Charu Rajput

Delhi, India
Email: charurajput2@gmail.com | Phone: +91-8368600640
LinkedIn: linkedin.com/in/charu-rajput2 | GitHub: github.com/charuu2

PROJECTS

Personalized Movie Recommendation System (Link) **Jan 2024 - Feb 2024**

- Implemented a content-based filtering movie recommendation engine using CountVectorizer and **Cosine Similarity**.
- Preprocessed TMDB dataset using Pandas and NLP techniques (stemming, keyword extraction) for efficient data handling.
- Integrated the model with Pickle for fast storage and retrieval of data.
- Constructed an interactive Streamlit UI for real-time recommendations and fetched dynamic content (posters, trailers) via TMDB API.

Customer Churn Prediction Model – IBM Telecom Dataset (Link) **Jan 2025 - Feb 2025**

- Deployed a **customer churn prediction model** using **Decision Tree, Random Forest, XGBoost, and Logistic Regression**.
- Performed **data preprocessing**, **feature engineering**, and **hyperparameter tuning** (using **GridSearchCV**), achieving **78% accuracy** and an **AUC-ROC score of 0.82**.
- Evaluated model performance using **confusion matrix, precision-recall analysis, and classification reports**.

Amazon Alexa Sentiment Analysis using Machine Learning **Feb 2025 - March 2025**

- Built an NLP-based sentiment analysis model using Amazon Alexa reviews to classify customer feedback as positive or negative.
- Conducted EDA (Exploratory Data Analysis), visualizing rating distributions, word frequencies, and review lengths using Seaborn and Matplotlib.
- Preprocessed textual data (stopword removal, stemming, tokenization) and optimized feature representation using CountVectorizer and MinMaxScaler.
- Achieved **93.76% accuracy** and **AUC score of 0.92** using **XGBoost**, **Random Forest**, and **Decision Tree Classifiers**.
- Applied **k-fold cross-validation** and **GridSearchCV** for **hyperparameter tuning** to enhance model performance.
- Created visualizations like word clouds, rating distributions, and confusion matrices to interpret results and evaluate model efficacy.

E-commerce Sales Dashboard | Power BI (Link) **Oct 2024 - Nov 2024**

- Designed an interactive Power BI dashboard to evaluate sales performance and customer satisfaction, visualizing data across countries, regions, order priorities, and shipping modes.
- Analyzed sales data for Furniture, Office Supplies, and Technology categories, delivering actionable insights to drive decision-making.
- Summarized key metrics: $1.47M total profit, 178K units sold, and $12.64M total sales.
- Created dynamic visualizations (pie charts, bar charts, line graphs) to identify trends and enhance business strategies.

EDUCATION

Bachelor of Technology (BTech)-Computer Science (Data Science)

Amity University • Noida, Uttar Pradesh • Oct 2022 – Present • 7.12/10 CGPA

XII (ISC)

The Frank Anthony Public School • New Delhi • May 2021 - July 2022

CERTIFICATIONS

NPTEL: Python for Data Science 2024

Data Science Job Simulation - Forage 2025

TECHNICAL SKILLS

- **Programming Languages**: Python, Java, SQL
- **Libraries & Frameworks**: NumPy, Pandas, Scikit-learn, Seaborn, Matplotlib, NLTK, XGBoost, imblearn
- **Databases:** MySQL
- **ML & AI Concepts**: Regression, Classification, Clustering, Feature Engineering, Feature Selection, Data Preprocessing, SMOTE
- **Model Evaluation**: EDA, Content-based Filtering, TF-IDF
- **NLP & Deep Learning**: Natural Language Processing, Tokenization, Word Embeddings, Deep Learning Fundamentals
- **Tools & Platforms**: Power BI, Git, GitHub, Flask, Streamlit, Google Colab, Jupyter Notebook, Microsoft Excel (VLOOKUP, Macros, Pivot Tables, Power Query)
- **Algorithms**: Decision Trees, Random Forest, XGBoost, SVM, Logistic Regression
- **Currently Learning**: PyTorch, Tensorflow, AWS (EC2), Large Language Models (LLMs)

ACHIEVEMENTS AND AWARDS

NPTEL Gold Medalist - Technical Communication for Engineers (Top **1%**,**90% Score**).

1st Place - Netflix-Code-Heist Hackathon (GFG, Amity University Noida, Sept 2024), showcasing strong problem-solving skills and coding proficiency.