

Abhishek Teja Goli

3706 Kettering Ct, Fairborn, OH- 45324 || g.abhishekteja@gmail.com || 937 443 7879

Linkedin : www.linkedin.com/in/abhishektejagoli

WORK EXPERIENCE

DataBeat

Telangana, India

Data Analyst Intern

Feb 2024 – July 2024

- Developed a deep learning solution using the ResNet architecture, achieving 94% accuracy in leaf disease classification for Envu. Fine-tuned hyperparameters through grid search optimization, reducing disease identification time by 30%.
- Redesigned Python-Django data ingestion pipelines, cutting processing time by 40% using multiprocessing and implementing best practices for data warehouse management with GAM including incremental loading patterns, API rate limiting management.
- Enhanced ETL workflows using Azure Data Factory by implementing metadata-driven templates and dynamic pipeline generation, improving pipeline efficiency by 50%, enabling real-time analytics for client dashboards.
- Designed an automated ad forecasting model for Nexstar Media Group, utilizing scikit-learn and Power BI, reducing manual reporting efforts by 70% and improving forecast accuracy by 15%.

PROJECTS

Organization-Specific LLM Application | [Link](#)

(Ollama Llama 3.2, Faiss, SQLite3, RAG, Python, APIs)

- Architected an enterprise-grade Retrieval-Augmented Generation (RAG) system enabling creation of custom LLMs using proprietary data, increasing information retrieval accuracy by 85%.
- Engineered a scalable vector database with FAISS and SQLite3, integrating IndexIVFFlat indexing for sub-500ms query responses on 10GB+ datasets. Implemented Llama 3.2 (3B parameters) with plans to upgrade to Llama 3.3 (70B) for enhanced performance.
- Developed a Python-based RESTful API processing multiple files formats and transforms them into optimized embeddings using intelligent document chunking algorithms.

Smart Movie Recommendation System | [Link](#)

(Tensorflow, NumPy, pandas, Flask)

- Engineered an emotion-based recommendation engine using Natural Language Processing and deep learning (LSTM, RNN) analyzing user sentiment to deliver personalized movie suggestions with 95% accuracy.
- Integrated hybrid filtering algorithms (collaborative and content-based), improving recommendation relevance by 35%, scaling to 100+ concurrent users using load balancing and database sharing.
- Developed an evaluation framework using precision, recall, and F1 score metrics, optimizing model performance.

FaceTrack: Real-Time Attendance Management System | [Link](#)

(OpenCV, NumPy, Scikit-learn)

- Implemented a real-time computer vision system using Haar Cascade algorithms and KNN classification that achieves 92% facial recognition accuracy, enabling simultaneous monitoring of up to 50 individuals for attendance tracking.
- Optimized the database structure, reducing identification latency by 65%, and automated reporting workflows, cutting administrative workload by 80% using efficient indexing techniques, such as B-tree and hash indexing.

Mental Health Severity Classification System | [Link](#)

(Scikit-learn, Pandas, NumPy, PCA, K-Fold Cross-Validation)

- Build up a machine learning-based system to classify mental health conditions into five severity levels (Normal, Mild, Moderate, Severe, Extremely Severe) using the DASS-42 and Ten Item Personality Inventory datasets.
- Applied Decision Tree, Naive Bayes, and K-Nearest Neighbor algorithms, with Decision Tree achieving the highest classification accuracy. Used PCA for dimensionality reduction and K-Fold Cross-Validation to evaluate model performance.

EDUCATION

Wright State University

Dayton, OH

Master of Science in Data Science | GPA: 4.0/4.0

Aug 2024 – Present

Mahatma Gandhi Institute of Technology

Telangana, India

B. Tech in Computer Science and Engineering (Data Science) | GPA: 3.06/4.0

Aug 2020 – June 2024

SKILLS

Programming Languages : Python, SQL, Java

Database : MySQL, PostgreSQL, MongoDB

AI/ML Expertise : Machine Learning, Deep Learning, LLMs, RAG, NLP, MLOps, OpenCV, Tensorflow

Cloud & Tools : Microsoft Office, AWS, Power BI, Azure, Docker, Kafka, Hadoop, Git

Soft Skills : Communication, Collaboration & Teamwork, Problem-Solving

CERTIFICATIONS

- Certified in **Data Science** by IBM.
- Google Data Analytics** from Google
- Certified by NPTEL in **Natural Language Processing (NLP)**.
- Earned **Python** and **Machine Learning** certification from **Kaggle**.