Aumkesh Chaudhary

aumkeshchaudhary@gmail.com
+91-7321999180
linkedin.com/in/aumkeshchaudhary
github.com/aumkeshchaudhary

Education

Indian Institute of Technology, Patna — Patna, India
Bachelor of Science (Honours) in Computer Science and Data Analytics — July 2023 - Present
CPI: 8.98

Relevant Experience

Capstone project: *CareerNavigator* — Feb.2024 - June 2024
Under the supervision of ***Dr. Kuldip Singh Patel***, Asst. Prof., IIT Patna

- Developed a Machine Learning model, ***CareerNavigator***, to assess candidates' employability by analyzing key attributes and predicting suitability for a job role.
- Cleaned, pre-processed, and performed feature engineering on the dataset containing **70k+** datapoints.
- Designed, trained, and evaluated multiple algorithms, utilizing performance metrics such as *accuracy*, *confusion matrix,* and *F1-score* to optimize model effectiveness.
- Selected kernel support vector machine as the best-fit model with the highest accuracy of **80%** and F1 score of **0.82**.
- Successfully showcased the project to faculty and industry experts, receiving recognition for its innovation and effectiveness.

Projects

AudeX: Vision-to-Voice Conversion Model

- Developed a dynamic web application combining Optical Character Recognition (OCR) and Text-to-Speech (TTS) technologies to improve accessibility.
- Integrated ***Tesseract*** for multi-language OCR and ***Web Speech API*** for TTS, enabling accurate text extraction from images and high-quality text-to-speech conversion.
- Designed a responsive interface with intuitive navigation, ensuring a seamless user experience.
- Enabled PDF export, word/character count, and keyword search for efficient document handling.
- Developed functionality for managing user profiles, such as signup, login, activity tracking, and editable data.

Advanced Text-to-Speech Optimization

- Fine-tuned Microsoft's SpeechT5 model to improve pronunciation of technical English terms focusing on modifying the phonetic representation to ensure precise pronunciation of abbreviations and acronyms.
- Achieved **25%** enhancement in speech quality of over the baseline TTS model, with significant improvements in handling technical terms.
- Optimized the baseline model to generate a native Italian voice by enhancing pronunciation, prosody, and stress patterns in line with the phonological rules of the Italian language, significantly improving speech quality and naturalness compared to other existing models.
- Harnessed tools like Transformers, PyTorch, and Hugging Face Datasets to implement advanced machine learning and NLP techniques, ensuring optimal model performance and reliability.
- Implemented dynamic ***quantization*** techniques to optimize model efficiency, reducing memory usage and boosting inference speed by up to **30%**.

Computer Vision Object Detection System

- Developed an object detection system utilizing ***YOLOv5*** and ***PyTorch***, designed to capture and process individual photos for object detection.
- Optimized the computer vision pipeline to achieve **30+ FPS** processing speed for 640x640 input resolution.
- Created a visualization system to render detection results with color-coded bounding boxes and confidence scores ranging from 0 to 1.
- Applied **non-maximum suppression** with **IoU** threshold of **0.3** for optimal detection accuracy.
- Added support for 80+ COCO dataset object types, with confidence threshold filtering at 0.3.

Skills

- **Programming skills** : Python, Java, R, C, MATLAB
- Web Development: HTML, CSS, JavaScript
- **Developer Tools** : Git, VS Code
- **Other Tools** : MS Office, Tableau, Canva
- **Professional skills:** Object Oriented Programming(OOPs), Machine Learning, and Deep Learning

Extracurricular activities

- **Piano:** 6+ Years of experience in classical and contemporary styles; actively compose original pieces that explore and innovate across diverse genres and elements of music.
- **Guitar:** 2+ Years of experience in Indian and western contemporary styles; exploring fingerstyle techniques and songwriting.