Pranav Murali

Machine Learning Engineer

mpranav4816@gmail.com +1 (312) 714-6556 Chicago, IL LinkedIn Portfolio GitHub

Summary

Machine Learning Engineer with 4.5+ years of experience designing and deploying scalable ML systems, data pipelines, and AI-driven solutions in production environments. Expertise in Python, PyTorch, NLP, and Generative AI, including RAG, LLM fine-tuning, and multi-agent systems. Proven track record of improving model performance, optimizing data workflows, and delivering high-impact solutions using AWS, MLOps, and real-time data processing. Strong background in end-to-end ML lifecycle, from feature engineering to deployment and monitoring, with a focus on scalability, reliability, and business impact.

Professional Experience

Comcast Jun 2025 – Present | Chicago, IL

Machine Learning Engineer

- Architect and deploy predictive maintenance models using Python and PyTorch to forecast broadband traffic patterns, proactively reducing network outages and enhancing service reliability at scale.
- Design and optimize high-throughput ETL pipelines using Apache Spark (PySpark) and Apache Airflow, processing 10M+ daily network log records and improving data pipeline efficiency by 35% for downstream analytics in AWS S3.
- Build and productionize a RAG system leveraging LangChain, Hugging Face Transformers, and Pinecone, significantly accelerating technical support resolution and improving knowledge retrieval accuracy.
- Lead end-to-end deployment of machine learning models using Docker and AWS SageMaker, integrating MLflow for experiment tracking and implementing robust data drift detection, achieving a 25% improvement in model stability in production environments.
- Develop interactive, real-time dashboards using Power BI and Plotly to monitor network health and model performance, improving decision-making speed by 30% for enterprise stakeholders.

CloudSphere IT Labs

Data Scientist Aug 2021 – Jul 2023 | Bengaluru, India

- Developed and optimized classification models using Python and Scikit-learn to predict user churn and transaction trends for FinTech and E-Commerce clients.
- Engineered complex behavioral features and designed optimized data pipelines using SQL and PostgreSQL, reducing query latency by 30% and improving data accessibility for analytics workflows.
- Built NLP-based text classification and sentiment analysis models using Hugging Face Transformers (BERT), automating support ticket routing and improving classification accuracy by 20%.
- Designed and deployed end-to-end machine learning pipelines on AWS, containerizing models with Docker and deploying on EC2, reducing deployment time by 40% and enabling scalable inference.
- Created interactive dashboards using Tableau and Matplotlib to present model insights and performance metrics, accelerating stakeholder decision-making across international enterprise clients.

Data Engineer Jan 2020 – Aug 2021 | Bengaluru, India

- Built and maintained robust ETL pipelines using Python and SQL to extract behavioral logs from MySQL and PostgreSQL databases, enabling seamless data transfer into centralized AWS S3 storage for downstream analytics.
- Implemented Apache Kafka-based streaming pipelines to capture and process real-time application events, error logs, and transactional metadata, improving data ingestion throughput by 40% without impacting primary web server performance.
- Orchestrated end-to-end data workflows using Apache Airflow, ensuring high system availability and reliable data delivery for client-facing reporting systems.
- Standardized data environments and automated testing workflows using Docker, reducing environment inconsistencies by 30% and improving deployment reliability across distributed engineering teams.
- Developed interactive dashboards using Power BI and Matplotlib to monitor database performance and pipeline health, enabling faster issue detection and data-driven decision-making during peak traffic periods.

Projects

Scalable ML Pipeline for Network Traffic Forecasting & Anomaly Detection

- Built a scalable machine learning system using Python and PyTorch to forecast broadband traffic patterns and proactively detect network anomalies, improving service reliability and reducing downtime risks.
- Developed high-throughput ETL pipelines with PySpark and Airflow, and integrated a RAG-based support assistant using LangChain and Pinecone to accelerate issue resolution and automate technical knowledge retrieval for operations teams.

GenAI-Driven DevSecOps Automation with Multi-Agent Virtual Teams

- Engineered an end-to-end DevSecOps automation pipeline integrating CI/CD, SAST, DAST, and IaC using Waterfall methodology, enabling structured delivery of secure, scalable infrastructure and automated deployment workflows.
- Developed and orchestrated a GenAI-powered multi-agent system using AutoGen and LangChain to simulate Scrum roles, driving adaptive task execution, improving collaboration, and enabling iterative enhancements across DevSecOps workflows.

AI-Based Secure Authentication System for Physically Challenged Users Using Morse Code

- Designed and implemented a contactless authentication system using computer vision and machine learning to detect facial and eye movements, enabling secure password entry for physically challenged users without traditional input devices.
- Engineered a Morse code-based input mechanism to enhance security against shoulder surfing and thermal attacks, ensuring accessible, reliable, and user-friendly authentication for individuals with motor disabilities.

Education

Master's in Computer Science Aug 2023 – May 2025 | Chicago, IL

Illinois Institute of Technology

Bachelor's in Information Science Aug 2018 – Jul 2022 | Bengaluru, India

Jain University

Skills

Programming & Querying: Python, SQL, R
Machine Learning & AI: Scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM, Deep Learning (CNN, RNN, LSTM), Natural Language Processing (NLP), Feature Engineering, Model Optimization, Model Explainability (SHAP, LIME)
Generative AI & LLMs: Hugging Face Transformers (BERT, GPT, T5), Prompt Engineering, RAG, OpenAI APIs, LLM Fine-Tuning (LoRA, QLoRA), LLM Evaluation & Optimization, LangChain, LlamaIndex
Data Engineering & Processing: Apache Spark, PySpark, Apache Kafka, Apache Airflow, ETL Pipelines
Cloud & Deployment: AWS (EC2, S3, SageMaker, Lambda), Azure (Azure ML, Azure Data Factory, Azure Blob Storage), GCP (BigQuery, Vertex AI, Cloud Storage), Docker, Kubernetes, Model Deployment, ML Pipelines
MLOps & Monitoring: MLflow, Feature Stores, Model Monitoring, Data Drift Detection, CI/CD (GitHub Actions)
APIs & Backend: FastAPI, Flask, REST APIs
Databases & Vector Stores: PostgreSQL, MongoDB, FAISS, Pinecone
Visualization: Tableau, Power BI, Matplotlib, Plotly

Certificates

- Oracle Cloud Infrastructure 2025 Certified Developer Professional
- Oracle Cloud Infrastructure 2025 Certified AI Foundations Associate
- AWS Fundamentals: Going Cloud-Native