Kshitiz Agrahari

DELHI, INDIA
+91-6387348632 kshitizvayu@gmail.com Linkedin github

EDUCATION

Madan Moahan Malaviya University of Technology, Gorakhpur **2019 – 2023**
B.Tech - Computer Science and Engineering - ***CGPA - 8.32*** *Gorakhpur, Uttar Pradesh*

EXPERIENCE

Tata Consultancy Services **December 2023 – Present**
Machine Learning Engineer *Delhi, India*

- Spearheaded the development of a machine learning model that stratified customers into 5 distinct segments, leveraging over **100,000+** purchasing records, which catalyzed a **30%** surge in customer engagement.
- Executed comprehensive data cleansing and **normalization** processes, addressing **5% missing values** and **3% outliers**, and engineered **10+** key features, resulting in a **15%** enhancement in model accuracy.
- Conducted in-depth Exploratory Data Analysis **(EDA)** on a dataset comprising **1 million** data points, uncovering 3 major trends and 5 significant correlations using advanced **visualization** techniques such as **histograms, box plots,and scatter plots**.
- Applied a suite of clustering algorithms, including **K-Means, Hierarchical Clustering, and DB-SCAN**, on a dataset of **100,000+** customer records, achieving a **25%** enhancement in clustering precision.
- **Amplified** conversion rates by **25%** and elevated the average order value by **20%** through meticulously targeted marketing strategies derived from model-driven insights.
- Reduced customer churn by 15% and enhanced **Customer Lifetime Value (CLV)** by **20%** through the deployment of data-driven retention methodologies.

TECHNICAL SKILLS

LANGUAGES: C, C++, Python, SQL

CORE: Data Structures, Algorithms, Machine Learning, Artificial Intelligence, Deep Learning, **Transformers**, Data Science, Computer Vision, Neural Networks, Generative Adversial Network(**GAN**), Natural Language Processing(**NLP**), Data Visualization

Technologies/Frameworks: Numpy, Pandas, MatplotLib, Scipy, spaCY, keras, **PyTorch, TensorFlow**, Scikit Learn, LSTM, Bi-LSTM, Encoder-Decoder, Fasttext, BERT, **LangChain**, HuggingFace, YOLO, **Auto-Encoder**, NLTK, **openCV** , Large Language Model(**LLMs**)

TOOLS: VS Code, **Docker**, Render, Git & Github, Jupyter Notebook, Google Colab, Kaggle, Streamlit, Roboflow, LabelImg

PROJECTS

Virtual Interview System **November 2024 - January 2025**

- **Technology: LLM, LangChain, Whisper, FAISS, Streamlit**
- Designed and developed an **AI-powered** Virtual Interview Application using Streamlit, LangChain, and GroqLLM to conduct automated technical interviews.
- Integrated GPT-based Large Language Models (LLMs) to dynamically generate context-aware, domain-specific interview questions using **PyPDF2** for candidate resume parsing and **FAISS** for **vector embeddings**.
- Implemented speech synthesis using **gTTS (Google Text-to-Speech)** to convert interview questions into audio, enhancing accessibility.
- Utilized **Whisper (OpenAI)** for high-accuracy **Speech-to-Text (STT)** conversion, enabling precise voice-to-text transcription of candidate responses
- Implemented **timer-based** interview workflow,session state management in Streamlit to track interview progress, question playback, and response submission.

- Integrated **LangChain Memory** to maintain context in **multi-turn interviews**, allowing follow-up questions based on previous responses
- Incorporated **JSON-based response** storage for **structured data collection** and efficient processing of candidate answers.

Real Time Face-Mask Detection Web App **September 2024 - November 2024**

- **Technology: YOLOv8, OpenCV, Roboflow, Ultralytics, Transfer Learning**
- Designed and implemented a **real-time face mask detection** system using YOLOv8 and OpenCV, capable of detecting masked and unmasked faces in **live video streams** with high accuracy.
- Created a high-quality labeled dataset by **manually annotating** images using **Roboflow**, ensuring precise bounding box placements and well-defined class distributions for effective model training.
- Trained a custom YOLOv8 for **100** epochs, achieving Precision: **83.4%**, Recall: **79.3%**, and mAP: **84.4%**, optimizing model performance for real-world scenarios.
- **Integrated OpenCV** for real-time frame processing, **face detection**, and overlaying detection results with bounding boxes and labels.
- Implemented **data augmentation** techniques, such as flipping, rotation, and brightness adjustments, to improve model generalization across various lighting conditions and perspectives.
- **Fine-tuned** the YOLOv8 model using transfer learning, leveraging **pre-trained weights** to improve detection performance on the custom face mask dataset.
- Developed a **scalable framework** that can be further extended for deployment on edge devices, **cloud platforms**, or integrated into real-world surveillance systems for public safety applications.

Error Log Analyzer **March 2024 - June 2024**

- **Technology: Fasttext, SVM, RandomForest**
- Developed an AI tool to analyze test case logs and accurately identify error types.
- Tested the tool using various classification algorithms, including **SVM and Random Forest**, to determine the best performing model and utilized **FastText** for text representation and classification.
- Evaluated the tool's performance on multiple parameters to ensure robustness and reliability.
- Achieved a model accuracy of **80%** through continuous testing, tuning, and optimization.

EXTRACURRICULAR

- Solved **600+** problems on various platforms like **LeetCode** and **GFG**
- Gold Medalist in Chess in **Inter-Branch Chess Competition**