

PARVATH REDDY

Data Scientist

Bengaluru, KA | +91-7675043309 | parvathreddy50@gmail.com | [linkedin.com/in/parvath-reddy](https://www.linkedin.com/in/parvath-reddy) | github.com/parvath-reddy

Summary

Data Scientist skilled in machine learning, deep learning, and AI development with internship experience in analytics. Hands-on expertise in building production-ready ML systems, RAG-based applications, and financial risk models. Strong foundation in Python, SQL, and advanced analytics to deliver data-driven business solutions.

Skills

Programming: Python, SQL, R, JavaScript

ML/DL & AI: Scikit-learn, XGBoost, TensorFlow, PyTorch, OpenAI API, RAG, NLP

Data Science: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Jupyter, Streamlit

Databases & Analytics: MySQL, PostgreSQL, Statistical Modeling, A/B Testing

Experience

Data Science Intern

August 2025 – September 2025

CodSoft (Remote)

Python, Scikit-learn, Random Forest, SMOTE, Regression Analysis

- Developed machine learning models for fraud detection and movie rating prediction using Python, Scikit-learn, and ensemble methods on 289,000+ transactions and 5,000+ films, achieving 95%+ precision, 95%+ recall, and 0.8 RMSE through advanced feature engineering, SMOTE oversampling, hyperparameter tuning, and cross-validation techniques that improved sales forecast accuracy by 85%.

Data Analyst Intern

February 2025 – April 2025

Mileto Foods (Remote)

Excel, Power BI, SQL, Data Analytics, DAX

- Analyzed 50,000+ rows of multi-state sales data using Excel, SQL, and Power BI to track regional KPIs and customer analytics across 5+ states and 3 product lines, designing DAX-powered dashboards and automated Excel templates with VLOOKUP and pivot tables that improved demand forecasting accuracy by 15%, increased quarterly revenue visibility by 12%, and reduced weekly reporting time by 35% across 20+ distribution points.

Projects

RAG AI Teaching Assistant | Python, OpenAI Whisper, BGE-M3, Ollama, Streamlit, FFmpeg

August 2025 [Live Demo](#)

- Engineered RAG-based video Q&A system with 6-stage pipeline achieving 85-90% query relevance accuracy and sub-3-second response time using Python, OpenAI Whisper, BGE-M3 embeddings, and Ollama LLM.
- Built production-ready Streamlit web application with semantic search capabilities handling multiple video formats, implementing cosine similarity algorithms and integrating local Llama 3.2 model serving 1,000+ educational chunks.
- Constructed automated data processing pipeline utilizing FFmpeg for audio extraction, OpenAI Whisper for speech-to-text conversion, reducing manual content indexing time by 95% and enabling instant access to video lecture information.

Fraud Detection System | Python, Scikit-learn, Streamlit, Plotly, Ensemble Methods

June 2025 [Live Demo](#)

- Implemented end-to-end fraud detection system using Python and Scikit-learn, training ensemble models on 1M+ financial transactions achieving 99.7% accuracy, 98.5% precision, and 97.8% recall metrics.
- Executed advanced feature engineering pipeline extracting 11 behavioral patterns from transaction data, reducing model training time by 60% while improving AUC-ROC score from 0.92 to 0.992.
- Deployed interactive web application using Streamlit and Plotly enabling real-time fraud detection with intelligent alerts and risk assessment dashboards serving 1,000+ predictions daily with 99.9% uptime SLA.

Credit Risk Prediction | Python, Scikit-learn, Extra Trees Classifier, Streamlit, Plotly

April 2025 [Live Demo](#)

- Architected end-to-end ML pipeline for credit risk assessment using Python, Scikit-learn, and Extra Trees Classifier, achieving 78.5% accuracy through feature engineering of 8 key financial indicators from 1,000+ credit records.
- Established production-ready Streamlit application with real-time prediction capabilities, interactive Plotly visualizations, and intelligent risk recommendations processing 500+ monthly assessments with 99.9% uptime.
- Optimized credit evaluation time by 60% through mechanized risk scoring system with confidence intervals, enabling data-driven lending decisions and improving loan approval consistency by 40%.

Certifications & Publications

Oracle Data Science Professional Certification

September 2025

Oracle Cloud Infrastructure - Data Science Specialization

Research Publication: YOLOv8 Pothole Detection System

April 2025

Computer Vision & Deep Learning

[SSRN](#)

Education

Bachelor of Technology - Computer Science Engineering

Expected: 2026

Vellore Institute of Technology, Bhopal

CGPA: 7.7/10.0