Malay Jain

06232155888|malayjain1234@gmail.com|Linkedin|Portfolio|GitHub

Education

Course	Institute	Year	Percentage/CGPA
B.Tech	Sagar Institute of Research and Technology, Bhopal	2026	7.61
12th	St. Joseph's Convent S.S. School, Sagar	2022	84.2%
11th	St. Joseph's Convent S.S. School, Sagar	2020	76.8%

Projects

Cardiovascular Disease Name Prediction

- This model supports early detection and diagnosis of heart disease types, enabling more accurate treatment decisions and proactive healthcare. It can serve as a foundational tool in clinical decision support systems to assist healthcare providers with diagnosis and improve patient outcomes.
- **Dataset Engineering**: Preprocessed heart disease dataset and encoded categorical variable.
- **ML Model** : Implemented **Random Forest Classifier** with **Gini criterion** and optimized parameters.
- **Deployment:** Deployed the prototype model using **Streamlit**, a library of python.
- **Performance Metrics**: **1) Precision/Recall** : 100% **2) Weighted Average Accuracy**: 81%

Expert - Candidate Matching System

- This model streamlines expert selection by providing precise, automated candidate-expert matching, saving time and improving relevance. It enhances decision-making in interviews and mentorships with data-driven recommendations tailored to candidate profiles, benefiting organizations.
- **Using NLP:** Utilized NLP techniques **(SpaCy NER)** to extract key phrases from expert and candidate profiles, focusing on qualifications, skills, research areas, and professional accomplishments.
- **BRET Algorithm:** Integrated BERT-based embeddings **(Sentence-BERT)** to generate semantic representations of profiles, enabling precise similarity calculations.
- Applied **cosine similarity** to rank and recommend the top 5 most relevant experts for each candidate.

Movie Recommendation System

- Developed a content-based Movie Recommendation System that recommends similar movies based on user input, enhancing user experience and engagement.
- **Data Engineering:** Merged multiple movie datasets by title, selecting essential features (genres, keywords, cast, director, overview) to refine content-based filtering for accurate recommendations.
- **Data Preprocessing:** Conducted extensive text preprocessing with tokenization, stemming, and text normalization to create cohesive tags, consolidating key attributes for each movie.
- **Algorithm:** Applied **CountVectorizer** to convert text data into vector form and calculated cosine similarity between vectors, efficiently ranking movies by similarity.
- **Deployment:** Deployed the model using **Streamlit**, providing a responsive interface for users, and utilized pickle to save the model and similarity matrix for fast, reliable data retrieval.

Community Leadership

1. **Google Developers Group Lead(GDG Lead)[Chapter SIRT] Sep 2024- Present**
2. **GeeksForGeeks Campus Mantri - Apr 2024- Present**
   - **Event Organization and Leadership:** Organized and led workshops, hackathons, and technical events on diverse technologies, fostering hands-on learning and increasing participant engagement.
   - **Strategic Marketing Initiatives:** Developed and executed strategic marketing initiatives across multiple platforms, boosting event attendance and cultivating an active tech community.
   - **Project-Based Learning Facilitation:** Project-based learning for students, enhancing their proficiency in various technologies while incorporating relevant industry courses to develop practical skills.
   - **Personalized Guidance and Support:** Provided tailored guidance to help members achieve project milestones, enhance their learning journeys, and build technical confidence.

Skills

- **Technical Languages:** Python, C++, C, SQL
- **Frameworks**: Flask, TensorFlow
- **Tools**: VS Code, Jupyter, MySQL, Power BI, Streamlit

Certifications:

- **Python:** HackerRank
- **SQL:** HackerRank