

SUMMARY

AI/ML Engineer with **3 years of experience** applying machine learning to real-world challenges in healthcare, finance, and telecom. Designed systems that improved customer interaction efficiency, reduced processing delays, and supported accurate decision-making. Contributed to automation efforts that minimized manual work and enhanced compliance in high-volume data environments. Played a key role in streamlining workflows for document handling, fraud detection, and customer service operations. Known for delivering practical solutions that align with organizational goals and improve team productivity across multiple departments.

TECHNICAL SKILLS

- Languages & Scripting:** Python, Java, C++, JavaScript, Bash
- Machine Learning & AI:** Regression, Classification, Clustering, SVM, XGBoost, Feature Engineering, Model Evaluation, Hyperparameter Tuning, scikit-learn
- Deep Learning:** Neural Networks, CNN, LSTM, TensorFlow, Keras, PyTorch, Reinforcement Learning
- NLP & LLMs:** OpenAI API, Hugging Face, LangChain, Prompt Engineering, NER, Text Summarization, Chatbot Development, RAG
- Data Engineering:** SQL, NoSQL, ETL Pipelines, Data Cleaning, Data Modeling, Query Optimization
- Databases:** MySQL, PostgreSQL, MongoDB, DynamoDB, Cassandra
- Cloud & DevOps:** AWS (EC2, S3, Lambda), Docker, FastAPI, REST APIs, GitHub Actions, CI/CD, Terraform (basic)
- Tools & Libraries:** Pandas, NumPy, Matplotlib, Seaborn, Scipy, Git, JSON, Regex
- Simulation & Systems:** Environment Simulation, Reward Engineering, LIDAR Object Detection
- Security & Compliance:** HIPAA, Secure Identity Protocols, Blockchain (Smart Chain)

PROFESSIONAL EXPERIENCE

Intel Corporation, USA

Jan 2025 – Current

AI/ML Engineer

- Developed NLP pipeline using LangChain and Hugging Face, improving chatbot accuracy by 32% and reducing customer service ticket resolution time by 40% across multiple internal departments.
- Trained reinforcement learning agents with PyTorch and LIDAR simulation data, improving object detection efficiency and reducing task completion time in warehouse simulations by 28%.
- Built REST APIs with FastAPI for patient data workflows, ensuring HIPAA compliance and reducing manual processing time for clinical information by 45%.
- Automated model deployment using GitHub Actions, Docker, and EC2, reducing CI/CD pipeline execution time by 50% and increasing build reliability across multiple releases.
- Integrated OpenAI-based RAG architecture to support legal document search, improving retrieval relevance by 35% across compliance teams and reducing research hours by 20%.
- Monitored production drift using model evaluation metrics, retraining LLMs as needed and maintaining consistent accuracy levels above 90% for high-volume NLP use cases.

Infosys, India

Jun 2021 – Jul 2023

AI/ML Engineer

- Designed fraud detection models using XGBoost and SVM, improving classification precision by 24% and supporting daily processing of over 500,000 insurance claims.
- Built ETL pipelines using Python and SQL, reducing data ingestion latency by 40% and increasing refresh frequency for real-time dashboards across business units.
- Trained CNN and LSTM models for OCR document classification, raising accuracy from 81% to 93% in loan application processing workflows.
- Created APIs with FastAPI for ML model endpoints, improving runtime access for production applications and reducing failure rates during deployments by 33%.
- Engineered MongoDB schema for text analytics, supporting faster query execution and reducing storage redundancy for document-based datasets by 18%.

Infosys, India | ML Engineer | Internship

Jan 2021 – June 2021

- Processed telecom data using Pandas and NumPy, improving data quality by 30% for churn analysis through cleaning, standardization, and anomaly removal.
- Built clustering models using K-Means and DBSCAN, identifying key user segments and improving marketing conversion rates by 17% through targeted campaign strategies.
- Created chatbot prototypes with OpenAI API, reducing manual customer interaction volume by 22% in internal service systems.
- Deployed TensorFlow models in Docker containers, enabling isolated testing environments and reducing model rollback issues during early development by 40%.
- Transformed SQL and JSON datasets for NLP use cases, decreasing preprocessing time by 36% for text summarization and classification tasks.

PROJECTS

LLM-Powered Medical Report Processor

Apr 2025

- Designed a full NLP system to extract and analyze medical report text using fine-tuned MedLLaMA with LoRA, enhancing diagnostic clarity.
- Deployed interactive Streamlit-FastAPI frontend supporting real-time report uploads, LLM selection, and response generation.

AI Agent for Database Querying with RAG

Feb 2025

- Built a RAG-based AI system that translates user questions into optimized SQL/NoSQL queries across diverse databases.
- Integrated logic for query optimization, reducing retrieval time and eliminating manual data interpretation efforts.

Lunar Lander Simulation Using Gymnasium

Jun 2024

- Developed a reinforcement learning agent using DQN and Q-learning to master lunar landing control in simulation.
- Fine-tuned hyperparameters and reward logic to improve landing consistency across diverse terrain conditions.

EDUCATION

Master of Science in Computer Science

Aug 2023 – May 2025

Illinois Institute of Technology, Chicago, IL

Bachelor of Engineering in Electronics & Communication Engineering

Aug 2018 – Jul 2022

Anna University, Chennai, India