Anubhav
Employee Code- INF-1288

Summary

Results-driven AI & Data professional with expertise in **Prompt Engineering, Python, and LLM APIs (OpenAI, Gemini, Hugging Face)**. Skilled in **data visualization (Tableau, Power BI)**, SQL, and automation pipelines. Designed and implemented **n8n workflows** to automate ad content using LLMs, generate marketing creatives with **Canva API**, and produce video content with **Veo3**, enhancing efficiency of digital campaigns. Built a **regression model to predict water pH values from multiple water quality features** and deployed it via a **Flask API** for practical use. Strong background in API development, prompt optimization, and translating complex insights into clear business outcomes.

Skills

- SQL
- Tableau
- Power BI
- Python Programming
- Machine Learning
- Deep Learning
- Python Libraries
- EDA
- AWS
- PyTorch
- Git
- NumPy
- SeaBorn
- Docker
- Flask

Experience

Infinity Soft Systems
ML Engineer
Dec 2022 – Present

Roles & Responsibilities-

- Python Programming for data processing, automation, and AI workflows
- Created visually compelling reports and dashboards using Tableau, increasing stakeholder engagement
- Presented key findings to both technical and non-technical stakeholders, improving decision-making speed Translated complex data trends and patterns into clear, accessible visualizations, enhancing communication
- Designed and implemented **n8n workflows** to automate ad content creation using **OpenAI/Gemini APIs**
- Automated generation of marketing creatives using the **Canva API** and video content using **Veo3**
- Built and deployed a **regression model to predict water pH values from multiple water quality features** via a **Flask API**
- Collaborated with cross-functional teams to align AI-driven solutions and reporting outputs with business requirements
- Developed efficient SQL queries to extract and analyze large datasets, reducing data retrieval time
- Translated complex data trends and patterns into clear, accessible visualizations, enhancing communication effectiveness

Projects

Netflix - Data Exploration and Visualisation

https://drive.google.com/drive/folders/1pP_bsRnG9vbw5YqVWElznKdXrFJepmF_?usp=sharing

- Conducted thorough EDA to detect and manage outliers, ensuring data integrity and enhancing analytical accuracy
- In the majority of top countries, movies constitute a predominant content category over TV shows.
- There was an upward trend in release dates around the year 2000. However, between 2017 and 2019, a downward trend in release dates was observed.
- Movies tend to be released towards the end of the month and towards the end of the year. In contrast, TV shows often see their highest release peaks in the middle of the month. Additionally, a notable concentration of TV show seasons occurs around the 8th month of the year on Netflix

- Initially, Netflix added an equal number of movies and TV shows. However, in recent years, there has been a noticeable increase in the addition of movies compared to TV shows overall.
- International movies, drama, and comedy genres dominate preferences across most countries. Technologies Used: Python, Pandas, NumPy, Matplotlib, Seaborn Target SQL

https://drive.google.com/file/d/1ZoMLG0phdcJGHFvSXfvRMGMK9UwVq7LK/view?usp=sharing

- Developed SQL queries to retrieve and analyze data for business insights and reporting purposes Managed operations across 4119 cities and 27 states, facilitating robust business expansion and order fulfillment nationwide consistent year-on-year growth in order volumes, with a substantial surge observed post-2016, reflecting impactful business expansion and increased market demand.
- Noted a significant rise in order costs from 2017 to 2018, particularly during the period from January to August
- Driver Churn Prediction System Ola Drivers

https://github.com/Anubhavsinghchauhan/Churn_predict_for_car_service_provider_Anubhav

Tools & Tech: Python, Streamlit, XGBoost, Scikit-learn, Pandas, NumPy, Pickle, KNN Imputer, EDA
https://anubhavsinghchauhan-churn-predict-for-car-servic-predict-k9too7.streamlit.app/

- Processed and analyzed 19,000+ Ola driver records across 14+ features to understand churn patterns and retention factors.
- Engineered key features like churn labels from Last Working Date, income/rating trend indicators, and performed missing value imputation using KNN Imputer.
- Performed extensive EDA and outlier analysis, identifying key churn indicators such as city, designation, and grade distribution.
- Built a robust XGBoost classification model, achieving 90% accuracy, 92% precision, and 0.93 F1-score on the test set.
- Designed and deployed a user-friendly Streamlit web app enabling real-time churn prediction based on user input.
- Anubhav Porter Delivery Time Forecasting using Deep Learning

https://github.com/Anubhavsinghchauhan/Churn_predict_for_car_service_provider_Anubhav/blob/main/Anubhav_porter.ipy nb

- Pre processed and engineered features from 175K+ structured delivery records, including time-based variables (hour, day, weekday, month).

- Applied Local Outlier Factor (LOF) to detect and remove anomalies, improving model consistency.
- Trained a deep neural network (DNN) using TensorFlow (ReLU, Adam) with metrics: R^2 = 0.997, MAE =0.37 min, RMSE = 0.46 min.
- Conducted detailed EDA to uncover trends like peak delivery hours and most common order configurations.

Tools: Python, Pandas, TensorFlow, Scikit-learn, Matplotlib, Seaborn

Education

BE/B.Tech/BS Chandigarh University.