Riddhiman Ghatak

Roll No.:210122048
B.Tech - Chemical Science and Technology
Indian Institute Of Technology, Guwahati

+91-6290995330
riddhimanghatak@gmail.com
g.riddhiman@iitg.ac.com
github.com/riddhiman-ghatak
linkedin.com/in/Riddhiman Ghatak
CodeForces | LeetCode | kaggle

EDUCATION

Degree/Certificate	Institute/Board	CGPA/Percentage	Year
B.Tech. Major	Indian Institute of Technology, Guwahati	8.00 (Current)	2021-Present
Senior Secondary	CBSE Board	94.8%	2020
Secondary	WBBSE Board	97.0%	2018

EXPERIENCE

- **Jeeva AI** *May 2024 - July 2024*
  AI Research Intern Location: Remote
  - Developed a model using the **Coswara** dataset to predict **smoking status** from **cough samples**.Used **librosa** to convert audio to **Spectrogram** images and used **UMAP** for visualization.
  - Implemented **CNN**, **ResNet50**, and **ensemble of EfficientNetB0 and EfficientNetB7** and achieved **0.82 accuracy**, **0.85 precision**, and **0.76 recall** on test data. Took Guidance from research paper **"Listen2Cough"** by **XUHAI XU university of Washington.**
- **Menrva Technologies** *Nov 2023 - Feb 2024*
  Data Science Intern Location: Remote
  - Developed a **resume scraping** and **shortlisting** pipeline. Implemented a **RAG** framework using **LangChain** and **ChromaDB** with **Mistral-7B-Instruct** to extract relevant information in **JSON** format.
  - Applied **Semantic Search** and **Re-ranking** algorithms, including **Cosine Similarity**, **Cross-Encoders**, **BM25**, and **LangChain's FlashRank** and achieved **BERT Score** of **0.56** for RAG and a **nDCG Score** of **0.47** for Re-ranking.
- **University of New South Wales** *Feb 2023 - April 2023*
  Research intern - advisor: Prof Rohitash Chandra Location: Remote
  - Developed a predictive model for GDP of South Asian countries. Applied **Exponential Smoothing** and conducted **Stationarity Checks**. Utilized **ARIMA** as a baseline model.
  - Then used **FBProphet** and **LSTM**, with hyperparameter tuning and Achieved **MAPE** of **11.41** , and **Theil's U Statistic** of **0.67**, demonstrating significant improvement over naive forecasting models.

PROJECTS

- **Quiz Crafter** *June 2024 - July 2024*
  Sparkathon, Walmart Github
  - Developed a pipeline using **Meta-Llama-3-8B** via **Groq API**, implementing a **RAG** framework Utilizing **ChromaDB** for generating MCQs from uploaded PDF study material and used **LangChain output parser** to parse output in json format .
  - Achieved **ROUGE** score of **0.47** and **BERT score** of **0.84**. Deployed entire pilpeline on **gradio**.
- **Super resolution-photo-mosaic** *Feb. 2023 - March. 2023*
  Coding club-IIT Guwahati, Techevince 9.0 Github
  - Developed a pipeline to generate a **high-resolution panorama** by stiching together fragmented, low-resolution images.
  - Initially used **LDMSuperResolutionPipeline** from **Huggingface Diffuser** to improve resolution then **Image Stitcher** to generate panoramic image. Achieved avg **CLIP-directional-similarity** score of **0.079**.
- **Library-seat-detection from CCTV footage** *Jan 2023 - March 2023*
  Consulting and Analytics club -IIT Guwahati Github
  - Trained a model using **YOLOv5 object detection model** to detect **empty seat** in library from live CCTV footage.
  - Used **YOLOv5 object detection model** on CCTV footage for object detection. It uses **centroid distance** to detect empty seat and achieved **mAP** score of **0.76** with **IoU thresold** of **0.65**.

TECHNICAL SKILLS

- **Programming** : Python, C/C++
- **Development** : HTML/CSS, JavaScript*, React.js*, Flask, Bootstrap, Material-UI*, Tailwind CSS*, Git, Docker*
- **Data Science** : Numpy, Pandas, Matplotlib, Seaborn, Scikit-learn, Pytorch, Power BI, Keras, Tensorflow, NLTK, TextBlob, OpenCV, Gradio, Streamlit, Langchain, Hugging Face, XGBoost, AdaBoost
- **DataBase Management** : MySQL
  * Elementary proficiency

ACHIEVEMENTS

- **Expert Kaggle:** : Currently rated as notebooks **EXPERT** in **kaggle** with **871** rank.   [Kaggle profile]
- **Knight Leetcode:** : Currently rated as **KNIGHT** in **Leetcode** with rating **1887**.   [Leetcode Profile]
- **Adobe Hackathon 2023:** : Got selected in final round in **Papyrus Nebulae Hackathon 2023**
- **Walmart Sparkplug Hackathon 2024:** : Got selected in final round amongst 14000+ particiapants accross india.
- **ENCODE 2023 Finalist** : Got selected in the **top 10 teams amongst 315+ teams across india**

KEY COURSES TAKEN

- **Computer Science:** : Introduction to computing(Theory and Lab), Data structures and algorithms(Udemy), Summer Analytics, Computational Chemistry(Theory and lab)