SHIVAM CHAUDHARY

+91 9628102333 shivachadhhary@gmail.com shivam-chaudhar shivamgithubok Kaggle:shivamchaudharys

Education

B-tech in Electronics and Communication, specialization in VLSI Engineering **2022 – 2026**
Indian Institute Of Information Technology, Senapati Manipur ***CGPA 7.69/10***

Projects

Automotive Valuation Intelligence: Car Price Prediction **GitHub**

- Developed a predictive model to estimate car prices based on various features using a multivariate regression approach.Performed data preprocessing, feature engineering, and exploratory data analysis to identify key factors influencing car prices.Utilized linear regression, decision trees, and ensemble methods, achieving high accuracy and evaluating performance with metrics like RMSE.

Spam Message Detection Link **GitHub**

- Implemented a Natural Language Processing (NLP) pipeline to classify messages as spam or ham using the SMS Spam Collection dataset.Applied machine learning algorithms like Naive Bayes and Support Vector Machines, achieving high accuracy and optimizing performance through hyperparameter tuning.

Movie Recommendation System Link **GitHub**

- Built a collaborative filtering-based recommendation system to suggest movies to users based on viewing history preferences. Utilized content-based filtering and matrix factorization techniques to improve recommendation quality. Integrated TMDB API to fetch movie posters and used the TMDB dataset for enhanced recommendations.

Sentiment Analysis of IMDB Movie Reviews Link **GitHub**

- Conducted sentiment analysis on IMDB movie reviews using machine learning models and NLP techniques, categorizing reviews as positive or negative. Implemented deep learning models like LSTMs and GRUs for advanced text processing, achieving significant accuracy improvements. **Fine-tuned a pre-trained DistilBERT model using the PyTorch library**, enhancing the system's performance on sentiment classification tasks.

Dialogue Summarization Using Fine-Tuned BART-Large-CNN **GitHub**

a text classification and summarization model by fine-tuning **Facebook/BART-Large-CNN** on the **SAMSum dataset** using Hugging Face Transformers and PyTorch. Utilized AutoTokenizer for efficient tokenization and preprocessing of conversational text. Fine-tuned the model to generate concise and coherent summaries while preserving key information. Optimized training with GPU acceleration for faster convergence. Evaluated performance scores to ensure high-quality outputs. Integrated the model into a user-friendly application for summarizing dialogues, enhancing productivity in tasks like meeting notes, customer support logs, and chatbot conversations.

Work Experience

Data Science Intern At Edunet Foundation **12 dec 2024 – 30 jan 2025** **Link**

- Worked on a project to develop a **plant disease detection system** using **deep learning techniques**,Preprocessed and analyzed a large dataset provided by Edunet Foundation, ensuring data quality for training and validation purposes.Achieved an accuracy of **95% on the training dataset** and **94% on the validation dataset**, showcasing the effectiveness of the developed model.Explored key metrics such as **precision, recall, and F1-score** to evaluate the model's performance and validate its reliability for real-world applications.Utilized tools like **TensorFlow**, **Keras**, and **Jupyter Notebook** to implement and fine-tune the model, demonstrating proficiency in machine learning frameworks.

Technical Skills

Machine Learning: Supervised Learning, Unsupervised Learning, Regression, Classification
Natural Language Processing: Embedding technique TFIDF, BOW ,Word2Vec, Navie Bayes
Large Language Model: Transfromers , Self Attension, Encoder-Decoder
Deep Learning: Artificial Neural Network, CNN, RNN, LSTM, GRU, GAN
Data Analysis: Data Preprocessing, Feature Engineering, Data Visualization(Mataplotlib, Seaborn)
Languages: Python, C, C++, JavaScript, HTML, CSS,
Web Development: Node.js, Express.js, EJS , React, Bootstrap,PostgreSQL, MySQL, MongoDB
Developer Tools: Jupyter, Google Colab, Postman, VS Code ,Linux, GitHub, Git
Libraries: Pandas, Numpy, Matplotlib, Sci-Kit Learn, XGBoost, Keras, Tensorflow, Streamlit

Certificates

Supervised Learning Specialization by by Andrew Ng **Coursera**
Advanced Algorithm by Andrew Ng **Coursera**
Unsupervised Learning by Andrew Ng **Coursera**
Full Stack Web Development Bootcamp by Angela Yu
Selected in top 30 in CodeRush by Codeforces Master and featured in the ***Hall of Fame.***