

# Jaya Raj Srivathsav Adari

☎ (201)398-5551 | ✉ ajsrivathsav3502@gmail.com | 🌐 Jaya-Raj-Srivathsav-Adari | 🎧 Srivathsav-max | 📧 Srivathsav.me

## Education

New Jersey Institute of Technology

Master in Computer Science

Sep 2023 – Dec 2024

3.71 GPA

Koneru Lakshmaiah University

Honors Bachelor's Degree in Computer Science

Aug 2019 – Mar 2023

9.08 GPA

## Technical Skills

**Languages:** Python, C++, JavaScript, GoLang, TypeScript, SQL

**Technologies & Frameworks:** :Apache Spark,Apache Kafka, Apache Airflow, Jenkins ,AWS Glue,Apache Hadoop, HDFS, MapReduce, PySpark, Pandas, Numpy, PyTorch, Keras, TensorFlow, Flask, FastAPI, Django, Redis, GraphQL.

**Cloud & DevOps:** Microsoft Azure, AWS, GCP, Git, JIRA

**Data Visualization Tools:** Tableau, Grafana

**Databases:** MySQL, MongoDB, Postgres, Azure Data Lake, Chroma, Hive

## Work Experience

Data Scientist

Mar 2025 – Present

WatchRx

Remote

- Built **speech-to-text pipeline** processing **1000+ hours** of healthcare calls monthly using VAD, speaker diarization, and Whisper STT, achieving **95% accuracy** for timestamped transcripts
- Shipped NLP summarization systems, **reducing manual documentation by 80%** and automating review for 500+ daily calls
- Implemented **HIPAA-compliant** data framework with role-based access controls and audit logging
- Improved transcription quality **by 25%** through confidence thresholds, language auto-detection, and intelligent re-decode queuing
- Established MLOps with MLflow tracking, drift monitoring, and CI/CD pipelines, **reducing deployment time by 60%**

Tech: Python, PyTorch, OpenAI Whisper, pyannote.audio, VAD, librosa, spaCy, scikit-learn, Pandas, NumPy, SQL, Airflow, Docker, Git

Data Scientist (AIOps / MLOps)

Sep 2024 – Dec 2024

Verizon

Jersey City, NJ

- Built anomaly detection on **5TB+ Kubernetes logs** daily using Isolation Forest and other custom machine learning approaches, achieved **70% accuracy**, reduced detection time by 45%
- Created **self-healing recommender** using RAG over 1000+ runbooks, **reduced manual intervention 80%**, cut resolution time from 2 hours to 15 minutes
- Orchestrated **ETL and ML** pipelines in Airflow/Spark, processing **10M+ daily events** with streaming and batch scoring
- Deployed Kafka ingestion with schema registry, **improved throughput 300%** and achieved <100ms end-to-end latency
- Implemented Grafana dashboards tracking detailed model metrics with automated SLO alerting

Tech: Python, PySpark, Airflow, Kafka, scikit-learn, Vector DB, Docker, Kubernetes, Grafana, MLflow

ML Engineer — Computer Vision

Jun 2021 – Jun 2023

Samsung

Hyderabad, India

- Optimized **ResNet** models across ONNX/TensorFlow/TFLite, **cut inference latency 40%** while maintaining 95% accuracy over 4000+ tests
- Implemented operator-level optimizations (fusion, graph optimization), achieving **30% memory reduction** and 25% faster execution
- Built benchmarking suite profiling small models, guided quantization/pruning for **2x speedup** on edge devices
- Automated **CI/CD** for model conversion with golden output validation, **reduced regressions 90%**

Tech: TensorFlow, TFLite, ONNX, Python, NumPy, CUDA, Android toolchain, Git

## Projects

SNOMED CT & ICD-10-CM Comorbidity Coding LLM

Jun 2025 – Present

[Python, PyTorch, Transformers, scikit-learn, spaCy, SQL, Spark, Airflow, MLflow, Vector DB, LoRA/QLoRA, UMLS]

Sponsored Project

- Trained a **multi-label LLM/classifier** to predict **high-comorbidity ICD-10-CM codes** from notes and primary diagnoses.
- Built a **terminology pipeline**: text normalization, **SNOMED CT** mapping, and ICD-10-CM crosswalks with synonym handling via **UMLS**.
- Modeled **comorbidity signals** using co-occurrence stats (PMI/lift) and graph embeddings to generate candidate code sets.
- Fine-tuned with **LoRA**, enforced valid code decoding; added **active learning** and evaluated with F1, top-k recall, and exact match.

Clinical Coding & Documentation Automation Suite

Apr 2025 – Jun 2025

[Python, PyTorch, Transformers, scikit-learn, spaCy, SQL, Airflow, MLflow, RAG, Vector DB, FastAPI, GCP]

Sponsored Project

- Built an **OASIS virtual nurse** (LLM + rules + validators) for structured capture; **reduced ops cost 40%**; added PII redaction and prompt evals.
- Shipped **predictive ICD-10-CM coding** on notes + claims with n-grams, embeddings, and comorbidity lift; improved coder throughput and accuracy.
- Delivered **CPT G0181 auto-completion** via structured extraction and validations; **reduced physician time 65%**.
- Logged inputs/labels and added **monitoring**: data quality, drift checks, precision/recall dashboards; tracked runs in **MLflow**.

Moxium — AI-Powered Educational Platform

Mar 2025 – Present

Technical Founder

StartUp Project

- Built an **adaptive learning engine** with student/skill embeddings and DKT/Transformer to serve next-best lesson.
- Designed a **recommendation system** for certification-competency fit using content tagging, SBERT embeddings, and learning-to-rank.
- Implemented **knowledge-gap detection** with real-time features; auto-adjusted paths; tracked Top-k Recall, MAP, and time-to-mastery.
- Set up **MLOps**: MLflow experiments, data quality checks, drift monitoring, and offline/online evaluation.

Tech: Python, Pandas, scikit-learn, PyTorch, Transformers, SQL/PostgreSQL, dbt, Airflow, MLflow.

## Certifications

- AWS Certified Developer – Associate
- TensorFlow Developer Certification