

Abhishek Teja Goli

(937)-443-7879 | abhishekteja45@gmail.com | Dayton, OH, USA | [Linkedin](#)

WORK EXPERIENCE

Topcon Healthcare

Remote

Data Science Intern

June 2025 – Present

- Optimized AutoMorph for Topcon fundus images, enhancing image-quality grading and vessel segmentation accuracy, leading to improved scalability for future development.
- Built Python-based data cleaning tools and tracking systems, reducing data processing time from hours to seconds, and a PPT generator that creates slides from Excel references and PNG files, eliminating manual reporting.
- Automated Big Data preparation with Python scripts, cutting dataset prep time by 60% and enabling scientists to analyze 3x more studies per week, accelerating clinical insights for patients.
- Evaluated logistic regression models with OCT-derived metadata, improving glaucoma detection with measurable gains in diagnostic accuracy.

Research Assistant, KASTLE lab

Wright State University

March 2025 – Present

- Conducting applied research under Dr. Cogan Shimizu, focusing on knowledge graphs, LLMs, and semantic technologies.

DataBeat

Telangana, India

Data Analyst Intern

Feb 2024 – June 2024

- Engineered deep learning models using ResNet, LSTM, and RNN to improve crop disease classification accuracy to 94%.
- Redesigned Python-Django data ingestion pipelines, cutting processing time by 40%. and implementing best practices for data warehouse management.
- Streamlined ETL workflows using Azure Data Factory and Python, resulting in 50% faster data processing and enabling real-time analytics for client dashboards.
- Built a predictive analytics model for ad forecasting (Nexstar Media Group), reducing manual reporting by 70% and improving forecast accuracy by 15%.

PROJECTS

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

- 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, optimizing embedding efficiency by 40%.

Smart Movie Recommendation System | [Link](#)

- 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.
- Enhanced precision-recall metrics through A/B testing, improving model accuracy by 20%.

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

- Developed an ML model for mental health severity classification, achieving 92% accuracy.
- Evaluated Decision Trees, Naive Bayes, and KNN, optimizing F1-score by 15%.

EDUCATION

Wright State University

Dayton, OH

Master of Science in Data Science | GPA: 3.6/4.0

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

TECHNICAL SKILLS

Programming & Database	: Python, SQL, PostgreSQL, Excel
Libraries & Frameworks	: Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, Flask, Django, Streamlit
Data Science and Analytics	: ETL Pipelines, Data Cleaning, Data Wrangling, Exploratory Data Analysis (EDA), Statistical Modeling, Hypothesis Testing, A/B Testing
Machine Learning & AI	: Regression, Classification, Clustering, Predictive Modeling, NLP, Generative AI, LLMs, RAG
Data Visualization:	: Matplotlib, Seaborn, Plotly, Tableau, Power BI
Cloud and Tools	: AWS, GCP, Azure Databricks, Git, Jupyter Notebooks, VS Code

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**.