

**Name: Vijaya Kumar Gorla**  
**Email: vijayakumarg596@gmail.com**  
**Phone: 7093683021**

### **Summary:**

- Data professional with about 3+ years of work experience in building solutions around data management and data processing
- Extensive work done on writing queries for business needs and data processing
- Extensively used SQL, Pandas, PySpark for Data Analysis and Model building.
- Explored Kafka, using Python, SQL, and database integration for real-time data processing and backend development.

### **Technology Skills:**

Language	:	Python
Databases	:	MySQL, Oracle SQL
Data Visualization :		Matplotlib
Libraries	:	Pandas, Pyspark, Kafka, OpenCV, Pillow, Boto3
IDE'S	:	Jupyter, VS code, PyCharm, IDLE
AWS Services	:	S3, EC2, Lambda, EMR, IAM, EventBridge
Operating Systems :		Windows, Ubuntu 20.04

### **Work Experience:**

Worked as a Data Engineer at Aroha Technologies from January 2024 to September 2025.

Worked as a Software Engineer at Calligo Technologies Pvt Ltd from January 2023 to October 2023.

Worked as an Associate Software Engineer at Rudra Coresoft Technologies Private Limited from Oct 2021 to Oct 2022.

### **Educational Background:**

<b>Qualification</b>	<b>Board/University</b>
B.Tech(CSE)	Siddharth Institute of Engineering and Technology, Puttur, AP, India. (Affiliated to JNTU Anantapur)

## **Project Details:**

### **Project 1:**

**Project Name:** CYSD

**Client:** CYSD (Centre for Youth and Social Development)

**Technology:** Python, Pandas, Azure Blob Storage, Virtual Machine, MySQL

### **Project Overview:**

CYSD is a non-profit organization working towards sustainable development initiatives. As part of a PoC, our objective was to process large volumes of survey and beneficiary Excel data stored in Azure Blob and transform it into structured formats for reporting and analysis.

### **Responsibilities:**

Developed and deployed ETL pipelines on a Virtual Machine to process Excel files from Azure Blob Storage.

Implemented complex data transformations including removal of Odia text, handling multiple sheets, and resolving multi-header row challenges.

Utilized Python and Pandas to clean, standardize, and structure the raw data for downstream processing.

Designed and loaded data into Dimension and Fact tables to support business intelligence and analytical reporting.

Performed data validation checks such as record counts, duplicate removal, and data consistency.

## **Project 2:**

**Project Name:** Protium

**Client:** CreditNirvana

**Technology:** Python, Pandas, PySpark, AWS (S3, EMR, Lambda), MySQL

### **Project Overview:**

Protium Finance Private Limited is a financial services company providing loans to customers. In collaboration with Credit Nirvana, Protium aims to enhance customer communication and reduce penalties through timely messages.

## **Responsibilities:**

As part of the project, I designed and implemented robust data pipelines using AWS S3 and EMR to ensure efficient data storage and processing. Leveraging PySpark, I managed large-scale data processing, enabling effective data transformation and analysis. I automated the data extraction, transformation, and loading (ETL) processes, which significantly reduced manual intervention and improved data accuracy. Using Pandas and PySpark, I managed and maintained large datasets. I integrated AWS services, including S3 and EMR, for scalable data processing and storage, and optimized data workflows to improve processing times and overall system performance. Additionally, I collaborated closely with the Credit Nirvana team to ensure the seamless delivery of messages to Protium customers, contributing to a reduction in customer penalties.

## **Project 3:**

**Project Name:** Automatic Number Plate Recognition (ANPR)

**Client:** Calligo Technologies Pvt Ltd

**Technology :** Python, pandas, OpenCV-python, yolo

### **Project Overview:**

Designed and developed a robust and efficient Automatic Number Plate Recognition (ANPR) system utilizing computer vision techniques and machine learning algorithms. The project aimed to accurately detect and recognize license plates from images and video streams, contributing to enhanced vehicle monitoring, security, and traffic management.

### **Key Responsibilities and Achievements:**

Data Collection and Preprocessing

Algorithm Selection and Model Architecture

License Plate Detection

Character Segmentation and Recognition

Real-time Processing and Integration

Performance Evaluation and Testing

## **Project 4:**

**Project Name:** Gopuff

**Client:** Rudra Coresoft Technologies Private Limited

**Technology :** Microsoft SSIS, SQL

GoPuff is a digital delivery service designed to deliver daily essentials within minutes.

### **Responsibilities:**

-We load data into the input tables from the generated CSV files and the data will be loaded into staging and output tables also, based on running some stored procedures and we validate data in the staging table by executing the Data Transformation package task.

- If we have loaded the data successfully into the tables then we move to the next occurring packages to be run, if not we must rectify the errors generated and we must continue further.