Nishant Gurunath

nishantgurunath@gmail.com | LinkedIn | Webpage

AREA OF WORK

Research and product development using Machine Learning and Natural Language Processing (NLP).

CAREER OVERVIEW

MOODY'S ANALYTICS Assc Dir - ML, 2020 - Present (4 yrs)

TEXAS INSTRUMENTS Design Engineer, 2016-2018 (2 yrs)

EDUCATION

CARNEGIE MELLON UNIVERSITY (CMU)
Master of Science in Electrical and Computer Engineering
Graduated December 2019
Pittsburgh, PA
GPA: 3.74 / 4

INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY (IIT BOMBAY)
Bachelor of Technology + Master of Technology in Electrical Engineering + Minor in Computer Science
Graduated June 2016
Mumbai, India
GPA: 8.57 / 10

SKILLS

PROGRAMMING

Python • C/C++ • MATLAB • SQL

SOFTWARE & PACKAGES

Pytorch • AWS • Scikit-Learn • LangChain • Databricks • Github • Elasticsearch • FastAPI • Docker • Pyspark • MLflow • RabbitMQ • Neo4j • Terraform • Flask • Streamlit

MACHINE LEARNING

Search • Recommendation • Reinforcement Learning • Transformers • Large Language Models • AI Agents • Embeddings • Knowledge Tags • Content Filtering • Entity Resolution • Entity Recognition • Information Extraction • Fine-Tuning and Transfer Learning

WORK EXPERIENCE

ASSC DIRECTOR - MACHINE LEARNING | Moody's | New York, NY
Feb 2020 – Present | Data Science (DS) and Engineering | Product Strategy

Know Your Customer (KYC) | KYC - Machine Learning | Ongoing

- Leading the design of a **RAG-based AI Agent** to automate the L1 screening process; AI agent reduces the false alerts by 80% and instrumental in driving client retention; Modeled the agent to auto-configure using client policy guidelines and reusable tools
- Lead the development of an **enrichment pipeline** to create entity risk profiles from news media; developed risk and duplicate **filters** to reduce analyst reviews by 40%; Setup **online testing and audit dashboards** to monitor model and pipeline operations
- Utilized and enhanced **data labeling** operation to setup **continuous training** workflow
- Guided the development of sync and async **model serving API** infrastructure on AWS
- Created keywords+**Gen-AI** models to apply **knowledge tags** to all news media data; steered a new product feature to provide news context for sanctioned entities
- Designed a **RAG**-based KYC **Investigation Assistant** chatbot that provides unified access to KYC data using natural language CWYD to facilitate and enhance the 2LoD
- Formed a DS group to innovate, collaborate, and instill industry best practices across the BU; kicked off two projects stemming from effective data and knowledge sharing
- Evaluated **entity resolution (ER)** tools - Senzing and Fincom, to consolidate KYC data
- Contributed to building big-data ML models in Spark for risk-oriented entity **alert systems**; implemented **experiment tracking** and **model governance** using MLflow

Task Force on Climate-related Financial Disclosures | TCFD 2021

- Created a machine learning pipeline for climate-related **information extraction**; extracted climate-related disclosures from company annual and sustainability reports
- Implemented a **pdf text-extraction** model using PyMuPDF and AWS Textract
- Designed an embedding-based **text retrieval** service using a Distilroberta model
- Fine-tuned an Electra cross-encoder model for **text ranking** and a Roberta model for **text classification** on climate-related relevance; obtained more than 80% F1-score

Commercial Real Estate (CRE) - News Pulse | Technical Lead

- Led a team of 5 engineers to deploy CRE - News Pulse on Moody's REIS website
- Designed an ML enrichment pipeline to provide real-time News around CRE markets
- Fine-tuned Transformer-based text classification models to **filter and enrich** (with CRE markets & sectors) relevant news articles; obtained greater than 85% F1-score
- Created **FastAPI** endpoints to provide real-time news for CRE markets

RESEARCH ASSISTANT | Separabl: Disentanglement in Speech | CMU
January 2019 - Dec 2019 | Prof. Alan Black | Prof. Richard Stern

- Proved that **multinode VAE** can be used to separate speech and music in audio for downstream speech applications such as speech synthesis using *found* data
- Experimentally determined the number of latent nodes required for **source separation**; showed that the same can be determined from the input data distribution
- Established improvement (by listening) in **speech synthesis** using separated speech

ACADEMIC AND PERSONAL PROJECTS

MULTILINGUAL CONVERSATION AGENT: LANGUAGE LEARNING ASSISTANT
Feb 2023 | Demo Application | Personal Project

- Created a bot that facilitates **open-ended conversations** to practice speaking in multiple languages; bridging the gap between language understanding and speaking
- Designed a Node.js application by integrating **Google's webspeech** (for speech recognition and speech synthesis) with **OpenAI text-davinci** (for chat completion)
- Deployed the web application on AWS ECS using Terraform cloud configuration