

# PRAJWAL KUMAR

AI / Machine Learning Engineer — Data Science — Generative AI

Pittsburgh, PA — +1 (878) 600-3424 — [prajwalk@andrew.cmu.edu](mailto:prajwalk@andrew.cmu.edu) — [LinkedIn](#) — [GitHub](#) — [Website](#)

## SUMMARY

AI/ML Engineer with hands-on experience building, evaluating, and deploying production-grade ML systems and data-driven applications. Strong background in supervised learning, deep learning, and generative AI, with end-to-end ownership across data processing, model training, experimentation, and cloud-native deployment using Python, PyTorch, AWS, and GCP.

## EDUCATION

Carnegie Mellon University	Pittsburgh, PA
Master of Science in Artificial Intelligence Engineering, GPA: 3.6	December 2025
Coursework: Data Engineering, Deep Learning, Machine Learning in Production, Generative AI, Trustworthy AI	
Maharshi Dayanand University	Rohtak, India
Bachelor of Technology in Computer Science and Engineering	June 2024

## EXPERIENCE

Infinite Computer Solutions	Irving, TX
AI/ML Summer Intern	May 2025 - August 2025
• Designed and deployed an agentic AI system to automatically extract, structure, and analyze telecom plan data across major U.S. carriers, reducing manual testing and enabling faster, data-driven market comparisons for enterprise stakeholders.	
• Built a production-grade retrieval-augmented generation (RAG) system backed by a Neo4j knowledge graph, enabling semantic search over large-scale telecom plan data and improving discoverability for product and analytics teams.	
• Containerized and deployed the end-to-end system using Docker and Kubernetes (Helm), delivering a scalable Streamlit-based analytics dashboard with data export to support faster, self-serve business analysis.	
Qriocity	Chennai, India
Machine Learning Developer Intern	January 2024 - February 2024
• Developed a healthcare-focused ML pipeline using TensorFlow to model medical ontologies and predict prescriptions from structured clinical data, achieving 99% accuracy and demonstrating real-world applicability in clinical decision support.	
• Built a multimodal ML system using TensorFlow-based DNNs and KNN classifiers to model emotional VAD signals from text features (TF-IDF), enabling reliable emotion classification for mental health applications.	
Zummit Infolabs	Bengaluru, India
Data Science Intern	November 2022 - March 2023
• Built and evaluated an NLP-based text classification pipeline in PyTorch to detect toxic content in social media comments, using feature preprocessing and supervised learning to support automated content moderation.	

## ACADEMIC RESEARCH & PROJECTS

End-to-End Movie Recommender with Kafka, Kubernetes, and A/B Testing - Carnegie Mellon ( <a href="#">GitHub</a> )	
• Built a scalable KNN-based movie recommender ( $\text{NDCG}@10 = 0.9983$ , 65K+ QPS) with Dockerized microservices, Kubernetes, CI/CD (Jenkins), and CRON-based retraining; integrated A/B testing and real-time monitoring with Grafana.	
Emotion-Aware Multimodal AI Companion - Carnegie Mellon University ( <a href="#">GitHub</a> )	
• Built a multimodal AI companion with speech-to-text (Google API), CNN-based SER (71.2%), and Transformer models; improved LLM empathy by 30% via psychoanalysis modules and deployed a Streamlit app for real-time emotion feedback.	
Cloud-Native Soccer Player Valuation Platform - Carnegie Mellon University ( <a href="#">Github</a> )	
• Designed an end-to-end data system on GCP to predict player value; built PySpark ETL pipeline from GCS to Cloud SQL (PostgreSQL) and trained regression models (SparkML, PyTorch) using Optuna/CrossValidator, deployed via Cloud Run.	
Neural Robot Dynamics: Linear-Time Inference and Long-Horizon Stability via Selective SSMs ( <a href="#">Report</a> )	
• Replaced Transformer dynamics with Mamba (SSM) and Jamba hybrid (SSM + attention) to enable linear-time inference and state-of-the-art sim-to-sim fidelity, limiting RL policy degradation to 1.37% over long-horizon rollouts.	
Advancing Image Security through Deep Learning and Cryptography in Healthcare Industry - IEEE ( <a href="#">Paper</a> )	
• Conference paper on a deep learning-based cryptographic framework for securing medical images in IoMT environments.	

## SKILLS

Programming & Data:	Python, SQL, Pandas, NumPy, PySpark, PostgreSQL, Neo4j, AWS, GCP, Streamlit
ML & GenAI:	PyTorch, TensorFlow, Scikit-learn, LangChain, OpenAI GPT, LLMs, RAG, Vector Databases, Playwright
Systems & MLOps:	Git, Docker, Kubernetes, Kafka, MLflow, Jenkins, CI/CD, Linux, REST APIs, Prometheus, Grafana