

# Vinay Kumar Camarushi

Data Engineer

📞 +1 (913)-227-9343 ✉️ vinay.kumar.cvk2999@gmail.com 📍 Overland Park, KS

🐙 [GitHub](#) 📄 [Portfolio](#) 🔗 [LinkedIn](#)

## Summary

Experienced Data Engineer with a track record of 1.5+ years in the industry, specializing in Python 3 and its supporting libraries, including Pandas, NumPy, Matplotlib, Keras, TensorFlow along with Java, and C. Known for adept problem-solving abilities and a strong commitment to maintaining industry-best coding practices. Skilled in collaborating effectively within multidisciplinary teams to achieve project objectives. Committed to continuously enhancing technical skills and knowledge in data engineering to drive innovation and deliver impactful results for the organization.

## Education

### University of Central Missouri

Master of Computer Science  
Aug 2022 – May 2024

### Sagi Rama Krishnam Raju

#### Engineering College

Bachelor of Technology  
Jul 2016 – Sep 2020

## Technical Skills

Python, Java, Machine Learning, Pandas, SQL, MongoDB, Microsoft Power BI, Tableau, Microsoft Azure, Qlik Sense, NoSQL, Git.

## Soft Skills

Collaboration, Problem-solving, Communication, Time management, Result-oriented

## Publications

International Journal of Advanced Science and Technology: Defect Detection in Printed Circuit Board (PCB) using Image Processing Vol. 29, No. 4s, (2020), pp. 1205-1210.

## Certifications

- Microsoft certified Data Analyst Associate (DA - 100)
- Microsoft certified Azure Fundamentals (AZ - 900)
- Microsoft certified Data Fundamentals (DP - 900)

## Work Experience

### TATA Consultancy Services

Data Engineer

Jan 2021 - Jul 2022

- Single handedly engineered python solutions to automate manual tasks, resulting in an 95% reduction in manual effort and earning recognition from both management and clients.
- Developed and monitored Power BI reports catering to Commercial and Performance metrics for the client, Diageo Great Britain Pvt. LTD.
- Crafted SQL Server Management Studio Views to enhance data comprehension in alignment with business requirements.
- Collaborated with cross-functional teams to understand data requirements and translate them into technical solutions.
- Optimized Tabular Model within Visual Studio and tuned database queries, data processing workflows, and infrastructure components to improve performance and efficiency.
- Deployed and managed data infrastructure in cloud and on-premises environments, ensuring reliability, availability, and security.
- Revamped errors within Azure Data Factory Pipelines, enhancing data flow efficiency and accuracy.
- Implemented data quality checks, validation mechanisms, and monitoring systems to ensure the accuracy, completeness, and consistency of data.
- Provided technical guidance, mentorship, and support to junior members of the team.

## Projects

### Online Grocery ecommerce Site

Aug 2023 - Dec 2023

- The Online E-Commerce Site project is an academic endeavor aimed at developing a fully functional e-commerce website using modern web development technologies.
- It utilizes Flask for backend development, MongoDB for database management, and HTML & CSS and JavaScript for the user interface.
- Additionally, the project utilizes AWS (Amazon Web Services) for hosting the application, ensuring scalability, reliability, and accessibility.

### Empirical Evaluation of Machine Learning Techniques for Text

#### Classification

Jan 2023 - May 2023

- This project assesses the performance of different machine learning algorithms in classifying text data into predefined categories.
- It aims to determine the effectiveness, efficiency, and scalability of methods like Naive Bayes, Support Vector Machines, Decision Trees, and Neural Networks in real-world text classification scenarios.
- The project aims to assess the effectiveness, efficiency, and scalability of these techniques in real-world text classification scenarios.

### Large Scale Matrix Multiplication using Hadoop MapReduce

Jan 2023 - May 2023

- Conduct large-scale matrix multiplication using Hadoop MapReduce to efficiently process vast datasets.
- Utilize Hadoop's distributed computing framework to parallelize matrix multiplication tasks across multiple nodes, enabling efficient computation and scalability for handling massive datasets.

### Airline Ticket Management System

Aug 2022 - Dec 2022

- A web-based application facilitating streamlined ticket booking, reservation, and cancellation processes for airlines, employing Flask, Python, MongoDB, HTML & CSS for efficient management.
- Comprehensive admin panel for airline staff to manage flight schedules, bookings, and generate reports.