Vinay Kumar Camarushi

Software Engineer, Overland Park, KS

vinaycvk01 @gmail.com • 913-227-9343

GitHub Portfolio LinkedIn

SUMMARY

- Experienced Software 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.
- Designed insightful dashboards using Power BI leveraging engaging data visualizations to convey insights and present data-driven narratives.
- Hands-on experience with data modeling concepts, including dimensional and relational modeling, such as Star Schema.
- Proficient in building predictive models using machine learning algorithms, including linear regression, logistic regression, decision trees, and random forests, to forecast trends and predict outcomes.
- Leveraged SQL to extract and analyze large datasets from diverse sources, uncovering key insights that supported data-driven business decisions.

EDUCATION

Master of Computer Science **Aug 2022 – May 2024**

University of Central Missouri, Lee's Summit, MO

B. Tech in Electronics & Communication Engineering **Jul 2016 – Sep 2020**

SRKR Engineering College, Andhra Pradesh, India

TECHNICAL SKILLS

- **Languages:** Python, SQL
- **Packages:** NumPy, Pandas, Matplotlib, SciPy, ggplot2, Scikit-learn, TensorFlow, Keras, Seaborn
- **Software/IDEs:** Visual Studio Code, PyCharm, Jupyter /Notebook, Google Collab
- **Visualization Tools:** Power BI, Advanced Excel (Pivot Tables, Lookups, Dates)
- **ETL & Cloud:** SSIS, Informatica, Azure (Data Lake, Databricks, Data Storage, Data Factory, Azure SQL Database, Azure Blob Storage)
- **Database:** MySQL, SQL Server
- **Operating System:** Windows, Linux, MacOS
- **Versioning:** Git

CERTIFICATIONS

- Microsoft Certified Data Engineer Associate (DP - 203)
- Microsoft Certified Data Analyst Associate (PL-300)
- Microsoft Certified Azure Fundamentals (AZ-900)
- Microsoft Certified Data Fundamentals (DP-900)

PROFESSIONAL EXPERIENCE

TATA Consultancy Services, India ***Jan 2021 – Jul 2022***

Software Engineer

- Single handedly automated the processing of excel files received from the market, which are used for data ingestion into the Data Lake. This included adding, deleting, updating, and removing duplicate records using Python libraries NumPy and Pandas, reducing manual effort by 95% 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.
- Deployed and managed data infrastructure in cloud and on-premises environments, ensuring reliability, availability, and security.
- Optimized Tabular Model within Visual Studio and tuned database queries, data processing workflows, and infrastructure components to improve performance and efficiency.
- 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.

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.

ACADEMIC PROJECTS

Online Grocery ecommerce Site ***Aug 2023 – Dec2023***

- 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 ***Jan2023 – 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.
- It utilizes Flask for backend development, MongoDB for database management, and HTML & CSS and JavaScript for the user interface.