

Mitravinda K M

+1 (312) 868-8623 | mitravinda462@gmail.com | [LinkedIn](#) | [GitHub](#) | [Personal Website](#)

EDUCATION

University of Illinois at Chicago (UIC) | Master of Science in Computer Science | GPA - 4.0/4.0

Aug 2023 - May 2025

Coursework: Advanced Machine Learning, Neural Networks, Data & Text Mining, Natural Language Processing, Computer Vision, Computer Algorithms

PES University | Bachelor of Technology in Computer Science

Aug 2018 - May 2022

Coursework: Machine Learning, Deep Learning, DBMS, Data Science, Data Analytics, BigData, Python App Programming, Information Retrieval

TECHNICAL SKILLS

- **Programming Languages:** C | C++ | Python | Shell Scripting
- **Front-end and Back-end technologies, Database:** HTML | CSS | JavaScript | Bootstrap | React | Node.js | PHP | .NET | MySQL | MongoDB
- **Tools:** AWS | Hadoop | PySpark | Android Studio | Git | Electron.js | Scilab | Arduino IDE
- **Python Libraries:** Pytorch | Tensorflow | Transformers | Langchain | NumPy | Matplotlib | Pandas | scikit-learn | Plotly | OpenCV | NLTK

EXPERIENCE

Data Scientist Intern | Amazon, USA | AWS, GenAI, LLM, Classification, Interpretability

Jun 2024 - Aug 2024

- Engineered data pipeline to curate & process ~3TBs security alert data from disparate sources, utilizing AWS services- **Lambda, S3 & SageMaker Studio**
- Trained **RandomForest** model to predict malicious & benign security alerts, F1 score & AUC **96.2%**, improving security engineering efficiency by **50%**
- Integrated **Claude LLM-powered explanation system** that leverages model outputs & interpretation metrics to explain model's behavior in natural language

Research Assistant - AI Developer | UIC, USA | Llama 2, HuggingFace, LangChain

Jan 2024 - May 2024

- Built a **chatbot** for querying CSV data from UIC's administrative department by integrating **Llama 2 model with LangChain**
- Developed a custom CSV data processing pipeline - generated embeddings with **HuggingFaceEmbeddings** & stored in a **FAISS** vector store

System Performance Analyst | IBM, India | CI/CD, Linux, KVM, Shell Scripting, Prediction, Data Visualization

Jul 2022 - Jul 2023

- Improved **database write performance** of PostgreSQL's benchmark pgbench by **2.5x** on IBM Power Systems
- **Optimized** Linux CI test data collection, run & analysis bringing **3.5x** improvement in **runtime** & **82%** reduction in workload **space consumption**
- **Built a data management & visualization tool** to manage & visualize workload-performance data across various Linux builds & **identify regression**

Systems Intern | IBM, India | CI/CD, Jenkins, Power System, Performance Analysis, Git

Jan 2022 - Jul 2022

- Analyzed performance of multiple **Linux benchmarks & cryptographic ciphers** across multiple **RHEL and SUSE** kernel releases on IBM Power systems
- Worked with **Hardware Management Console(HMC) & Virtual I/O servers(VIOS)** to test IBM Power systems' performance
- Worked on **CI pipeline** using **Jenkins** that automates workload-execution, obtaining performance output and **identifying regressions**

Teaching Assistant - Blockchain & Statistics for Data Science | PES University, India

Aug 2021 - May 2022

PROJECTS

Quora Duplicate Question Detection: Comparative Analysis | PyTorch, NLTK, Sklearn, NumPy, Pandas, Matplotlib

- Applied text **preprocessing** techniques: tokenization, stop-word removal, stemming & lemmatization on Quora question pairs
- **Embedded** the question pairs using **BagOfWords, Word2Vec and TF-IDF** representations on **batches** of the huge question-pair dataset
- **Classified** duplicate question pairs using traditional models **SGD-Classifer** (acc: 73.24%), **Naive Bayes Classifier** (acc: 74.06%), **XGBoost** (acc: 81.99%)
- **Fine-tuned BERT** model(acc: 78.265) & analyzed its performance against **traditional models**; compared **impact of embedding** on model performance

Face Sketch-Photo Synthesis & Recognition | Tensorflow, OpenCV, PIL, NumPy

- Built a framework to **convert face-photos to face-sketches** using **Two Scale Image Decomposition with Bilateral Filtering**
- Trained a 9 layered **Convolutional Neural Network**, post preprocessing, on the **celebA** database to **convert face-sketches to photos**
- Employed **Fisherface Linear Discriminant Analysis** to perform **facial recognition** of face-photos with an **accuracy of 91.875%**

Twitter Sentiment Analysis | PyTorch, Transformers, Transfer Learning, NumPy

- Processed **tweets** on **presidential debate** between Obama & Romney, handled mentions & URLs, tokenized tweets using **RoBERTa tokenizer**
- **Fine-tuned RoBERTa** model to **classify sentiments** expressed in the tweets about Obama & Romney into **positive, negative & neutral**
- Built framework with **64%** accuracy & F1 by training **2 fully connected layers** added on frozen RoBERTa layers, using cross-entropy loss & Adam optimizer

Character-level Text Generation LSTM | PyTorch, NumPy

- **Implemented** an LSTM trained on a dataset of names post preprocessing the input into a length-11 sequence of 27 dimensional vectors
- **Softened** the algorithm by **extracting top-10 most probable predictions** and selecting a **random sample** from them to predict the next letter

Data Analytics on Mental Health in Tech & Tech Employees | Sklearn, NumPy, Pandas, Matplotlib, Plotly

- **Predicted** possibility of being diagnosed with a mental health issue using **Gradient Boost Classifier** with an accuracy of **93.939%**
- **Clustered** employees into 3 risk-clusters, high, medium and low using **Spectral Clustering** with a Calinski-Harabasz index of: **316.76**; Computed **risk-score**
- Analyzed the **impact of workplace factors**; Performed **multi-year study** on pandemic's impact & the **mental health scenario** in tech **pre & post-COVID-19**

PAPER PRESENTATIONS & PUBLICATIONS

- Mitravinda, K. M., et al. "Face Sketch-Photo Synthesis and Recognition" *International Conference on Image Processing and Capsule Networks*. Cham: Springer International Publishing, 2022. [DOI](#); Presented at the [3rd International Conference on Image Processing and Capsule Networks](#)
- Mitravinda, K. M., Devika S. Nair, and Gowri Srinivasa. "Mental Health in Tech: Analysis of Workplace Risk Factors and Impact of COVID-19" *SN computer science* 4.2 (2023): 197. [DOI](#); Presented at the [3rd International Conference on Adaptive Computational Intelligence](#)
- Mitravinda, K. M., and Sakshi Shetty. "Employee Attrition: Prediction, Analysis Of Contributory Factors And Recommendations For Employee Retention" *2022 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE)*. IEEE, 2022. [DOI](#)