

PRAJWAL KUMAR

Pittsburgh, PA — +1 (412) 728-2613 — prajwalk@andrew.cmu.edu — [LinkedIn](#) — [GitHub](#) — [Website](#)

EDUCATION

Carnegie Mellon University **Pittsburgh, United States**

Master of Science in Artificial Intelligence Engineering - Information Security, GPA: 3.6/4.0 December 2025

Coursework: Adversarial Machine Learning, Data Engineering and Pipelines, Intro to Deep Learning, Machine Learning in Production (S3D SCS), Generative AI and Large Language Model (LLM).

Leadership: Internal Vice President of CMU AI Club, Vice President of GOINI student group

Maharshi Dayanand University **Rohtak, India**

Bachelor of Technology in Computer Science and Engineering, GPA: 8.27/10 June 2024

Coursework: Big Data Analytics, Data Science, Machine Learning, Neural Networks, Data Structures and Algorithms.

Leadership: Intel® Student Ambassador for oneAPI, Google Developer Student Club Lead '23 (Chapter MDU)

EXPERIENCE

Qriocity **Chennai, India**

Machine Learning Developer Intern **January 2024 - February 2024**

- Engineered medicine prescription ontology using TensorFlow and RDFLib with graph-based knowledge representation.
- Developed cataract classification system (LensCraft AI) using CNN to track surgical tools with attention mechanisms.
- Implemented email phishing detection model using ensemble of Logistic Regression classifiers with TF-IDF vectorization for feature extraction, achieving 96% accuracy with focus on minimizing false negatives.
- Created Speech-to-Text interview system with SpeechRecognition and natural language processing pipeline, incorporating probabilistic model for recruitment likelihood prediction with Streamlit deployment.
- Designed emotion detection system using transfer learning in TensorFlow with K-NN for psychological indicators.

ScriptEdge Pvt. Ltd. **Akola, India**

Machine Learning Intern **July 2022 - August 2023**

- Implemented generative multimodal system using Stable Diffusion with ControlNet for conditional image generation, integrating OpenAI Whisper for speech-to-text and custom prompt engineering for domain-specific outputs.
- Developed resume parser using Regex and SpaCy with custom entity extraction for CV information retrieval.
- Engineered multilingual text classification system using bidirectional LSTM networks with attention mechanisms.
- Built RASA conversational AI with 30+ intents, custom NER pipeline, and database integration for enterprise deployment.
- Designed computer vision system for QR code artistic integration using conditional GAN architecture with ControlNet, implementing perceptual loss functions to maintain QR functionality while maximizing aesthetic quality.

Zummit Infolabs **Bengaluru, India**

Senior Data Science Intern **November 2022 - March 2023**

- Led development of driver drowsiness detection using CNN-LSTM with Keras and OpenCV (94% accuracy, 2800 images).
- Designed Squeeze-Excitation Network for image quality classification with channel attention mechanisms (92% accuracy).
- Developed toxic content classifier using PyTorch with BERT embeddings and adversarial training (96% accuracy).
- Engineered NLP chatbot with hierarchical attention networks, contextual memory, and entity recognition.

ACADEMIC RESEARCH & PROJECTS

Cardiovascular Disease Prediction Using Genetic Algorithm - IJARESM, November 2022 ([Research Paper](#))

- Engineered a cost-effective heart disease prediction system using ANNs and binary classification algorithms on critical health variables, implementing genetic algorithms for model optimization with enhanced diagnostic reliability.

Advancing Image Security through Deep Learning and Cryptography in Healthcare - July 2024 IEEE ([Link](#))

- Developed secure EHR protection by integrating deep learning with cryptographic techniques (98.7% success rate), utilizing chaotic systems and optimized neural networks for medical image encryption, while addressing IoMT security challenges.

Data Pipelining and Machine learning using SparkML on FIFA 22 Dataset - Carnegie Mellon ([GitHub](#))

- Implemented scalable data pipelines using Apache Spark for the FIFA 22 dataset (73,975 player records, 100+ features), performing preprocessing and feature engineering to build predictive models for player market value (MAE: 0.016).

Multilingual RAG Pipeline for Research Paper Recommendation using Mistral AI - Carnegie Mellon ([Github](#))

- Developed a multilingual RAG pipeline with ChromaDB for semantic search across research papers. Integrated Mistral LLM with LangChain, compared embedding models and deployed a Streamlit application for paper recommendations.

SKILLS

Languages/Packages: Python, SQL, Pandas, NumPy, NLTK, Keras, Scikit-learn, TensorFlow, PyTorch, PySpark, OpenCV

Tools: Git, AWS, Azure, Google Cloud (Speech-to-text, BigQuery, Dialogflow, Vertex AI, Kubernetes), Flowise AI, Mistral AI, LangChain, Linux, PostgreSQL, Docker, Neo4j, Kafka, ChromaDB, Streamlit, Jenkins