KISHORE SAMPATH

(857) 891-4581 sampath.ki@northeastern.edu linkedin.com/in/s-kishore/ github.com/kishore-s-15

Education

Northeastern University **Sep. 2023 – Expected May 2025**
Master of Science in Artificial Intelligence *Boston, MA*
Coursework: Machine Learning Operations, Explainable AI, Unsupervised Learning and Data Mining, Bias/Ethics Implications in AI, Algorithms.
Gradute Teaching Assistantship: Artificial Intelligence

SRM Institute of Science and Technology **Aug. 2016 – May 2020**
Bachelor of Technology in Mechanical Engineering — (Top 5%) *Chennai, India*
First rank in the university examinations in 2019 and awarded a performance-based merit scholarship.
Related Courses: Neural Networks and Fuzzy Systems, Probability and Statistics, Calculus of Variations and Non-linear programming.

Technical Skills

Languages: Python, C, C++, JavaScript, Java **Tools & Technologies**: Linux, Docker, LLM, Generative AI, MLOps
ML Frameworks: PyTorch, Tensorflow, Keras, Scikit-Learn, OpenCV, NLTK, Transformers, LangChain, OpenAI Gym
Dev Frameworks & DBs : Django, FastAPI, ReactJS, SQL, MongoDB, ChromaDB, FAISS, Streamlit

Professional Experience

Khoury College of Computer Sciences, Northeastern University **August 2024 - Present**
Graduate Student Researcher *Boston, MA*
- Building a Disease-Informed Ethical & Trustworthy Report Generation LLM model using Multimodal EHR data.

GITAA Private Limited (IIT - Madras Incubated Startup) **May 2021 – July 2023**
Full-Stack Developer and Team Lead *Chennai, India*
- Developed a data-driven software application to mimic the induction furnace process to improve it's operational efficiency using the client-server architecture. (Stack: ReactTS, Django, Postgres, Pandas, Redis, Tensorflow, Scikit-Learn, Docker)
- Implemented the Insights module for performing exploratory data analysis, visualizing trends, and inspecting the descriptive statistics of the induction furnace's operational data.
- Built the Analytics module for training ML & ANN models on the operational data of the induction furnace using a custom user-defined data preprocessing pipeline to predict the optimal process parameters under given constraints.
- Automated the Cost Sheet generation process for the sales team and improved their productivity by reducing the time taken to generate a cost sheet by 75%.

Center for Industrial Consultancy and Sponsored Research, IIT - Madras **October 2022 - March 2023**
Machine Learning Engineer - Outsourced *Chennai, India*
- Developed an NLP-based Customer Research Analysis tool using the client-server architecture for a Fortune 500 client and deployed it in their HPC cluster. (Stack: ReactTS, Django, MySQL, Redis, Celery, NumPy, PyTorch, Docker)
- Engineered an Extractive Summarization module using Sentence BERT and the Lex Rank algorithm to effectively condense information from customer interview transcripts.
- Built a Context-aware Sentence Tagging module to tag relevant sentences from transcripts based on the user-defined keywords using Zero-Shot Text classification with BART-Large-MLNI model and reduced the processing time by 80%.

Projects & Publications

Navigating Intersectional Bias: Enhancing Fairness in Multimodal Clinical Predictions
- Developed a cutting-edge multimodal fair classifier for clinical decision-making tasks using Electronic Health Records by integrating a custom object detection pipeline and GPT-2 for image modality encoding, alongside a novel method for encoding structured data using MedBERT.
- Analyzed model performance variations across diverse intersectional demographic groups and mitigated biases using innovative post-processing ensembling techniques.
- The research work is currently under review for AAAI 2025 main conference.

Kaggle - Autism Prediction Challenge | (Link) **February 2022 - March 2022**
- Developed a high-performing two-level stacking ensemble classifier combining XGBoost, LightGBM, CatBoost, and Random Forest classifiers as base estimators and a final Logistic Regression Model.
- Trained the model using stratified k-fold cross-validation and optimized it's hyperparameters using Optuna to achieve a top 5% ranking in the competition with a test set AUC-ROC score of 0.943.

Honors and Contributions

- Contributed to the Deepchecks (PR links) and Hugging Face Transformers (PR links) open source libraries.
- Awarded the "Rising Star - 2021" for exceptional performance at GITAA Private Limited.
- Top 4% worldwide contributor on Stackoverflow in 2021, impacting over 15000 people (Link).