NISTHA MITRA

nistha.mitra28@gmail.com nmitra28 nmitra28

WORK EXPERIENCE

Software Engineer

Oracle Cloud Infrastructure, Speech AI

July 2022 – Present Seattle, Washington

- Spearheaded the development of cutting-edge cloud-based technologies within Oracle's Speech Artificial Intelligence Team.
- Proficiently constructed, tested, and meticulously maintained services utilizing **Java/J2EE** and web-based technologies such as **HTTP, JSON, WebSockets, and RESTful Web Services.**
- Built innovative **Health/Medical solutions for Oracle Cerner**, successfully **implementing Automated Speech Recognition models** for diverse use cases.
- Led the creation and management of a **cloud-native data collection tool**, essential for assembling audio and text data-sets to train ASR models.
- Took charge of product database integration and management, utilizing **PL/SQL** procedures and dynamic queries to optimize database interactions, ultimately **reducing both database calls and latency by a remarkable 32%**.
- Pioneered innovative research and proof-of-concept projects, demonstrating the practical application of **Vector Databases, Large Language Models, various AI frameworks for fine-tuning models**, and a diverse array of AI skills in the dynamic realm of Speech AI.

Research Assistant

UMD Computer Science Department, Big Data Visualization

Jan 2021 – Nov 2021 College Park, Maryland

- Co-authored a paper with Dr. Leo Zhicheng and team. The paper was accepted in the EuroVis Conference 2023 :
  A Comparative Evaluation of Visual Summarization Techniques for Event Sequences
- Conducted an insight-driven crowdsourcing experiment to assess visual summarization techniques (CoreFlow, SentenTree, Sequence Synopsis) on big data. Compared their summaries across tasks, datasets, granularity levels. Sequence Synopsis excelled in quality but incurred longer comprehension time. Participants evaluated content and interpretability, with implications for future techniques.

Data Science Intern

Oracle Cloud Infrastructure

June 2021 – Sept 2021 Remote

- **Tech Stack** : OCI Services (Object Storage, AutoML and Accelerated Data-science SDK) Keras, TensorFlow, Seaborn, Numpy

Data Science Intern

TATA Steel, TATA Group

May 2019 – Aug 2019 Tatanagar, India

EDUCATION

B.S. in Computer Science, General Business (minor)

University of Maryland

Aug 2018 – May 2022

Computer Science : Computer Networks, Data Science, Machine Learning, Computer System, Advance Algorithms.
Business : Financial Management, Managing People and Organizations, Strategic Management of Human Capital, Foundations of Marketing, Foundations of Accounting.

PROJECTS

AI and Machine Learning

- **Autonomous Multi-Agent Systems using LLMs**
  Designed and built autonomous multi-agent systems to execute complex, sequential workflows using locally hosted open-source LLM. Using **Llama 2, llama.cpp, LlamaIndex, Hugging Face**.
- **CT-Scan Prediction : 3D Convolutional Neural Network**
  Used a 3D Convolutional Neural Network to predict lungs scarring on CT scan images of COVID Patients. Used **TensorFlow**.
- **Fraud Detection Classification**

Data Science and Visualisation

- **Attendance Prediction : Decision Tree Model**
  Used supervised learning models to analyze 98000 employee data points to predict attendance in various training programs offered by the company.
- **Black Lives Matter : Fatal Police Shooting Analysis**
- **Moneyball Analysis (Oakland Athletics)**

Network Protocol

- **Chat Server**
  Implemented a chat server that allows clients to engage in group chats in chat rooms and send private messages to one another. Used language **C++ and UDP Protocol**.

TECHNICAL SKILLS

- **PROGRAMMING LANGUAGES :**
  JAVA | C++ | C | Python | SQL | HTML | CSS | JSON
- **FRAMEWORKS :**
  Tensorflow | PyTorch | Hugging Face Transformer | Langchain | Spring | Flutter | Vert.X | Pandas | Scikit-learn
- **DATABASES :**
  Oracle DB | MySQL | MariaDB | AWS RDS | Apache Cassandra | Vespa Vector DB
- **TOOLS :**
  Postman | Swagger | Mockito | Hibernate | JPA