Saarthak Gupta

linkedin | Portfolio | Github | Medium

Email : Saarthak.gupta.mec22@itbhu.ac.in
Mobile : 7303277192

EDUCATION

- **Indian Institute of Technology(BHU)** — Varanasi, India
  Integrated Dual Degree in Mechanical Engineering; CGPA: 8.33 — 2022 - 2027
- **S.R.D.A.V Public School** — Delhi, India
  CBSE(Class XII); CGPA: 93.8% — 2022

SKILLS

- **Languages:** Python(pandas, numpy, matplotlib, seaborn), SQL
- **Frameworks:** Langchain, HuggingFace, pytorch, scikit learn, tensorflow, keras
- **Tools:** git, Mysql, powerBI

PROJECTS

CLV PREDICTION FOR AUTO-INSURANCE COMPANY | LINK

Developed a comprehensive Customer Lifetime Value (CLV) prediction model **to improve customer retention and marketing strategies**, leading to more informed business decisions.

- Leveraged machine learning techniques to predict CLV, optimizing hyperparameters with GridSearchCV for accurate predictions.
- Conducted comprehensive data preprocessing **followed by detailed EDA using statistical summaries, distribution analysis, and visualizations** to uncover insights.
- Created a web interface and dashboard for CLV predictions and customer insights, and **implemented a Q&A platform using Google's Gemini LLM to facilitate natural language queries, enhancing marketing strategies** and customer retention efforts

MACHINE TRANSLATION USING TRANSFORMER | LINK

Developed a machine translation system **using the transformer architecture in PyTorch** to translate **English to Hindi**, utilizing the IIT Bombay English-Hindi dataset.

- Built the **entire transformer model from scratch using PyTorch**, including attention mechanisms and positional encoding, to effectively capture language patterns.
- preprocessed the data to **create efficient tokenized inputs** for bilingual translation.
- Assessed the **model's performance using BLEU scores** to ensure translation accuracy and improve model performance through iterative tuning.

EKYC | LINK

The objective was To develop an advanced EKYC (Electronic Know Your Customer) verification system using RCNN and DeepFace for robust customer authentication and identity verification.

- Implemented contour detection to extract ID card images. Enhanced OCR accuracy with **RCNN (easyocr) for text detection and CRNN-CTC for text recognition.**
- Utilized **DeepFace for accurate face comparison, ensuring seamless verification** and reducing fraudulent activities.
- Built a user-friendly interface using Flask to streamline the EKYC process, enabling efficient customer registration and verification.

HONORS AND AWARDS

- 2nd runner up in HackFest, a National data science competition conducted by Tredence Analytics.
- Finalist in Convolve (PAN IIT Data Science event).
- Data Science author for level up coding on Medium.