

Ramakrishna Reddy Vennam

Data Scientist

ramakrishna.r@careernb.com | +1 (656) 208-4865 | FL, USA. | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Professional Summary

- Results-driven Data Scientist with almost **4 years** of experience in extracting, transforming, and analyzing complex datasets to deliver actionable business insights.
- Adept at designing and implementing scalable data models, advanced statistical methods, and customized reporting frameworks to support strategic decision-making.
- Experienced in applying a range of statistical techniques such as regression, classification, clustering, sentiment analysis, and ensemble methods across structured and unstructured data sources.
- Proven track record of enhancing process efficiency and supporting cross-functional teams through data-driven solutions and clear visual reporting.

Education

Masters in Computer Science, University of South Florida, FL, USA.
2025

May

Work Experience

Charles Schwab, FL | Data Scientist.

Jan 2025 -

Current

- Applied advanced **Natural Language Processing** techniques and anomaly detection methods to analyze transaction narratives and customer behaviour, significantly improving the accuracy of fraud and risk segmentation.
- Employed **SQL, Python, and Pandas** to perform large-scale data cleansing, transformation, and validation across datasets exceeding **3 million records**, leveraging Hadoop and parallel processing to ensure high data quality and performance.
- Engineered and optimized predictive features, implementing machine learning models such as **K-Nearest Neighbours (KNN)** and **Gradient Boosting**, which improved credit risk classification accuracy by **25%**.
- Conducted comprehensive **Exploratory Data Analysis (EDA)** using **Matplotlib and Microsoft Excel** to uncover patterns and correlations in financial datasets, supporting data-driven investment decisions.
- Deployed machine learning models via **Flask-based REST APIs**, utilizing Flask SQLAlchemy for seamless database interaction and integrating AWS services including IAM and RDS for secure and scalable deployment.
- Designed interactive financial dashboards in **Power BI**, delivering real-time performance tracking and executive reporting that enhanced operational efficiency by **30%**.
- Integrated **LLaMA-based large language models** into financial analysis pipelines to enable advanced inference, contextual analysis, and intelligent retrieval of domain-specific insights.

Accenture, India | Data Scientist.

May 2020 – Jul

2023

- Engineered custom **PyTorch Data Loaders** to handle complex data formats, enhancing ingestion workflows and reducing pre-processing time by **40%**.
- Built, trained, and evaluated machine learning models using **Scikit-learn**, **accelerating experimentation cycles** and shortening development timelines.
- Developed and deployed advanced classification models—such as **Decision Trees, Random Forests, and SVM**—to support targeted marketing efforts, resulting in a **10% increase in average customer lifetime value**.
- Processed and structured large-scale unstructured text data using key NLP techniques, including tokenization, lemmatization, and stop-word filtering.
- Optimized **SQL queries** and resource allocation strategies to improve data processing efficiency, achieving reduction in overall execution time.
- Designed and delivered interactive Tableau dashboards to visualize critical health metrics in real time, supporting leadership with actionable insights and improving strategic decision-making by **25%**.
- Implemented **scalable MLOps pipelines** leveraging FastAPI, Docker, and CI/CD integrations to ensure reliable deployment of machine learning models into production environments.
- Developed resilient data pipelines in **Azure using Data Factory and Data Lake**, enabling robust batch and real-time processing with distributed frameworks like Hadoop and Apache Spark.

Skills

- Programming Languages:** SQL, Python, R, Scala, C++
- Database:** MySQL, Mongo DB, PostgreSQL, NoSQL
- Tools:** R Studio, Jupyter Notebook, Hadoop, Advanced Excel-Analytic Solver Platform, Docker, R-shiny.
- Python Libraries & Data Analysis:** Pandas, Numpy, Scikit-learn, Py-Torch, Tensor-flow, Keras, Matplotlib, Seaborn, Django, Flask, Flask SQLAlchemy.
- Machine Learning Skills:** Deep Learning, Decision Trees (CART), Ridge Lasso Regression, Res Net, Naïve Bayes, KNN, SVM, Boosting Random Forest.
- Cloud Platform:** Amazon Web Services (IAM, EC2, RDS), GCP (Big Query), MS Azure (Azure data factory, Azure Data Lake)
- Methodologies:** Waterfall, SDLC, SCRUM, Agile.
- Software Packages:** Microsoft Office (Excel, Word, PowerPoint), SharePoint, Google Suite.
- Data Skills:** Data Manipulation, Data Visualization, Data Analysis

- **BI Tools:** Tableau, Power BI
- **Soft Skills:** Analytical Thinking, Curiosity & Continuous Learning, Problem-Solving Ability, Adaptability, Time Management

Projects

Financial Insight Engine using LLMs (LLaMA) Charles Schwab	Jan 2025
– Present	
<ul style="list-style-type: none">• Built an AI-powered NLP solution leveraging LLaMA-based large language models to extract and summarize insights from unstructured financial documents in real-time.• Applied prompt engineering and integrated with Flask APIs and AWS RDS, embedding outputs directly into Power BI dashboards for dynamic business intelligence.• Achieved a 40% reduction in analysis time, accelerating decision-making and improving financial reporting workflows.• Collaborated with analysts to refine domain-specific prompts and ensure compliance with financial terminology.• Skills/Keywords: LLMs, NLP, Prompt Engineering, Flask, AWS, RDS, Power BI, Real-Time Inference, Document Intelligence	

Customer Segmentation for Marketing Optimization Accenture	May 2020 – Jul 2023
<ul style="list-style-type: none">• Developed machine learning models (SVM, Decision Trees) for behavioral customer segmentation, improving campaign conversion and targeting precision.• Created end-to-end pipelines using SQL, Pandas, and Tableau for visual reporting; improved customer LTV by 10%.• Overcame data heterogeneity issues with structured preprocessing and enhanced cluster accuracy from 70% to 80%.• Skills/Keywords: Customer Segmentation, Clustering Models, Behavioral Analytics, SQL, Pandas, Tableau, ML Pipelines	

MLOps Pipeline for Production Deployment Accenture	May 2020 – Jul 2023
<ul style="list-style-type: none">• Designed and deployed an end-to-end MLOps framework using FastAPI, Docker, and GitHub Actions for CI/CD automation of ML models.• Integrated with Azure Data Factory and Azure Data Lake to manage and orchestrate data workflows across environments.• Reduced ML model deployment time by 60%, from several weeks to under 2 weeks, enhancing reliability and release velocity.• Skills/Keywords: MLOps, CI/CD, FastAPI, Docker, GitHub Actions, Azure Data Factory, Data Lake, Model Monitoring, Containerization	

Certifications

- **Artificial Intelligence Foundations: Machine Learning**
- **Microsoft Certified: Azure Fundamentals**
- **Data Science and Machine Learning Bootcamp with Python**
- **Workshop in Machine Learning – Pantech E-Learning**