Aravinda Raman Jatavallabha

Raleigh, NC | [aravindaraman14@gmail.com](mailto:aravindaraman14@gmail.com) | (919) 327-0958 | [github.com/aravinda-1402](https://github.com/aravinda-1402/) | [linkedin.com/in/aravinda-jatavallabha](https://www.linkedin.com/in/aravinda-jatavallabha/)

# EDUCATION

**Master of Computer Science (Data Science Track) |** North Carolina State University, Raleigh, NC Aug 2023-May 2025 (expected) Courses – Data Science, Natural Language Processing, Neural Networks, Database Management Systems **CGPA: 4.0/4.0**

**B. Tech in Information Technology |** Manipal Institute of Technology, Manipal, India Jun 2019-Jul 2023 Minor: Big Data Analytics; Courses - Data Mining, Machine Learning, Pattern Recognition, Algorithms **CGPA: 8.64/10.0**

# TECHNICAL SKILLS

* **Programming Languages, Technologies & OS:** Python, SQL, Docker, AWS (S3, Sagemaker, Lambda), Power BI, Snowflake, Git, Linux
* **Libraries:** Pandas, NumPy, Matplotlib, Scikit-learn, Keras, PyTorch, TensorFlow, NLTK, PyG, SpaCy, SciPy, LangChain
* **Machine Learning:** Time Series Analysis, Classification, Regression, Convolutional Neural Networks (CNN), Natural Language Processing (NLP), Graph Neural Networks (GNN), Retrieval-Augmented Generation (RAG), Large Language Models (LLMs), A/B testing, Data Mining
* **Training & Certifications:** [Deep Learning (deeplearning.ai),](https://www.coursera.org/account/accomplishments/specialization/DFFQB96C6FJ2?utm_source=link&utm_medium=certificate&utm_content=cert_image&utm_campaign=sharing_cta&utm_product=s12n) [Machine Learning (Stanford Online),](https://www.coursera.org/account/accomplishments/verify/9AB6K4JPZSQ9?utm_source=link&utm_medium=certificate&utm_content=cert_image&utm_campaign=sharing_cta&utm_product=course) [AI Summer School](https://drive.google.com/file/d/1jPhjHO8WfoukW9IKUp8yKpRyoLKbnjFB/view)

# WORK EXPERIENCE

**Data Science Intern |** SmartProtect Public Safety Solutions, Wilmington, DE May 2024-Present

* Automated emergency data processing with **AWS Lambda** & **S3**, securing 5+ service agreements and boosting client satisfaction by 25%.
* Developed **Power BI dashboards** to track key performance metrics for 911 centers, using **sentiment analysis** on dispatcher reports to evaluate communication effectiveness, directly influencing operational strategies and enhancing efficiency by 15%.
* Implemented **LSTM** and Prophet models on **AWS SageMaker** for 911 call forecasting, conducted **A/B testing** to compare pipelines, and deployed the best-performing model to reduce scheduling errors by 20%.
* Designed and optimized staff scheduling algorithms based on call volume trends, improving shift planning, and aligning staff, with simulations suggesting an 18% reduction in overtime costs.

**Machine Learning Intern |** Defense Research and Development Organization, Bengaluru, India Jan 2023-Jun 2023

* Engineered a Temporal **Graph Neural Network** (**GNN**), leveraging continuous temporal data and node features to predict future user interactions on online platforms, increasing model accuracy by 2% over current benchmarks [[Paper].](https://drive.google.com/file/d/158F9V9XCoOoz8T-kh5VBvL4I0_ZgaJ4a/view?usp=sharing)
* Developed and integrated Incremental **BERT** (iBERT) with **Temporal GNN** to capture semantic drift and enhance real-time semantic understanding of evolving text data, reducing data processing time by 40%.
* Achieved a language model perplexity of 3.19 in masked language prediction, surpassing the state-of-the-art score by 6%, with this work published in Springer ICPR 2024 [[Paper].](https://rdcu.be/d2zF1)

**Data Science Intern |** Merkle Inc., Bengaluru, India May 2022-Jul 2022

* Led a team of 4 in developing **predictive models** (**XGBoost, LightGBM, LSTM**) for revenue optimization strategies, performing data transformations, and applying **SQL indexing techniques** on 10M+ data points, driving a 10% revenue increase.
* Processed 16M+ rows of Home Depot sales data using **PySpark**, improving query performance by 40% through advanced data handling techniques, generating actionable pricing insights.

**AI Research Assistant |** Manipal Institute of Technology, Manipal, India Jan 2021-May 2022

* Devised a Bone Age Assessment model by enhancing VGG-16, MobileNet, Inception V3, and XceptionNet architectures to determine skeletal vs. chronological age differences achieving state-of-the-art accuracy [[Paper].](https://arxiv.org/abs/2207.10169)
* Designed and deployed **LSTM** and Bidirectional LSTM models to predict channel link quality on IoT network traces, achieving 99.4% accuracy, and automated the **ETL pipeline** using **Apache Airflow**, **AWS S3**, and **SageMaker**, cutting manual tasks by 60% [[Paper].](https://ieeexplore.ieee.org/document/10767139)

# PROJECTS & PUBLICATIONS

* **Legal Query AI Assistant** [[Code]](https://github.com/aravinda-1402/Legal-Query-AI-Assistant) : Built a chatbot with **RAG** capabilities using **OpenAI GPT** and **Llama 2**, integrating **Pinecone Vector Database** and **HuggingFace** embeddings to automate document retrieval and text chunking via LangChain, enabling precise, context-aware legal query responses through a **Streamlit** interface.
* **Diabetes Prognosis using Machine Learning** [[Paper](https://ieeexplore.ieee.org/document/9985487) | [Code](https://github.com/aravinda-1402/Diabetes-Prediction)] : Presented at IEEE ICIRCA 2022 on a **Random Forest** model for diabetes classification, containerized using **Docker**, achieving 3.8% higher accuracy than previous models, benchmarking against K- Nearest Neighbors (**KNN**) and Artificial Neural Networks (**ANN**) to enhance reproducibility and scalability across platforms.
* **Wolf Parking Management System** [[Code](https://github.com/aravinda-1402/Wolf-Parking-Database-Management-System)] : Built a **SQL**-based parking management system using **BCNF** and **3NF** normalization, improving query efficiency by 35% with **indexing** and advanced reporting. Implemented SQL **triggers** and **ACID**-compliant transactions, enhancing performance by 11% and automating vehicle and permit management.
* **Customer Churn Prediction** [[Code](https://github.com/aravinda-1402/Customer-Churn-Prediction)] : Developed an end-to-end customer churn prediction pipeline using **Apache Airflow**, automating data processing, model training, and deployment with **AWS SageMaker**. Built a **Dockerized Flask web app** and **REST API** for real- time predictions using **logistic regression**, achieving 92% accuracy on the Telco Customer Churn dataset to support proactive retention strategies.