Mansi Kulshrestha

+91 6395756963 — mansikulshrestha87@gmail.com — www.linkedin.com/in/mansi-kulshrestha-38388121a — ://github.com/mansikulshrestha

Skills

Visualization	Tableau, Power BI, Seaborn	**Database Management**	SQL
Languages	Python, C++, SQL	**Machine Learning**	Python (Scikit-learn, TensorFlow),NumPy Pandas
Data Analysis Tools	Excel, Python, SQL, Tableau, Power BI		

Education

Vellore Institute of Technology (Expected Graduation: 2025)
Bachelor of Technology in Computer Science and Engineering
Specialization: Artificial Intelligence and Machine Learning
CGPA: 8.79

Relevant Coursework:Operating Systems, DBMS, Object-Oriented Programming

Projects

Movie Recommender System **Jul 2024 – August 2024**
Tech Stack: Python, Pandas, NumPy, Scikit-learn, Streamlit, Render, Pickle

- **Data Preprocessing: Cleaned and processed a dataset of 5000 movies, extracting key features like genres, keywords, cast, and crew.**
- **Feature Engineering: Applied `CountVectorizer` to convert textual features into numerical data and performed cosine similarity to recommend similar movies.**
- **Model Implementation: Implemented content-based filtering for movie recommendations using cosine similarity between movies.**
- **Deployment: Deployed the recommendation system as a real-time Streamlit app on Render, accessible at https://cinesuggest-1.onrender.com, allowing users to receive instant movie suggestions.**
- **Optimization: Improved model performance by utilizing key features and reducing dimensionality using `CountVectorizer` for better recommendations.**

Email Spam Classifier **Nov 2024 – Jan 2025**
Tech Stack: Python, Pandas, NumPy, NLTK, Scikit-learn, WordCloud, Matplotlib, Seaborn, XGBoost, Pickle

- **Data Preprocessing: Cleaned and processed a dataset of 5,000 emails, handling missing values, duplicates, and renaming columns.**
- **Text Preprocessing: Applied text normalization techniques including tokenization, stopword removal, stemming, and special character removal using `NLTK`.**
- **Exploratory Data Analysis (EDA): Analyzed text features such as character count, word count, and sentence count; visualized distributions of 'ham' and 'spam' emails using `WordCloud` and histograms.**
- **Model Implementation: Built and tested multiple classification models including Naive Bayes, Logistic Regression, SVM, and `XGBoost` for spam classification.**
- **Evaluation: Assessed model performance using accuracy, precision, and confusion matrix; fine-tuned models for optimal performance.**
- **Deployment: Saved the final model and TF-IDF vectorizer using `Pickle` for future use in email spam prediction systems.**

Sales Insights using Tableau **Jul 2024 – Aug 2024**
Tech Stack: Tableau, Excel, SQL

- **Data Collection: Extracted and processed sales data from multiple sources for analysis.**
- **Data Visualization: Created interactive dashboards in Tableau to analyze sales trends and product performance.**
- **Key Metrics: Visualized sales growth, top-performing products, and regional distribution.**
- **Trend Analysis: Created time-series charts to track sales trends and identify opportunities.**
- **Reporting: Automated monthly reports and presented insights using Tableau's storytelling feature.**

Co-Curricular Activities

- **Event Organizer: Coordinated workshops and guest lectures, enhancing student engagement.**
- **Networking: Built connections with industry professionals and alumni, gaining valuable insights.**

Hobbies

- **Gym, Doodle, Singing (at times)**