



# VASANTH KUMAR

## SUMMARY

Data engineer seeking to leverage 4+ years of experience in building data pipelines. Good understanding of analyzing big datasets, finding patterns and insights within unstructured and structured data pipelines as ETL Process. Knowledge in Handling Components and Big Data Optimization. Good team player and able to work in the individual environment

## CONTACT

PHONE NUMBER:  
7639-55-4131

LINKEDIN:  
<https://www.linkedin.com/in/vasanth-kumar-kuppusamy-598367196>

MEDIUM:  
<https://medium.com/@vasanth370>

EMAIL:  
Vasanth370@gmail.com

## ACTIVITIES AND INTERESTS

Exploring new technologies  
Sports activities  
Article Blog writing  
Block chain

## WORK EXPERIENCE – 4 YEARS – (SERVING NOTICE PERIOD)

**Senior Software Engineer**  
**Photon Interactive, Chennai.**  
May 2021 – Present 2022

Worked for client Walgreens company (FMCG), which operates the 2nd-largest online E – commerce and retail pharmacy store chain in the USA. It specializes in filling prescriptions, health and wellness products, health information, and photo services.

- Designed and implemented a real-time data pipeline to process semi-structured data by integrating 150+ million raw records from 14 sources.
- Integrated COVID19 dataset across the world and another dataset consist of various life factors, scored by the people living in each country data from Walgreens mobile and web app. merge these two datasets to see if there is any relationship between the spread of the virus in a country and how happy people are, living in that country.
- Lead the role as Airflow DAG developer Admin and managed entire team ETL Tasks.
- Architected the solutions to the Omniture tracking for Mobile and web app's which has more than 10+ million users to collect real-time tracking data. helps to improve future product sales strategies.
- Automated ETL process across billions of rows of data, which saved 45hours of manual hours per month.
- Have configured connection setup with Data Management Portal with REST API call for extracting semi-structured data in JSON formatted data into a raw zone (Pre-landing zone).
- Raw data extracted to Landing and Landing merge layer based on logical architecture design with dynamic schema evaluation and have maintained in S3 Lake by increased performance and scaled from 0 to 80,000+ hourly users.
- Have working experience in Databricks, PySpark, Spark-SQL, Airflow, Python, Java, S3, ETL, VCS - Git - Big Bucket- Git Lab, REST API, GraphQL. Having Good Knowledge in Handling Both Front-end UI Optimization and Big Data Optimization.

**Software Engineer**

## **RMADE Technologies Pvt, Ltd. Chennai**

May 2018 – May 2021 - 3 years

- RMADE Technologies is Product based GPS Tracking company with its own GPS manufacturing and developing Applications for Tracking vehicles with live movements of your fleet in real-time with 50K+ users.
- We collect millions of data every day and use big data platforms for storage and analytics to improve sales and future product innovation.
- Designed the data pipeline architecture for a new product that quickly scaled from 0 to 60k daily users using spark, SQL, Python.
- Closely worked with Architects to Improved Application performance and cost-cutting by converting from native frameworks to hybrid cross-platform frameworks.
- Published a few articles Blogs related to Spark and Front - end development in The MEDIUM helps many developers to learn basics easily.

## **EDUCATION**

---

### **Mahalingam College of Engineering and Technologies**

2014 - 2018

Electronic and communication Engineering

#### **Projects:**

- Smart Class Room using Raspberry Pie. This project aims at replacing the blackboards with TFT Liquid Crystal Display for delivering lectures to the students in order to make class hours more efficient and to aid the handicapped. Duration of 12 months using Linux software and Raspberry Pie Board. Which was Totally funded by our college for R&D and implementing the concepts in future classes.

## **KEY SKILLS AND CHARACTERISTICS**

---

- Pyspark, Spark
- SQL
- Python, Java,
- Airflow, Databricks, Jupiter notebook
- S3, ETLs
- VCS – GitHub, Gitlab, Big Bucket
- Rest API, GraphQL
- TypeScript, JavaScript
- Html, CSS