GUNIK LUTHRA

+1(201) 492-7921 | gunikluthra@outlook.com | New York Metropolitan Area | LinkedIn | GitHub | Portfolio

EDUCATION

Stevens Institute of Technology | Hoboken NJ, United States
Master of Science, Computer Science: GPA 4.0/4.0 ***MAY 2025***
Relevant Coursework: Generative AI, Machine Learning, Deep Learning, Big Data, Data Mining

Vellore Institute of Technology | Vellore TN, India
Bachelor of Technology, Computer Science ***APRIL 2023***
Relevant Coursework: Database Management Systems, Data Structures and Algorithms

SKILLS

Languages: Python, Spark, C++, Java, JavaScript

AI: Generative AI, RAG, AI Agents, Prompt Engineering, LLMs, Langchain, LangGraph, CrewAI

Machine Learning: NumPy, Pandas, Scikit-learn, PyTorch, TensorFlow, NLP, Decision Trees, Logistic Regression, SVM, Predictive Modelling, Feature Engineering, Transformer models, NER, K-Means clustering, Spatial clustering

Programming and Development: Git, Kafka, AWS, GCP, Apache Airflow, Docker, Flask, Streamlit

Data Analytics and Visualization: SQL, Power BI, Matplotlib, Seaborn, MS Office Suite, MS Excel, RDBMS

PROJECTS

Fin AI - Financial Report Generator *(link)* ***Sept 2024 – Dec 2024***

- Created an AI agent using LangChain with chain-of-thought reasoning to automate 10-K financial data extraction
- Designed 10 agentic tools to transform financial data and generate insights, reducing report generation time by 85%
- Implemented Streamlit dashboards, boosting visualization and decision efficiency by 60%

Comparing NER and LLM for Analyzing VAERS DATA *(link)* ***Sept 2024 – Dec 2024***

- Engineered a preprocessing pipeline for VAERS text data, improving accuracy by 31% through abbreviation expansion
- Evaluated NER and LLM approaches on 15K+ datapoints, showing LLM's superior accuracy in symptom extraction
- Applied DTW, LCS, and Kendall Tau correlation to demonstrate LLM's symptom extraction is 2.6x more effective

Chubb (Capstone Project) - Real-Time Stock Data Analytics *(link)* ***Jan 2024 – May 2024***

- Developed a real-time data pipeline using Kafka, AWS EMR and Spark to process data from 10 tickers, implemented parallel processing for scalable data ingestion and created real-time visualizations using Streamlit
- Employed Spark Structured Streaming to compute KPIs maintaining a Data Quality Index (DQI) of 0.9
- Integrated event logs, metadata capture and ganglia for optimizing resource utilization by 17%

NYC Traffic Collision Analysis *(link)* ***Jan 2024 – May 2024***

- Incorporated data cleaning, feature engineering and ETL for analysis on 5M+ data tuples from MV collision data
- Leveraged spatial clustering to identify accident hotspots and calculate severity scores (0–1), and developed predictive models with MLlib, achieving 84% accuracy in risk forecasting
- Executed scale out on GCP Dataproc resulting in 73% drop in training time

ChatML - Chatbot for Machine Learning *(link)* ***Sept 2023 – Dec 2023***

- Collaborated with a cross-functional team to design an end-to-end chatbot using GPT-3.5 API, RAG, and vectorization techniques to query an inbuilt knowledge base of machine learning models
- Automated model execution based on user-defined parameters, reducing manual intervention by 30%
- Optimized hyperparameter tuning, reducing fine-tuning time by 40% and boosting accuracy by 25%

WORK EXPERIENCE

Stevens Institute of Technology | Hoboken NJ, United States
Graduate Assistant ***June 2024 – Present***

- Enhanced ID verification using advanced data analytics, enhancing reservation process efficiency
- Addressed technical inquiries with IT expertise and data-driven analysis, optimizing operational processes

HCL Technologies LTD | Noida UP, India
Software Engineering Intern ***Feb 2023 – Mar 2023***

- Designed crucial features for e-commerce website using C# and .NET, focusing on performance and maintainability
- Utilized Entity Framework for data access, enabling efficient and secure interaction with databases
- Created RESTful APIs using ASP.NET Core, and integrated them with front-end applications

PUBLICATIONS

- Comparison of ML Algorithms on Colour Quantization Techniques, 2022 IEEE International Conference, Bhopal, India, DOI: 10.1109/CCET56606.2022.10080406
- Road Accident Monitoring System and Dynamic Insurance Pricing Using Fog Computing, 2022 International Conference on I-SMAC, Dharan, Nepal, DOI: 10.1109/I-SMAC55078.2022.9987333