Jayantha Nanduri

Boston, MA | (+1) 617-992-4852 | jayantha.nanduri@gmail.com | linkedin/in/jayantha-nanduri | github/jaynanduri

Education

Northeastern University, Teaching Assistant - CS5800 Algorithms — Boston, MA
Master of Science in Computer Science; GPA: 3.85/4.0 — *Jan. 2023 – May 2025*

- **Relevant Courses** : NLP, Large Scale Data Processing, Machine Learning, Design Patterns, Web Development.

Experience

Information Technology, Advanced Applications Co-op — Aug 2024 – Present
Boehringer Ingelheim — *Athens, GA*

- Reduced analysis time from **days-to-minutes** to summarize long unstructured documents by engineering a chatbot application using **streamlit & Langchain**.
- Addressed **context window limitations** by developing **map-reduce workflow**, splitting documents into small batches, summarizing each batch, and then combining them into a final summary.
- Eliminated **LLM hallucinations** with a **RAG vector** store to verify output against source documents.
- Enhanced operational efficiency by **30%** by automating file tracking at manufacturing site through a full-stack application built with **Next.js, PostgreSQL, and NFC tags**.
- Accelerated development speeds by **20%** with streamlined workflows, **containerizing** application code and database into separate services, leveraging **Docker** for improved modularity.
- Integrated **PowerBI** dashboard into the application, providing stakeholders with real-time insights into operational efficiency and key performance metrics.

Machine Learning Engineer — Jan 2020 – Dec 2022
Jocata Financial Advisory & Technology — *Hyderabad, India*

- Optimized **time-to-credit** from **2-5 days to under 6 hours** by developing **AI-powered scoring model**, evaluating SME's key factors like sales growth, segment stability and the dependency on buyers and suppliers.
- Reduced API response latency by **30%** by integrating a **Redis caching** layer in **Django**, consistently handling 10,000+ daily requests.
- Developed scalable **ETL pipeline** in **Apache Spark** to ingest tax filing data from government APIs, and store the processed results in **PostgreSQL** database for efficient data management.
- Improved text extraction by **40%**, accurately detecting different languages in images, filtering out relevant text by training a custom **Faster R-CNN** model using **PyTorch, MLFlow**.
- Decreased **Storage Footprint by 40%** through Efficient Image Preprocessing using **AWS Lambda and AWS Batch** to perform real-time image resizing, compression, and format conversion.
- Automated hyper-parameter tuning with **MLflow, HyperOpt**, optimized performance through experiment tracking, and reduced development times by **10%**.

Projects

Distributed Graph Analysis | *PageRank, Python, Java, PySpark, Map Reduce, AWS*| Github

- Implemented **PageRank** in **MapReduce & Spark** to analyze **1M** nodes dense web graph on **EMR** cluster with I/O handling in S3 buckets.
- Employed lineage information to enhance debugging, performance optimization, and result validation, improving **accuracy of outputs by 20%**.

Caption Craft | *Full Stack, GenAI, React.js, Node.js, MongoDB, MaterialUI, Express, GraphQL*| Github

- Optimized the noise scheduling process in the pre-trained **stable diffusion** model, enabling the generation of 10,000 high-quality images with an **FID score below 10**.
- Developed APIs using **GraphQL** to optimize data fetching, enabling flexible, efficient queries and real-time updates, improving performance and scalability in user connections, recommendations, and content search features.

Technical Skills

Languages: Python, Java, SQL, Scala, Golang, TypeScript.
Frameworks: NextJS, FastAPI, NumPy, Pandas, PyTorch, Scikit-learn, LangChain.
Databases: Kafka, MongoDB, PostgresSQL, Spark, Redis.
Cloud platforms: AWS, Azure, Vercel.
Cloud and Software tools: PowerBI, Docker, Openshift, Jenkins, Git, Jira.