

**Decisive, strategic and performance-driven professional**, targeting assignments in **Data Engineering** with an organization of high repute for mutual growth.

+919959029291



## PROFILE SUMMARY

- ❖ **Result oriented professional** offering **9 years of total experience and 6 years** of relevant experience in **data engineering and cloud development** as well as in **Application Development focusing on Big Data Technologies and Mainframes** in domains such as Retail, COMMS/MEDIA & BNFS.
- ❖ Proficient in **leveraging Big Data technologies** such as Apache Spark, Airflow, AWS, and Snowflake to develop robust data pipelines and data marts, driving actionable insights and business growth.
- ❖ **Built** scalable, efficient and fault-tolerant **data pipelines, and built enterprise data products** by leveraging AWS services such as **S3, EC2, EMR, Glue, Lambda, Cloudwatch** etc..
- ❖ Skilled in building **customer data mart**, successfully implemented **customer data platform(CDP)** for enhanced customer insights, analytics and provide **omni-channel** experience to customers.
- ❖ Involved and successfully implemented **data integration** with other platforms viz., salesforce marketing cloud (**SFMC**) thus powering marketing emails to customers for product recommendations, personalised discounts etc.
- ❖ Expertise in data ingestion, data processing, data warehousing and **data modelling**, thus building **enterprise data products**.
- ❖ Instrumental in consuming data from files, databases, API's, kafka topics for batch processing applications by leveraging efficient load strategies and load quality, aggregated data into data warehouse.
- ❖ Excellent in bringing wide variety of data into data lake, while maintaining a focus on data quality and business logic by using efficient and optimized, compressed storage file formats such as parquet, avro.
- ❖ Collaborated with cross functional teams to drive the project end-end by following **functional and non-functional requirements** such as scalability, reliability, data integrity, data security etc..
- ❖ Successfully modelled the data by following dimensional modelling techniques and implementing star, snowflake schema for reporting, analytics, and machine learning applications.
- ❖ Involved and **created KPI's** such as Search KPI, Viewership KPI etc..by leveraging events data thus helping business stake holders to take actionable insights from reporting dashboards.
- ❖ Successfully implemented and integrated data to our DIY website to display personalised discounts to discount sensitive customers by exposing model generated data using Mulesoft after performing load-testing, stress-testing.
- ❖ Skilled in leveraging **S3 storage integration, Snowflake stored procedures, tasks** etc. for building robust data pipelines in snowflake.
- ❖ Led the successful execution of critical projects, including **"Iris," "Discount Preference," and "HyperX,"** displaying a deep understanding of data processing, ML model implementation, and database management.
- ❖ Involved in consuming google analytics data, maintaining production data pipelines and addressing them in-case of failures.
- ❖ Rich & qualitative experience in **Data Governance**; establishing standards, policies & procedures.
- ❖ **A keen planner, strategist & implementer** with expertise in devising **framework & roadmap** for Data Management, Data Migration, Data Ingestion, Data Integration, Data Governance, Data Quality, Master Data Management & Data Security.



## CORE COMPETENCIES



Data Analysis & Management



Data Modelling & Data Curation



Data Integration & Management



Data Integration



Data Warehousing / API Integration



Data Engineering



Cloud Development



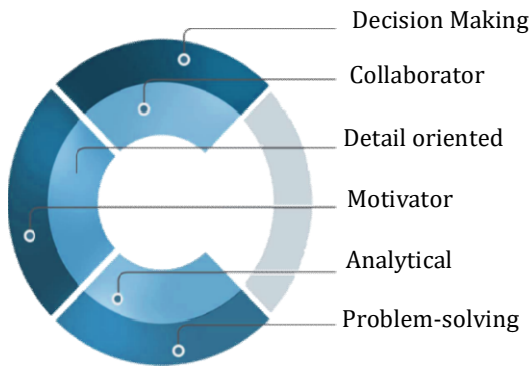
Cloud Migration



Data Security & Data Quality



## SOFT SKILLS



## EDUCATION



**Bachelor of Electrical and Electronics Engineering**  
from GMR Institute Of Technology in 2014



## IT SKILLS

- ❖ Pyspark, Python, Snowflake, AWS, Airflow, Git, Jenkins, Hive
- ❖ Scala, SQL, NoSQL, Data Modelling and Azure



## WORK EXPERIENCE

**Since Jan'21: Advance Auto Parts**

**Growth Path:**

Since Jan'21 as

**Senior Data Engineer**

**Key Result Areas:**

- ❖ Developing and implementing robust data architecture solutions to meet business requirements.
- ❖ Designing and optimizing data models for efficient storage, retrieval, and processing.
- ❖ Actively creating and managing data integration processes to ensure the seamless flow of data across various systems.
- ❖ Administering and maintaining databases, ensuring high performance, availability, and security.
- ❖ Collaborating with cross-functional teams to ensure effective API implementation.
- ❖ Implementing data security measures to protect sensitive information and ensure compliance with regulations.



## PREVIOUS EXPERIENCE

**Jun'20 – Dec'20: Tavant as Senior Software Engineer**

**Feb'19 – Mar'20: Cognizant as Associate Projects**

**Jan'15 – Jan'19: Syntel as Software Engineer**



## PERSONAL DETAILS

**Date of Birth:** 13<sup>th</sup> September 1993

**Languages Known:** English, Hindi & Telugu

**Address:** Hyderabad, India

*\*Refer to Annexure for Major Projects*

## ANNEXURE MAJOR PROJECTS

### Projects Undertaken:

#### Title: IRIS

**Description:** The objective of the identity resolution project is to establish a Customer Data Platform (CDP), facilitating the aggregation of customer information from diverse touchpoints, including POS, e-commerce, loyalty, and marketing channels.

A unified identifier, also known as an omni ID, is generated for each customer by consolidating data from various sources. The Customer Data Platform (CDP) holds significant potential, empowering stakeholders and businesses to align the unified ID with their specific use cases. This alignment enables the recommendation of products, generation of reports, development of dashboards throughout the enterprise, targeted user engagement, and driving sales across omni-channels. Ultimately, this initiative aims to deliver a seamless omni-channel experience to customers.

#### Highlights:

- ❖ Designed and implemented a customer data-mart (CDM).
- ❖ Created data pipelines that consumed data from multiple sources using pyspark and orchestrated the entire pipelines using Airflow.
- ❖ Conducted data profiling and cleansing, loading the cleansed data into S3 in parquet format, and into Snowflake tables, enabling data science and decision science teams to run models, build reports, dashboards, & so on.
- ❖ Created a data dictionary for the project.
- ❖ Developed custom attributes for email engagement, loyalty customers, & so on.
- ❖ Created sfmc ingestion pipelines, enabling the marketing team to send emails to customers who were 'ENGAGED' based on their current OPT status.
- ❖ Orchestrated data pipelines to pull in incremental data on a daily basis and sent them to SFMC, facilitated the team in launching marketing campaigns, & so on.

#### Title: Hyper Rec

**Description:** This project facilitates the recommendation of personalized products to online customers. Utilizing in-house developed models like NPTB, FTJ, Trending Now, etc., the project leverages feed files obtained from the e-commerce team to make product recommendations. These feed files are also shared with a third-party vendor and SFMC, responsible for sending emails to customers. This approach enhances the overall customer experience across online and offline channels, contributing to increased revenue compared to generic emails sent to all customers.

#### Highlights:

- ❖ Designed and implemented data pipelines that sent data to SFMC and a third-party vendor for an email program.
- ❖ Created a data pipeline that brought in e-commerce data and loaded it into S3 and Snowflake on a weekly basis.
- ❖ The vendor initiated an API call to AAP-developed in-house models via the mule layer, receiving SKU information, which was then passed to SFMC.
- ❖ SFMC already possessed meta-data information for the recommended SKUs and sent emails.
- ❖ Collaborated with multiple cross-functional teams and stakeholders in setting up the data in the enterprise data lake and data extensions in SFMC.

#### Title: Google Analytics

**Description:** This project aids in monitoring the customer journey on our B2C DIY website. It allows us to retrieve customer information through Tealium tagging mechanisms. These Tealium tags gather additional data beyond the default information provided by Google and transmit the data to Google BigQuery. We process this data, categorize it into subject areas, and then load it into Snowflake. Various teams throughout the enterprise utilize this data to develop use cases, reports, custom dashboards, leadership KPIs, and more.

#### Highlights:

- ❖ Designed and implemented a data model for GA universal, thereby constructing the Google data mart.
- ❖ Created data pipelines that consumed raw JSON data from Google BigQuery and loaded the last 3 days' worth of data into S3.
- ❖ Created data pipelines that fetched raw data from S3 into corresponding subject areas, thus loading the data into Snowflake tables for analytical purposes.
- ❖ Conducted data profiling, cleansing, and loaded the data into Snowflake tables, allowing data science and decision science teams to run models, build reports, dashboards, etc.
- ❖ Created a data dictionary for the GA datamart.
- ❖ Recommended products to customers based on user activity, specifically during checkout or when adding items to the cart, resulting in increased revenue, and so on.
- ❖ Decision science could build leadership KPIs, such as user count, transaction revenue, etc.

**Title: Max Search KPI**

**Description:** This project plays a crucial role in assessing the effectiveness of the search bar on the HBO MAX platform. It involves the analysis of search results using various search attributes. User activity within the platform is monitored to acquire customer insights and comprehend customer behavior, ultimately aiding the business in reaching its targets. Metrics are computed based on search behavior for a product, and its performance is evaluated for future projections. Key search attributes, such as playback\_conversion, watchlist\_conversion, and search initiation, are considered in this evaluation. The overarching goal is to enhance the business by delivering improved search performance to customers

**Highlights:**

- ❖ Designed and created a design document through interactions with the business.
- ❖ Loaded raw user activity data from the S3 bucket to the Snowflake intake schema by establishing a stage in Snowflake.
- ❖ Flattened the raw Telegraph JSON data and loaded the transformed data into a Snowflake table.
- ❖ Created SnowSQL scripts to transform the data based on requirements, utilized Snowflake-based transformations.
- ❖ Created and scheduled a DAG to monitor the daily activity of a user at the session level.
- ❖ Created a summary table to upsert the data from a daily temporary view and shared the table with the business to comprehend the search KPI.
- ❖ The business generated a Looker dashboard based on a summary table to analyze search performance for user sessions.