ingestion tools like Kafka, spark streaming, and Sqoop.
- Experience with Big Data distributions like Hortonworks (Hortonworks 2.1) with Ambari.
- Hands-on experience in application development using Java, RDBMS, and Linux shell scripting.
- Experience with Tableau and Power BI to generate reports.
- Good understanding of Dimensional Modelling, Star and Snowflake Schemas
- Experience with Informatica Power Centre for data integration and ETL.
- Experience in deploying SSRS Reports on Report Manager using SSDT.
- Good knowledge on OLAP Concepts and SQL Server.

AI & ML

- Working experience with data science python libraries like (NumPy, Pandas, Kera's, Theano, Caffe, Open AI Gym, SciPy, Tensor flow…. etc.).
- Created models like regression, classification, LSTM, CNN (Convolutional neural network), ANN (Artificial neural network), Clustering model, N degree polynomial model…. etc.)
- Did some explorer towards problem solving using AI.
- Emotion AI: Detecting facial expressions and declaring emotion of human with CNN, ANN and ResNet
- AI in health care. (Classifier model): To localize brain tumors and giving treatment early to avoid serious injuries and helpful to save lives.
- AI in business marketing (K – means clustering) Understanding the business in market like customer interests.
- Creative AI (Deep dream algorithm): Applying AI in culture, creating paintings, write compelling stories and compose a new music.
- Ability to adapt any new technology quickly.

Professional Experience:

- Worked as a Associate at "Wipro Ltd." from Oct 2019 – Dec 2021.
- Working as a Azure data engineer at "Luxoft a DXC Technology Company" from Oct 2021 – Till Date

PROJECT DETAILS:

Project #1: EDWD - Client: VFS Global, Role:	
PySpark, python, Unity Catalog, data bricks, ADF, Azure, Delta, Autoloader, spark Streaming, MS Fabric, Medallion Architecture- Environment	
Role	Azure Data Engineer
Description:	VFS Global, also known as "Visa Facilitation Services Global," is an outsourcing and technology services company for governments and diplomatic missions worldwide. The company manages visa and passport issuance-related administrative and non-discretionary tasks for its client governments
	1.Developed pipelines in Azure Data Factory to fetch data from sources and load into Azure Synapse Analytics. 2.Implemented Medallion architecture in 3 layers: Bronze, Silver, Gold. 3.Experienced with Unity Catalog for data governance and access control. 4.Worked with Azure DevOps for infrastructure setup, deployment, and testing. 5.Implemented incremental load strategy and Autoloader for streaming data. 6.Utilized Delta tables and Azure Data Lake Storage (ADLS) for data processing. 7Automated code with workflow tasks and Data Factory notebook activities. 8.Created Linked services, datasets, and pipelines in Azure Data Factory. 9.Developed stored procedures using TSQL and created copy, lookup, and metadata activities. 10.Monitored pipelines, reported issues, and performed bug fixing. 11.Created data flows for transforming data to Azure using Azure Data Factory. 12.Implemented dynamic dataset components and triggered pipelines based on schedules. 13.Prepared external tables, dimension, and fact tables in Azure SQL. Developed business logic using stored procedures in Azure SQL Analytics

Project #2: HPI- Tera Restatement, Client: HPI	
• Environment	• Azure data factory, Azure data bricks, Pyspark, Unity Catalog, Spark Streaming, Delta Live Tables (DLT), Azure SQL DB, ADLS 2, Talend (7.3), Hive, Hadoop, Linux,
Description:	The Business process involves adding three more restated fields and loading of data from Ingestion Hive to Distillation Hive table. There after data is exported in the form of Parquet file and pushed to Azure Data Lake Store Gen2 – Landing container. Parquet file is then loaded into ADW and ADB layer which is consumed by Power BI tool to generate dashboards which are used by end users for analysis and decision making.
• Role	• Azure Data Engineer
Roles and Responsibilities	ETL workflow using Azure ADF, Databricks, and PySpark. Data governance with Unity Catalog for data sharing, access, and masking. Data transformation and loading into ADLS Gen2 using PySpark. Automated code with Azure Data Factory and notebook activities. Worked on Autoloader for streaming data and Azure SQL DB. Data migration from On-Prem to Azure Cloud with Databricks. Parameterized hive scripts and executed using Talend. Extracted data from hive and transformed using Talend components. Implemented performance tuning logic and optimized targets and sources. Deployed and ran TAC jobs for ingestion and distillation. Loaded data from files to tables using Talend components. Used various Talend components for data processing and loading. Shifted partitions in hive using Talend components. Scheduled Talend jobs using TAC and monitored daily loads.

Project #3: SDM	
Role & Environment	Data Engineer , Talend 7.2 Big Data (HDFS, Hive), UNIX.
DESCRIPTION: -	what they want most. And their real- time approval technology helps them do it all on the spot. • Segmentation data mart (SDM) delivered core capabilities to support the Email and Direct mail marketing campaigns. Core capabilities includes Universe of marketable contacts, Account association, Job title categorization, Sales rep assignments, Suppressions, 3rd party list loads and integration to Eloqua.
Roles and Responsibilities	• Creating the Jobs to extract the data from source to target using Talend tool. • Extracted data from Oracle database and Flat files and staged into a single place and applied business logic to load them in the Hive. • Implemented performance tuning logic on targets, source to provide maximum efficiency and performance. • Parameterized the Talend Jobs and Increased the re-usability. • Applied slowly changing dimensions like Type 1 effectively to handle the Delta Loads. • Developed Talend jobs and published into TAC.

Project #: EPICX	
Role & Environment	Data Engineer., Talend 7.3, Unix, Aws, Sftp, Snowflake, Redshift, Athena, S3, Sap
DESCRIPTION:	Integrating the data from different source systems to SFTP location and making available this data to all business users across the world and respected stake holders for their business decisions.
Roles and Responsibilities	1. Prepared jobs using Talend components (tFineInOutDelimited, tFileList, etc.). 2. Dynamic column mapping with tfineinoutdelimited using dynamic schema. 3. Captured pipeline logs and execution statistics. 4. Added triggers for job execution in TMC. 5. Extracted data from Azure Data Lake Storage, transformed with Talend (tmap, tjava, etc.). 6. Loaded transformed data into Oracle. 7. Implemented dynamic data set components. 8. Scheduled job triggers. 9. Created external tables, dimensions, and fact tables in Azure SQL. 10. Resolved bugs.

EDUCATIONAL QUALIFICATION

- B. Tech from Jbiet.

TECHNICAL SKILLS:

- **AZURE**: Azure Data Factory (ADF), Azure Data Bricks, Unity Catalog, Delta lake, Azure Data Lake Services (ADLS), Azure Blob Services, Azure SQL DB, Azure Active Directory (AAD), Azure Devops. MS Fabric
- **Languages**: Scala, Core Java, Python
- **Databases**: Azure SQL, My SQL, Postgres SQL, Oracle.
- **No SQL Databases**: HBase, MongoDB.
- **Data warehouse**: Hive, Azure Synapse, redshift, Snowflake.
- **Big data** Technologies: Apache Hadoop, Apache Hive, Apache Spark.
- ETL Tools: Azure Data Factory (ADF V2), Talend 6.4/7.3/ Big Data Enterprise & open Source), Informatica
- Data ingestion: Sqoop, Kafka, Spark Streaming
- Data Visualization**:** Tableau, power BI

---------Thanks.,

G Navaneeth Rao