Neeta Kannu
neetatechy@gmail.com | +1-719-233-4623 | LinkedIn | GitHub | Portfolio

CAREER SUMMARY

AI/ML Engineer with 8 years in software engineering and 2.5 years in artificial neural networks (ANNs), deep learning, large language models (LLMs), speech recognition, reinforcement learning (RL), and natural language processing (NLP); proficient in Python, PyTorch, TensorFlow, Docker, Kubernetes, Google Cloud (GCP), and Microsoft Azure; designing and scaling production-grade, cloud-native AI systems that deliver up to 30% faster inference, 15–20% higher model accuracy, and measurable business impact across NLP, computer vision, and RL domains.

EDUCATION

- **University of Colorado at Colorado Springs, USA – M.S. in Computer Science** **Aug 2023 – Dec 2025**
  Relevant Coursework: Computation Theory, Data Structures and Algorithms, Artificial Intelligence, RL, ANN, Software Product Management

- **University of Mumbai, India – B.S. in Computer Science** **Jun 2012 – May 2015**
  Relevant Coursework: Operating Systems, Linux, C, Java, Assembly Language, Software Engineering

PROFESSIONAL EXPERIENCE

Wireless Network Assistant, UCCS OIT, Colorado Springs, CO, USA **Jun 2024 – Aug 2025**
- Automated **network provisioning using Python**, reducing manual configuration time by 40% and streamlining deployments
- Monitored **Windows and Linux systems** with **SolarWinds** and **ClearPass**, cutting **downtime incidents by 25**% and enhancing system reliability.

Research Assistant, UCCS College of Engineering and Applied Science, CO, USA **Aug 2023 – May 2024**
- Processed and visualized large-scale genomic datasets, streamlining ML workflows and delivering a **30% faster pipeline** for accelerated research.
- Containerized end-to-end **ML systems** using **Docker**, **CUDA**, and **GCP**, enabling scalable, repeatable experimentation in distributed environments.
- Refactored existing code into modular Python scripts with embedded feature engineering, improving maintainability and reducing onboarding time.

Senior Software Engineer, Accolite Digital, Bangalore, India **Aug 2021 – Jan 2023**
- Engineered **C# backend and Selenium test suites**, enabling scalable deployments, boosting release **speed 30%** and **cutting defects 25%**.
- Orchestrated **Jenkins and Azure DevOps CI/CD**, reducing manual deployment effort 50% and ensuring reliable, production-ready releases.
- Optimized **monitoring and deployment scripts**, improving system visibility, reducing downtime, and supporting reproducible data workflows.

Software Consultant, Capgemini, Bangalore, India **Nov 2019 – Aug 2021**
- Developed and tested **REST APIs with Java, Selenium**, and **Postman**, accelerating QA turnaround 40% and improving software reliability.
- Containerized applications using Docker and Kubernetes, enhancing horizontal scalability and minimizing environment drift in production.
- Migrated large-scale systems to cloud architectures, integrating **CI/CD** with **Git and Jenkins** to enable continuous, reliable delivery.

Systems Engineer, Tata Consultancy Services, Bangalore, India **Jun 2015 – Nov 2019**
- Delivered **high-availability backend services in Java**, **C++, and Shell**, ensuring 24/7 enterprise uptime and system reliability.
- Tuned **complex SQL queries**, improving **database throughput by 25%** and accelerating application responsiveness.
- Automated delivery pipelines with **PowerShell and Jenkins**, cutting **release time by 40%** and enhancing deployment operations.
- **Mentored three junior** engineers in automation and requirements analysis, accelerating **onboarding by 30%** and strengthening team capability.
- Resolved **concurrency and threading issues** in distributed systems, enhancing stability and ensuring SLA compliance.

PROJECTS

Prompt Engineering for LLMs [Project Link] **Jan 2025 – Present**
- **Enhanced LLM accuracy** by 35% on **GSM8K, SVAMP, and MAWPS** using **prompt engineering, reducing hallucinations.**
- **Increased retrieval relevance by 22%** by integrating **semantic search** in RAG pipelines for faster, more accurate enterprise responses.
- Led **fine-tuning of LLaMA/ChatGPT-3.5** on **GPU-optimized GCP**, reducing inference costs **by 20%** while maintaining performance.

Accent Identification Model (ANN) [Project Link] **May 2023 – Jan 2024**
- Constructed a **multilingual accent classifier (Speech Accent Archive)**, boosting **ASR accuracy by 20%** and expanding global coverage.
- Engineered **audio-text fusion preprocessing**, cutting misclassifications and improving model generalization across diverse accents.
- Scaled **ML experiments on GCP**, implementing reproducible workflows with **Weights & Biases (W&B)** to accelerate model deployment.
- Owned implementation of **PyTorch, TensorFlow, Hugging Face** to deliver high-performance, production-ready speech models.

Heart Disease Prediction [Project Link] **Jan 2024 – May 2024**
- Produced **SVM** and **Logistic Regression** models on Cleveland dataset, achieving **85% accuracy** for early cardiac risk, enabling faster care.
- Applied **PCA** for dimensionality reduction, accelerating model inference by 30% and improving deployment efficiency.
- Deployed models on **GCP** using **Docker** containers, **facilitating faster iterations and reliable deployments.**
- Leveraged **Python, Scikit-learn, PCA, Docker, and GCP** to build production-ready, high-impact healthcare AI solutions.

3D Chromosome Reconstruction (RL) [Project Link] **Jan 2024 – May 2024**
- Implemented **Python, TensorFlow, CUDA, Docker, and HPC** to build high-performance, reproducible genomics AI workflows.
- Designed **3D chromosome folding** simulation using **Reinforcement Learning** on Hi-C data, improving spatial prediction.
- Reduced **training cycles 40%** via **CUDA parallelization** and scalable container orchestration on HPC clusters, speeding experimentation.

TECHNICAL SKILLS

- **Programming Languages:** Python, Java, C/C++, SQL, Bash, PowerShell, HTML, JSON

- **ML / AI Frameworks**: scikit-learn, XGBoost, LightGBM, CatBoost, PyTorch, TensorFlow, Keras, Hugging Face Transformers, spaCy, NLTK, OpenCV, ANN, CNN, RNN, LSTM, Reinforcement Learning, Multimodal AI, Conversational AI, Prompt Engineering, Feature Engineering, Hyperparameter Tuning, Time Series Forecasting, Model Deployment

- **Cloud & DevOps:** GCP, Azure, AWS, Docker, Kubernetes, Jenkins, CI/CD, MLflow, Airflow, Kafka, Spark, Hadoop, Data Pipelines, BigQuery

- **Tools & Certifications:** Jupyter, Google Colab, Git, GitHub, VS Code, PyCharm, Selenium, Maven, TestNG, .NET, ASP.NET | Python for Data Science and AI (IBM – Coursera), Overview of Data Visualization (Coursera)