DHANUSH CHANDRA SHEKAR

+1 8128227923 dchandr@iu.edu linkedin.com/in/dhanushchandrashekar

Education

Indiana University Bloomington **August 2024 – May 2026**
Masters in Data Science *CGPA 3.71/////4*

Bangalore Institute of Technology, Bangalore **August 2020 – May 2024**
B.E in Computer Science and Engineering *CGPA 8.51/10*

Technical Skills

Languages: Python, SQL, R
Libraries & Frameworks: Pandas, NumPy, Scikit-Learn, StatsModels, Seaborn, Matplotlib, TensorFlow, Keras
Data Platforms & Tools: Snowflake, Tableau, Power BI
Analytics & ML Techniques: A/B Testing, Regression, Classification, Clustering, Feature Engineering, Predictive Modeling, KPI Development, Attribution Modeling
Data Engineering: ETL Pipelines, Data Cleaning, Transformation, SQL Joins, Window Functions, Data Aggregation
Statistics: Hypothesis Testing, Time Series Analysis, ANOVA, MSE Analysis, Recursive Feature Elimination (RFE)

Projects

Finance Article Q&A App | ***Python, Streamlit, FAISS, LangChain, Hugging Face*** **Dec 2024 – Jan 2025**

- Built a Streamlit-based Q&A web app that lets users enter a finance article URL and ask related questions, returning answers using a **Retrieval-Augmented Generation (RAG)** system.
- Used **Groq's LLM** through **LangChain** to generate answers by pulling relevant article content with help from **FAISS** and **Hugging Face** text embeddings.
- Created a scraper using `requests` and `BeautifulSoup` to fetch and clean article text—worked successfully on selected sites like Moneycontrol.
- Handled possible issues like bad URLs, missing content, and API errors with built-in error checking and clear user messages.

Pneumonia Detection Model | ***VGG16, Python, Keras*** **October 2022 – December 2022**

- Built a **CNN**-based image classifier using **VGG16** and transfer learning to detect pneumonia from chest X-rays with 95% accuracy.
- Preprocessed over 5,000 X-ray images by resizing, normalizing, and applying data augmentation to improve model generalization.
- Compared models like custom CNNs, **ResNet**, and **VGG16**, and selected VGG16 for its strong performance in precision and recall.
- Designed simple python visualization dashboards using streamlit with accuracy/loss plots and confusion matrix to help staff interpret results more easily.

Netflix Content Analysis | ***Python, Pandas, Tableau,Seaborn, Scikit-Learn*** **June 2024 – July 2024**

- Performed exploratory data analysis on 1.5M+ viewer interactions to uncover engagement patterns and inform release strategies, contributing to a 35% increase in viewership.
- Improved recommendation performance by applying **Recursive Feature Elimination (RFE)**, resulting in a 20% boost in user engagement.
- Reduced model error by 15% through predictive model tuning using **Mean Squared Error (MSE)** analysis.
- Analyzed correlations among genre, release timing, and engagement metrics to guide content curation decisions.
- Developed interactive dashboards in both **Tableau** and **Seaborn** to communicate actionable insights to stakeholders.

Experience

Enterprise Building Training Solutions (EBTS) **October 2023 – November 2023**
Data Science Intern *Bangalore, India*

- Built a real-time facial recognition attendance system using **Flask**, **OpenCV**, and **KNN**, reaching 95% accuracy and cutting manual work by 70%.
- Used OpenCV's **Haar cascades** for live face detection and a KNN classifier for recognizing and verifying student identities.
- Managed user data and attendance records with a file-based system and automated CSV logs for easy tracking and export.
- Created a simple and responsive web interface for users to mark attendance, making the system easy to use and deploy.

Certifications

AWS Certified AI Practitioner [Credential] **April 2024**