Nithish Reddy Konakati

nithishreddy.konakati@gmail.com | Nithish Reddy K | NithishAjju

PROFESSIONAL EXPERIENCE

Research Analyst | Texas Tech University | Texas **Jan 2024 -present**

- Leveraged an **MLP** model using historical student data to predict graduation likelihood, achieving 82% accuracy and 87% ROC-AUC, enabling early academic interventions for at-risk students.
- Applied **DistilRoBERTa** for aspect-based sentiment analysis and LDA for topic modeling on student feedback, boosting the **F1 score** by **15%.** Extracted actionable insights to improve course quality and teaching effectiveness.

Data Scientist | Switch Mobility | Chennai, India **Jan 2021 – June 2023**

- Developed a battery anomaly detection model using unsupervised machine learning. Reduced high-dimensional data by **60%** with an **autoencoder** model, then applied **k-means clustering**, achieving **78%** accuracy in detecting anomalies.
- Designed a modular Python framework to automate the processing and analysis of CAN logs, **reducing workload by 80%.**
- Implemented an **XGBoost** model for driver behavior classification, achieving **85%** accuracy in predicting safety scores on a 5-point scale and improving energy consumption by **20%**, using a **sliding window approach** for fair score updates.
- Engineered a fully **automated EDA pipeline** for EV fleet analytics, retrieving data from S3, computing key performance metrics, and generating monthly reports, cutting manual effort by **90%**.
- Created a comprehensive suite of **Tableau dashboards** to analyze EV data, achieving a **15%** improvement in battery efficiency, a **5–10%** reduction in subsystem energy usage, and optimizing motor performance through data-driven insights.

Data Analyst | Ashok Leyland, Product Development | Chennai, India **July 2019 – Dec 2020**

- Executed text analytics using **NLP** techniques such as **tokenization** and **LDA topic modeling** to classify unidentified IT support tickets into relevant categories, achieving a **coherence score** of **0.6** and reducing ticket categorization effort by **60%**.
- Enhanced the torque retention of H6 engine cylinder head bolts by **23%**, reducing gasket failures by applying **Lean Six Sigma** tools and **statistical analysis** (**Hypothesis Testing, ANOVA**) in Minitab.

PROJECTS

PhishShield-Scalable ML-Based Phishing Detection | Python | Scikit-learn | mlflow | AWS | Docker | GitHub Actions | Matplotlib

- Designed and implemented a robust **end-to-end machine learning pipeline** to detect phishing websites, achieving **96%** accuracy.
- Utilized **AWS S3** for cloud storage, **AWS EC2** for scalable compute, **Docker** for containerization, and integrated **MLflow** for efficient model tracking, version control, and experiment management.
- Automated **CI/CD** workflows with **GitHub Actions** to streamline model deployment and ensure seamless updates.

Q&A Network Security Bot | Python | Flask | LangChain | Llama-3.2-3B | Chroma DB | Hugging Face Embeddings

- Implemented a **Retrieval-Augmented Generation (RAG) pipeline** for an AI-powered Q&A Bot using Python and Flask, integrating **Llama-3.2-3B** for real-time query processing and Chroma DB for efficient document retrieval.
- Engineered a vector-based semantic search system using Hugging Faces **DistilBERT** embeddings and **cosine similarity** for document retrieval. Developed an interactive web interface with HTML, CSS, and JavaScript for seamless user interaction.

COVID-19 Data Augmentation using GANs | Python | PyTorch | Torchvision | Pandas | Matplotlib | CUDA

- Built a **Generative Adversarial Network (GAN)** model in Python using **PyTorch** for synthetic COVID-19 image generation, leveraging **CUDA** for GPU acceleration and **Torchvision** for efficient image transformation.
- Enhanced model performance by improving **IoU** by **15%** through hyperparameter tuning, optimizing image segmentation accuracy.
- Optimized generator and discriminator architectures with **convolutional layers** and **batch normalization** to enhance image realism.

YouTube Transcript Summarizer | Streamlit | Python | LLaMA-2 7B | LoRA | YouTube Transcript API

- Developed a Streamlit web app to extract YouTube transcripts and generate concise, context-aware summaries.
- Fine-tuned the **LLaMA-2 7B** model using **LoRA** and optimized prompt engineering, achieving a **10% improvement in BERT Score** for enhanced summarization accuracy.

SKILLS

Programming Languages & Visualizations: Python, SQL, HTML, CSS, JavaScript, C++, MATLAB, Tableau**,** PowerBI, Excel, Linux
ML & Data Science: LLMs, MLflow, Transformers, Neural Networks, Deep Learning, NLP, Statistics, Big Data**,** A/B Testing
Python Libraries & version control: NumPy, Pandas, Matplotlib, Seaborn, Plotly, Scikit-learn, TensorFlow, Keras, PyTorch, GitHub
Database & Frameworks, : MySQL, PostgreSQL, MongoDB, Hadoop, Spark, LangChain, Fast API, Flask, SDLC, Kafka
Cloud Technologies: Google Cloud Platform (GCP), Amazon Web Services (AWS), Azure, Docker, Kubernetes, CI/CD, Git, Airflow

EDUCATION

Master of Science in Computer Science, *Texas Tech University | CGPA: 3.9/4* **Aug 2023 — present**
Relevant Coursework: Neural Networks, Intelligent Systems, Advanced Database Management Systems, Analysis of Algorithms
Bachelor of Technology in Electrical Engineering, *National Institute of Technology, Kurukshetra, India* **Aug 2015 — Apr 2019**