ANANDHARAJU DURAI RAJU

Ph.D. Candidate, Computing Science, Simon Fraser University, Burnaby
+1 (604) 518-3116 Vancouver, Canada anandharaju@ieee.org

linkedin.com/in/anandharaju
anandharaju.github.io
Google Scholar

SUMMARY

ML Research Engineer specializing in building end-to-end LLM, GNN and CNN systems, with 7+ years of professional experience in leading Retail and Telecom bigdata projects. Passionate about solving real-world problems with a focus on AI efficiency and acceleration. A cross-functional collaborator who is highly adaptable to emerging AI trends

SKILLS

- **Languages:** Python, Apache Spark, Java (JSP, JSF, Springboot), Javascript, PySpark, C++
- **ML/DL Frameworks/Libraries:** PyTorch, Tensorflow, Keras, HuggingFace, Spark MLlib, Unsloth AI
- **Parallel Computing:** Distributed Data Parallel, Python multiprocessing, Pandas Dask, MPI (C++), HF Accelerate
- **Databases:** Postgres, HBase, Oracle, DB2
- **Tools/Packages:** Docker, Ollama, vLLM, SGLang, Jinja, NLP NLTK, SpaCy, Postman, Kafka, Grafana, REST, Git
- **AI Agent Frameworks:** SmolAgents, LangChain, LangGraph, LlamaIndex, Azure AI, AWS Bedrock, GCP

RESEARCH EXPERIENCE

Research Assistant, Simon Fraser University | **Canada** | Prof. Ke Wang ***Jan 2019 – Present***

- Surpassed benchmarks in automated table semantics understanding task by 6% (F1-score) via **LLM-guided Graph Attention Networks** trained on embeddings generated by Mistral, Mixtral, Llama and Qwen LLMs
- Performed zero-shot, few-shot and fine-tuning based benchmarking of SOTA LLMs on Table understanding
- Trained **novel** Transformers/xLSTMS hybrids on long sequences with 96x less GPU memory using CNN extractors
- Optimized GPU memory usage (22x less), time (50% less), and carbon footprint (7x less) **without performance loss** in training malware classification CNNs on ultra-long sequences (>250M timesteps), achieved via a **novel retroactive pruning algorithm** and **memory-efficient backpropagation** ***– Published in ACM CIKM 2024*** [PDF]
- Surpassed state-of-the-art performance by 2-9% TPR @ 0.1% FPR using a **novel boosting algorithm** designed for efficiently learning sequential representations with minimal false detections ***– Published in IJCNN 2022*** [PDF]
- Optimized LLM/DL GPU usage via gradient checkpointing, offloading, quantization and LoRA/QLoRA

ACADEMIC EXPERIENCE

- Ranked 1st on Question-Answering task by improving BERT via semantic sentence similarity-based **input pruning**
- Certified in **Agentic AI** (HuggingFace) - built Google Gemini-based agents tackling GAIA dataset with observability
- Trained Generative AI models (Variational AutoEncoder) with 97.6% accuracy for **credit card fraud detection**
- Improved Llama with human feedback-based **reinforcement learning** (RLHF) using Google Cloud Vertex AI
- Built and pre-trained (GPT, Llama) **from scratch**, fine-tuned **multi-modal LLM** for speech and visual QA tasks
- Enhanced NDCG metric for **learning-to-rank** article popularity achieving >92% performance on unseen articles

INTERNSHIP EXPERIENCE

Research Intern, Huawei Canada | **Canada** | **Data Privacy & Protection Tech. Lab** ***Jan 2021 – Dec 2021***

- Designed, developed and deployed **end-to-end** CNN malware detector with <100ms detection latency via Docker
- Developed **efficient pipelines** with 80% reduced pre-processing time with Dask parallelization to extract, preprocess and load/store sequence data (pickle, JSON, protobuf) accelerating the ML training and deployment
- Developed a compact top-performing residual neural network-inspired malware detector with 97% accuracy
- Successfully published a **pioneering survey paper** on cross-architectural IoT malware threat hunting [PDF]

PROFESSIONAL EXPERIENCE

Technology Lead, Infosys Limited | **India** | **AI & Automation Services** ***Sep 2011 – Dec 2018***

- Led 14-member feature team (on-site + offshore) under agile-based software development delivering end-to-end real-time order events-processing big data solutions using Spark, Kafka and Java-based backend services
- Experienced in unit, integration and smoke testing, and owning production software with 24/7 gold SLA support
- Won client's **"AWARD OF EXCELLENCE"** in 2016 and 2017 for tackling high priority incidents and change requests
- Championed SSH/JSch-based automation that reduced 47 manual hours/month in monitoring InfoVista servers
- Experienced in achieving effective client discussions, code reviews and mentoring of junior team members

AWARDS AND MISCELLANEOUS ROLES

- Garnered **"GOLD MEDAL" (Top 1%)** at state level in my undergraduate studies from Anna University
- Played the role of **"STUDENT CHAIRMAN"** of Computer Science department for undergraduates
- Reviewer in Conferences and Journals – KDD, ICDM, ICDE, WSDM and IEEE Access