Aditya Kumar

+1 (770) 706-2214 | 2512.adityasingh@gmail.com | linkedin.com/in/aditya-kumar-gatech | github.com/aditya2512
sites.google.com/view/aditya-kumar-gatech

Education

Georgia Institute of Technology — Atlanta, GA
Master of Science in Computer Science (Machine Learning) GPA: 3.85 / 4.0 — *Aug 2024 - May 2025*

CMR Institute of Technology — Bangalore, India
Bachelor of Engineering in Information Science GPA: 8.76/10.0 — *Aug 2016 - May 2020*

Experience

Allstate India Private Limited — Bangalore, India
Software Engineer - Machine Learning — *Feb 2021 – Aug 2024*

- Designed and scaled ML pipelines using Python and Spark, deploying time series models (ARIMA, Holt-Winters) on Azure ML, optimizing storage forecasting and reducing cloud operational costs by 35% through automated resource scaling.
- Developed and optimized deep learning models for anomaly detection using BERT with CI/CD pipelines on Docker/Kubernetes, reducing Mean Time to Repair (MTTR) by 40% through real-time alerts and parallel processing of 50TB+ server logs.
- Implemented A/B testing frameworks for personalization model evaluation, enabling offline experiments and performance benchmarking for recommendation algorithms.
- Built and optimized large-scale distributed data pipelines with PySpark and Flink, improving query performance by 60% via schema optimization, efficient handling of structured/unstructured data, and distributed data processing.
- Developed predictive resource allocation algorithms using T5 and Gemini models, reducing resource scaling time by 24 hours, contributing to operational efficiency.
- Enhanced recommendation systems using Retrieval-Augmented Generation (RAG) techniques, improving chatbot response accuracy by 50% and reducing response time by 40%.

Research & Projects

LLM Security — Nov 2024 – Present

- Researching neural network security vulnerabilities, identifying attack vectors, and designing mitigation strategies for safeguarding ML models.
- Implementing reverse engineering techniques to extract weights/biases from LLMs, including Falcon 40B and GPT-4, using advanced cryptanalytic techniques in black-box settings.

LLM-Driven Multilingual Medical Diagnosis System — Sept 2024 - Dec 2024

- Designed a pipeline for multilingual medical diagnosis, processing Hindi inputs and delivering accurate diagnoses using ensemble LLMs (Mistral AI, Phi 3.5, Llama 3.2-1B), achieving 95% accuracy.
- Implemented model versioning and deployment pipelines using MLflow for tracking, evaluation, and scalable inference.

Multimodal Disease Diagnosis and Treatment — Sept 2024 - Dec 2024

- Developed a multimodal ML pipeline integrating structured healthcare data and echocardiographic videos (DICOM files) to predict cardiac conditions and Right Ventricular Ejection Fraction (RVEF).
- Implemented transformer-based models and Temporal CNNs for accurate disease classification, optimizing model inference speed and reliability.

Technical Skills

Programming: Python (Pandas, NumPy, Sklearn, Matplotlib), SQL (PostgreSQL, MySQL), PyTorch, TensorFlow, CUDA Python, R, Keras, Java, Scala, ETL.

Software: Tableau, Power BI, Git, Hadoop, Spark, Bash, Linux, Terraform, ServiceNow, Flink, Hadoop, Docker, Kubernetes, CI/CD, Jenkins, AWS (EC2, S3, Lambda, SageMaker, EMR), Unity.

ML & AI: LLMs (GPT-3.5, GPT-4, BERT, Llama 2, Falcon 40B, T5, Gemini), Recommendation Systems, NLP, GenAI, Personalization Algorithms, A/B Testing, RAG Pipelines.

Coursework: Big Data Analytics, Data and Visual Analytics, Machine Learning, Reinforcement Learning, NLP, Deep Learning.