Aditya Aryan

Fresher

Email Id: aryaditya007@gmail.com

Mobile Number: +91-9631542960

LinkedIn: www.linkedin.com/in/adityaaryan1

GitHub: https://github.com/aditya-aryan123

Skills

Programming languages and Database Tools - Python, SQL(SQL Server), MongoDB.

Statistics - Descriptive Statistics, Inferential Statistics, Probability Distributions, Hypothesis Testing.

Machine Learning: Linear Regression, Logistic Regression, Regularization(Ridge, Lasso), K-Nearest Neighbors, Support Vector Machine, Decision Tree, Extra Tree, Ensemble - Random Forest, Gradient Boosting, Naive Bayes - Gaussian Naive Bayes, Bernoulli Naive Bayes, Multinomial Naive Bayes. Deep Learning - Neural Networks(DNN, RNN, LSTM).

Unsupervised Learning - Hierarchical and K-Means clustering, PCA, Weight of Evidence and Information Value, Imputation.

Data Visualization - Tableau.

Natural Language Processing - Sentiment Analysis, Text Cleaning and Preprocessing, Word2Vec, BERT.

Personal Projects

Project 1: End-to-End Fake News Detection System

GitHub Link - https://github.com/aditya-aryan123/streamlit-fakenewsdetectionsystem-app

1. Extracted and analyzed Fake News to detect spam emails using various Machine Learning models.

2. I used regex and Port Stemmer for data cleaning and stemming and Multinomial Naive Bayes as my final model.

3. Deployed the system using streamlit cloud.

Project 3: Climate Change Analysis and Prediction

GitHub Link - https://github.com/aditya-aryan123/Berkely-Temp-Analysis

1. A monthly mean temperature analysis for the USA was conducted based on data from Berkeley Earth and CRU.

2. Seasonal analysis, and time-series analysis to capture trends and events. Anomaly detection to find extreme trends in the data. Using the Mann-Kendall Test to find the overall trend of temperature.

3. Multiple time-series models were trained and tested, such as ARIMA, SARIMA, and Holt's Model. Facebook's Prophet model was found to be the most accurate in giving predictions.

Project 4: End-to-End Hindi Text Summarization

GitHub Link - https://github.com/aditya-aryan123/text-summarization-flask

1. The dataset for Hindi text was obtained from Kaggle. The dataset was cleaned and preprocessed before tokenizing it.

2. Calculated weighted word frequency and sentence score to generate a summary.

3. Deployed the project using streamlit.

Project 5: Dengue Outbreak Prediction

GitHub Link - https://github.com/aditya-aryan123/Dengue-Outbreak-Prediction

1. Data was gathered from the NOAA website. Analyze the dataset, and apply feature engineering to add broader context to provide a more straightforward perspective on the input data**.**

2. I used matplotlib, seaborne, and plotly to visualize and find specific trends in the dataset.

3. I used a stats model for the final modeling, and found negative binomial regression to be the best fit for the data, with a mean absolute error of 25 and 6 for San Juan and Iquitos respectively.

Education

2016 - 2022

Manipal University Jaipur

Civil Engineering

2014 - 2016

Doon Public School

2012 - 2014

Don Bosco Academy School

ACCOMPLISHMENTS

Machine Hack

GitHub Link - https://github.com/aditya-aryan123/MachineHack2022-Hackathon

Participated in the Analytics Olympiad 2022, a Machine Hack Hackathon. Finished 60th on the public leader board and 106th on the private leader board out of 1006 competitors. Achieved a log loss of 0.6803 by introducing an Extra Tree Classifier model for the Classification task.

Analytics Vidhya

GitHub Link - https://github.com/aditya-aryan123/AnalyticsVidhya-Job-A-Thon-2022

Participated in Analytics Vidhya Job-A-Thon 2022. Finished 142nd on the public leader board out of 6388 competitors. Achieved a root mean squared error of 307 by introducing the LightGBM Regressor.