

Saikiranmansa Sunnam

Machine Learning Engineer

✉ ssaikiranmansa2023@gmail.com ☎ 7164150744 📍 Buffalo 🌐 LinkedIn 📁 Portfolio 🤖 Hugging Face

SUMMARY

Machine Learning Engineer with expertise in NLP, computer vision, and anomaly detection, leveraging PyTorch, TensorFlow, and Scikit-learn to build scalable ML pipelines and deploy APIs via FastAPI/Docker. Proven success in fraud detection (Isolation Forests), Transformer fine-tuning, and RAG-based QA systems, with strong MLOps and cloud (AWS) experience. Currently pursuing an M.Sc. in Computer Science at the University at Buffalo, blending academic rigor with hands-on ML deployment.

PROFESSIONAL EXPERIENCE

Machine Learning Intern, Rivach LLP

05/2022 – 03/2023 | India

- **Designed an end-to-end machine learning pipelines** across multiple domains such as **computer vision, regression, and anomaly detection**, utilizing **TensorFlow, PyTorch**, and **Scikit-learn** to deploy models with **98% accuracy** and **R² scores > 0.95** for regression.
- **Constructed robust data preprocessing systems** for **20,000+ labeled images** and **100,000+ structured records**, integrating **SMOTE**, **label encoding**, and **feature scaling** to reduce model training errors by **30%+** and eliminate data skew.
- **Engineered scalable CNN-based deep learning architectures** for image classification, improving predictive accuracy by **22%**, while reducing model inference time by **18%** using architectural pruning techniques.
- **Implemented ensemble learning models**, including **Random Forest** and **Gradient Boosting**, resulting in a **15% increase in F1-score** and **12% gain in recall**, significantly outperforming traditional classifiers.
- **Tuned model hyperparameters** through cross-validation and grid search strategies (e.g., `n_estimators`, `max_depth`, `learning_rate`), reducing **false positive rates by 25%+** and improving **model variance by 20%**.
- **Built and defined unsupervised anomaly detection pipelines** using **Isolation Forest**, **One-Class SVM**, and **Local Outlier Factor**, enhancing fraud detection precision by **19%** on highly imbalanced datasets.
- **Conducted deep regression modeling** on public health datasets using **PyTorch**, optimizing for **MSE = 0.034** and achieving an **explained variance of 0.96**, resulting in accurate life expectancy predictions.

TECHNICAL PROJECTS

Domain-Specific QA System Using DeepSeek and RAG

- **Built a production-grade question-answering system** leveraging **DeepSeek** and **Retrieval-Augmented Generation (RAG)**, combining **FAISS**-based dense vector retrieval and a language model to deliver **30%+ improvement in response accuracy**.
- **Processed and indexed over 1 million domain-specific documents** using **FAISS** and **Elasticsearch**, enabling efficient, low-latency (<250ms) query retrieval at scale.
- **Enhanced semantic search quality** by integrating **Sentence-BERT** embeddings, boosting **BLEU** and **ROUGE** scores by **15–20%**, and improving relevance in context-sensitive queries.
- **Benchmarked QA performance** using **BLEU**, **ROUGE**, and **Exact Match**, achieving **30%+ lift over baseline retrieval and generation models**, validating end-to-end model effectiveness.
- **Launched a scalable QA pipeline on AWS** using **FastAPI** and **Docker**, reducing document retrieval time by 50%.

Advanced Anomaly Detection and Text Classification Using Deep Learning

- **Engineered an anomaly detection framework** using three autoencoder variants to detect patterns in a time-series dataset of **5,315 records**, achieving a maximum **R² of 0.9916** and detecting **25–74 anomalies** depending on the model.
- **Developed a Transformer-based text classifier** using **PyTorch** and fine-tuned it on the **AG News dataset**, increasing classification accuracy from **90.08% to 90.53%** via **L2 regularization** and **dropout**.
- **Preprocessed and visualized 10,000+ tokens**, leveraging **tokenization**, **normalization**, and **data visualization** with **Matplotlib** to identify trends and feature importance for classification.
- **Optimized training using advanced regularization techniques**, fine-tuning dropout rates and learning rates, which resulted in a **0.45% improvement in accuracy** and **12% reduction in overfitting**.
- **Conducted comprehensive model evaluation** using **R²**, **precision**, **recall**, **F1-score**, **confusion matrix**, and improved **ROC AUC from 0.82 to 0.91**, validating model performance and generalization.

EDUCATION

State University of New York at Buffalo, Master of Science in Computer Science

12/2024 | Buffalo, NY

Courses: Machine learning, Deep learning, Computer Vision & Image Processing, Operating Systems, Algorithms Analysis and Design, Data Intensive Computing, Computer Security, Data Mining and Query Language, Software Engineering

SKILLS

Core ML & AI: Generative AI, LLMs (LLaMA, GPT, BERT), RAG, NLP, Computer Vision, Anomaly Detection

Frameworks & Libraries: PyTorch, TensorFlow, Scikit-learn, Hugging Face, Keras

Programming: Python, SQL, R, Java, C

Data & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Plotly

DevOps & Cloud: Docker, FastAPI, REST APIs, AWS (EC2, S3, Lambda, SageMaker, DynamoDB, CloudWatch), CI/CD

Big Data & Databases: Hadoop, Spark, MySQL, Oracle Database

Tools & Collaboration: Git, Jupyter Notebooks, Data Warehousing, Testing, Code Reviews, Documentation

PUBLICATIONS

Granite classification using machine learning and edge computing

Published in **F1000Research**: [Link](#) 🔗