Jai Surrya J

E-mail | LinkedIn | GitHub| Leetcode| Portfolio

Summary

I'm a Data Science enthusiast who loves exploring data, building machine learning models, and solving real-world problems using Python. I'm passionate about data preprocessing, feature engineering, predictive modeling, and also enjoy designing and developing scalable software solutions,and deploying end-to-end ML systems.

Education

Shiv Nadar University Chennai, India
B.Tech Computer Science (Internet of Things) || CGPA – 8.36* *Aug 2023 – May 2027*

Relevant Coursework

- Data Structures
- Discrete Mathematics
- OOPs
- Operating Systems
- Software Design
- Database Management
- Computer Networks
- Computer Architecture

Projects

Kindle Review Sentiment Analysis | nltk , sikitlearn, Pandas , Numpy GitHub

- A Natural Language Processing (NLP) project that classifies Amazon Kindle product reviews as positive or negative using text preprocessing and a Logistic Regression classifier. This model achieved an accuracy of 82% on the test dataset.
- Implemented text preprocessing techniques including tokenization, lemmatization, and stopword removal using NLTK, and applied Word2Vec for feature vectorization.
- Performed model evaluation using precision, recall, and F1-score metrics, and visualized insights with Matplotlib and Seaborn.

Churn Prediction | Tensorflow , Keras , Pandas , Scikit learn , Streamlit ,Seaborn GitHub

- Customer Churn Prediction App A web-based interactive tool powered by Streamlit and a Deep Learning (ANN) model that predicts the likelihood of a customer churning from a bank. With an accuracy of 93%, the model leverages user demographic and banking activity to make predictions in real-time.
- Build a deep learning model to predict churn,Use structured customer data including demographic, credit, and account usage details,Create an intuitive and responsive Streamlit web app to allow interactive predictions.

Airbnb Price Prediction System | Pandas,Scikit learn,seaborn,XGBoost GitHub

- This project aims to build a robust machine learning model that predicts Airbnb listing prices based on features like room type, property type, number of accommodates, amenities, and more. The system achieves 98% accuracy on the test set, showcasing strong predictive performance.
- Conducted exploratory data analysis (EDA) to extract insights and visualize feature correlations with price.
- Handled missing values, encoded categorical variables, applied Box-Cox transformation to normalize skewed features.

Technical Skills

Skills: Deep learning , Machine learning , Data visualization , Natural language processing.
Languages: Python, Java, C++.
Tools: Git, VS Code, GitHub.
Libraries: Tensorflow , keras , nltk,Pandas , Numpy , Scikit learn , Seaborn , Matplotlib , Streamlit.
Databases: MongoDB, MySQL.
OS: Linux, Windows.

Certifications

- **Python for Data Science – IBM via Coursera**
  Covered Python programming, data structures, NumPy, pandas, data wrangling, and visualizations using matplotlib/seaborn.
- **Machine Learning – Stanford University (Coursera)**
  Guided by Andrew Ng. Implemented linear regression, SVMs, neural nets, and k-means using Octave.