

EDUCATION

University of British Columbia

Master of Data Science – Computational Linguistics

GPA: 4.0 / 4.0

Vancouver, Canada | 2021 - 2022

Vellore Institute of Technology

B. Tech. – Computer Science and Engineering

CGPA: 9.31 | Top 5% of the Engineering class

Vellore, India | 2016 - 2020

SKILLS

Languages / Tools: Python, R | Java | Git, GitHub | Docker | MySQL | PostgreSQL | Spark | OpenSearch | Redis

Technical Skills: - AI | Deep Learning | ML | NLP | LLM | ML Pipelines | ETL Pipelines | CI/CD Pipelines | Language Modeling | NER | Supervised Learning | Embeddings | Information Extraction | Data Visualization | Regression

Frameworks: - Neural Networks | HuggingFace Transformer | LangChain | LangGraph | Agents and Tools | OpenAI API | GPT, BERT | PyTorch | Tensorflow | scikit-learn | PySpark | pandas | Kafka | FastAPI | AWS


EXPERIENCE

SS&C Technologies | Machine Learning Scientist

Aug 2022 - Present

- Identified and rectified anomalies in fund P&L using advanced data analysis techniques to ensure accuracy and reliability in financial reporting.
- Designed and implemented RAG (Retrieval-Augmented Generation) pipelines using the OpenAI API and LLMs like GPT, LLaMA, Falcon, and Vicuna to provide contextual explanations of price movements, boosting the team's analytical depth and aiding business teams in evaluating fund performance.
- Built Dockerized ML and ETL pipelines triggered by Kafka and FastAPI endpoints to fetch raw data, preprocess it, engineer features, and apply anomaly detection models.
- Utilized OpenSearch, ArgoCD, Rancher, and Grafana to deploy, monitor, and manage containerized applications in production, ensuring high availability and observability.
- Refined the trade verification process by implementing automation and using GitHub for version control and collaborative development, maintaining clean, auditable CI/CD workflows that reduced operational risk and enhanced organizational resilience.

SAP | NLP Engineer

ClearlyDefined Summarizer  - UBC Capstone | May - Jun 2022

- Built a **Text Summarizer** to summarize Licenses from [ClearlyDefined](#) using custom **TextRank** and **T5**.
- Used **Multi-task BERT** model and **Doc2Vec** to classify term labels; deployed model on [HuggingFace Spaces](#).

FindMind Analytics | Machine Learning Engineer

Internship | Dec 2019 - Jul 2020

- Led a team of junior interns; worked on [LegalMind](#), a research and collaboration platform to analyse legal terms of a judgment and summarize it, thus helping Lawyers save time; monitored end-to-end project.
- Built scrapers and pre-processed text. Trained Named Entity Recognition (NER) models to extract entities.

PUBLICATIONS

- [Impact of Sequence Length and Copying on Clause-Level Inflection](#) - Association for Comp. Linguistics

ACHIEVEMENTS

- Won Annual Software Development Hackathon by [SCOPE, VIT](#); led a team of four.
- Runner up in Database Design and Implementation Hackathon; led a team of two.