

EDUCATION

Carnegie Mellon University (CMU) | Master of Science in Artificial Intelligence and Innovation | CGPA: 3.93/4.0 May 2024

Relevant Coursework: Introduction to Machine Learning, Foundations of Computational Data Science, Coding Boot Camp, AI Engineering, ML for Large Datasets, Introduction to Question Answering, Advanced Natural Language Processing, On-device ML

Vellore Institute of Technology (VIT) | B.Tech in Computer Science Engineering | CGPA: 9.30/10.0 Jul 2022

WORK EXPERIENCE

BNY Mellon | Data Scientist (Capstone) ([Github](#)) ([Confluence](#)) Remote | Sep 2023 - May 2024

- Designed and developed a visualization platform using Neo4j, Cypher, and NeoDash, streamlining data summarization, analysis, and visualization of earnings calls. This platform reduced data gathering and report generation time, enhancing user productivity by updating a new file within 17 minutes on the AuraDB database and retrieving data within a few seconds.
- Leveraged Prompt Engineering and RAG (Retrieval Augmented Generation) techniques on Large Language Models to create a Topic Modeling tool and a QA chatbot, enabling efficient and insightful data retrieval from a Financial Knowledge Graph.

Introduction to Deep Learning | Teaching Assistant ([Website](#)) Pittsburgh | Aug 2023 - May 2024

- Crafted assignments, organized Kaggle competitions, and structured course content for 400+ students under Prof. Bhiksha Raj.
- Presented recitations, managed office hours, supported term projects, graded assignments, and quizzes, covering topics like MLP, CNN, RNN, Transformers, and Attention using PyTorch.

Juniper Networks | Software Engineering Intern California | June 2023 - Aug 2023

- Built a knowledge graph for supply chain network data, for understanding interdependencies between multiple components, SKUs, manufacturers, etc.
- Implemented Graph Neural Networks to create an AI-based solution for scenario analysis within supply chain, enabling predictive insights into node-based risk for fluctuations in supply chain data, thus enhancing supply chain resilience.

Bosch x CMU Corporate Startup Lab | Product Intern Remote | Jan 2023 - May 2023

- Identified potential customers, conducted competitive analysis, analyzed product-market fit, and mapped stakeholders.
- Facilitated over 15 interviews with multiple stakeholders and delivered weekly product updates to the client.

Tech Mahindra | Data Science Intern Remote | Jan 2021 - July 2021

- Analyzed potential fraudulent claims and processed comments to classify issues for warranty analytics.
- Developed a resume ranker system by curating and refining information from various files, streamlining the hiring process.

PROJECTS

ML Ops for Movie Recommendation System

- Implemented and evaluated surprise models and tracked provenance for a movie streaming service hosting 20k+ movies.
- Performed unit testing and data quality checks for input data.

On-Device Voice Enhanced QA system

- Collaborated on developing a voice-enhanced insights system on Nvidia Jetson Nano focusing on Voice Recognition and Q&A.
- Applied optimization techniques, including quantization and pruning, achieving a 35% reduction in model size while maintaining accuracy at a 4% reduction and improving latency for seamless on-device deployment.

Specific Entity Recognizer

- Scraped and parsed over 100 papers from multiple conferences, pre-processed and manually annotated 5 files for 7 specific entities, creating a comprehensive dataset. Trained a BERT, SciBERT, and RoBERTa model to accurately identify entities.

SKILLS

Languages and Tools: Python | Label-Studio | AWS | GCP | Databricks | ML Flow | Neo4j | Cypher | SQL

Libraries: NumPy | Pandas | Scikit-learn | NLTK | Spacy | PyTorch | Networkx | Pyvis | Surprise | Beautiful Soup | SciPDF | PySpark

Blogs: [Snorkel](#) | [Product-Owner](#)

PATENTS AND PUBLICATIONS

- Sign Language Interpreter.** Filing Number: 202141037366A. [Patent Link](#). [IRJET-Research](#).
- Missing Person Detector.** Filing Number: 202141037347A. [Patent Link](#).
- One-Place Solution to Mental Health and Workload Management.** Filing Number: 202141043594A. [Patent Link](#).
- Anomaly Detection using Unsupervised ML Algorithms:** [Springer Book Chapter Link](#).
- Phishing Website Detection using ML models:** [SSRN Conference Paper](#).