**Website traffic analysis Project Design and Innovation**

|  |  |
| --- | --- |
| **Date** | **04-10-2023** |
| **Team ID** | **718** |
| **Project Name** | **Website traffic analysis** |

**Table of Contents**

|  |  |
| --- | --- |
| 1 | Introduction |
| 2 | Problem Statement |
| 3 | Design and Innovation Strategies |
| 3.1 | Data Collection and Feature Engineering |
| 3.2 | Data Pre-processing |
| 3.3 | Model Selection and Training |
| 3.4 | Geographic Analysis |
| 3.5 | Market Sentiment Analysis |
| 3.6 | Explainable AI (XAI) |
| 3.7 | Continuous Learning |
| 4 | Conclusion |

**1. Introduction**

Website traffic analysis is the process of studying visitor data and their interactions on a website, including metrics like page views, unique visitors/, traffic sources, bounce rates, and user behaviour. This analysis helps businesses make data-driven decisions to enhance their online presence and user experience. This document will cover the problem statement, the steps to solve it, and the design thinking approach for our project.

**2. Problem Statement**

To analyse website traffic data with the aim of gaining insights into user behavior, identifying popular pages, and understanding traffic sources to assist website owners in enhancing the user experience. The dataset contains various metrics, including page load, average session time, row,day, day.of.week,date,page.Loads,Unique.visits,First.time,returning visits.

**3. Design and Innovation Strategies**

**3.1. Data Collection and Feature Engineering**

Data collection and feature engineering are pivotal in website traffic analysis. The dataset contains five years of daily time series data on website traffic measures, including page loads, unique visitors, first-time visitors, and returning visitors for statforecasting.com, collected using StatCounter from September 14, 2014, to August 19, 2020 .Feature engineering for this dataset may involve creating additional features based on day-of-week and academic calendar seasonality patterns to improve forecasting models for 1-day-ahead, 7-day ahead, and entire-next-week predictions of unique visitors.

**3.2. Data Pre-processing**

Data pre-processing for website traffic analysis involves cleaning, transforming, and organizing data to make it suitable for analysis. The dataset contains daily time series data spanning five years, measuring various website traffic metrics page loads, unique visitors, first-time visitors, returning visitors for an academic teaching notes website, with complex seasonality linked to day-of-the-week and academic calendar, collected via StatCounter.

**3.3. Model Selection and Training**

Select the right machine learning model, prepare data through cleaning and feature selection, train it, evaluate performance, and interpret results for website traffic analysis. Deploy the model for ongoing monitoring, empowering data-driven decisions to enhance website performance and user experience.

Data Preparation : Prepare the dataset by renaming columns to 'ds' for the date and 'y' for the target variable (unique visitors).

Model Initialization : Initialize a Prophet model, considering the complex seasonality patterns mentioned in the dataset description.

Model Training : Fit the Prophet model to the dataset using the fit method.

Forecasting: Generate forecasts for the desired horizon using the make future dataframe and predict methods.

Evaluation: Evaluate the model's performance using appropriate metrics and visualize the forecasted values to assess accuracy. 3.4. Geographic Analysis.

**3.5. Market Sentiment Analysis**

Using the website traffic dataset, a Market Sentiment Analysis can be conducted using the Prophet machine learning model. By leveraging the historical daily traffic data, Prophet can capture the complex seasonality and trends in website visitors, making it a valuable tool for predicting future user engagement. This analysis can help website owners and marketers make data-driven decisions and optimize their strategies to align with the fluctuations in user interest and engagement over time.

**3.6. Explainable AI (XAI)**

Using the provided website traffic analysis dataset, you can apply Explainable AI (XAI) techniques with the Prophet machine learning model to gain valuable insights into the complex seasonality patterns and trends in daily website traffic. Prophet, developed by Facebook, is particularly suitable for time series forecasting tasks like this. It provides interpretable results by decomposing the time series into components such as trend, seasonality, and holidays, making it easier to understand the factors influencing website traffic variations over time. By leveraging Prophet's transparency and the rich dataset, you can develop accurate forecasting models and extract actionable information for making informed decisions related to website traffic optimization and content planning.

**3.7. Continuous Learning**

Continuously learning from this website traffic analysis dataset is essential for adapting to changing online user behaviors and optimizing website performance. By regularly updating and refining forecasting models, businesses can stay ahead of trends, identify growth opportunities, and tailor content strategies to meet user demands. This ongoing learning process allows for timely adjustments, better engagement, and improved decision-making in the dynamic online landscape, ultimately enhancing the user experience and driving sustainable traffic growth.

**4. Conclusion**

Website traffic analysis dataset offers a valuable resource for understanding and forecasting user behavior over five years, considering complex seasonality patterns. By applying machine learning models like Prophet, businesses can extract meaningful insights and adapt their strategies for better engagement and optimization. Continuously learning from this dataset allows organizations to stay competitive and responsive in the ever-evolving digital landscape, ultimately leading to improved user experiences and sustained growth.

