

Website Traffic Analysis

– Phase 2

TEAM MEMBERS :

KAMARAJ	2021115047
KARTHIKA	2021115049
KARTHIKEYAN	2021115050
KAVIYA	2021115051
JEEVESH	2021115312

INTRODUCTION:

In the previous phase, we clearly established our end goals, and we also explained what insights we intend to derive from this project.

In this phase, we will introduce cutting-edge techniques and tools to advance our comprehension of WEBSITE TRAFFIC ANALYSIS. The objective of this phase is to harness the power of machine learning algorithms to forecast service interruptions and extract insights from passenger feedback, ultimately enhancing the overall service quality and user experience.

Below is an in-depth description of the design and algorithms applied in this project.

DATA INTEGRATION:

Data integration is the process of collecting, transforming, and unifying data from various sources, such as web analytics tools, databases, and server logs, to create a coherent and standardized dataset for analysis within IBM Cognos. It involves identifying sources, collecting data, standardizing formats, mapping relationships, and loading data securely while ensuring data quality and synchronization. Effective data integration is fundamental for accurate and reliable insights in a website traffic analysis project, enabling seamless decision-making based on a consolidated and harmonized dataset.

DATA PREPARATION:

Data Preparation is a pivotal step in a website traffic analysis project using IBM Cognos. In this phase, raw data collected from various sources, including web analytics tools, server logs, and databases, undergoes rigorous cleaning, transformation, and structuring. The primary aim is to ensure data quality and consistency. This involves identifying and rectifying data inconsistencies, handling missing values, and standardizing data formats. Furthermore, data preparation may involve filtering out irrelevant or erroneous information to create a clean dataset that is ready for analysis. This process lays the foundation for accurate insights and informed decision-

making, as it ensures that the data is reliable and free from anomalies.

DATA MODELLING:

Data Modeling is the subsequent step in the analysis project, where the prepared data is organized into a structured framework within IBM Cognos Framework Manager. A well-designed data model defines the relationships between tables, establishes key metrics, and provides a foundation for efficient querying and reporting. For website traffic analysis, it's crucial to create a time-based dimension, allowing for trend analysis over specific periods. This data model serves as the backbone for generating reports and dashboards that deliver meaningful insights. Effective data modeling simplifies the complex task of data retrieval and presentation, ensuring that users can efficiently access and understand the relevant metrics and KPIs for website traffic analysis.

DESIGNING REPORT :

Designing reports is a pivotal phase in this project using IBM Cognos, where data insights are transformed into visually engaging and informative presentations. During this process, you select appropriate visualizations, such as tables, charts, and graphs, to represent key performance indicators (KPIs) and metrics. Effective report design is characterized by clarity, relevance, and user-friendliness, ensuring that stakeholders can quickly grasp the most crucial insights. Additionally, thoughtful layout and intuitive navigation

enhance the user experience, making it easier to explore and understand the data. By employing IBM Cognos tools like Cognos Analytics or Report Studio, you can tailor reports to meet specific objectives, providing a dynamic and interactive platform for data-driven decision-making and continuous improvement in the context of website traffic analysis.

ANALYSIS :

In the analysis phase, data is scrutinized to uncover patterns and insights, revealing user behavior and performance metrics. These insights lead to concrete actions, which could involve content optimization, marketing strategy adjustments, or resolving identified issues. This iterative process ensures the website evolves in alignment with organizational goals and user satisfaction.

MACHINE LEARNING MODELS :

Time Series Forecasting Models:

- ARIMA (AutoRegressive Integrated Moving Average): ARIMA models are suitable for predicting time series data, making them ideal for forecasting future website traffic patterns, such as page views or user sessions.
- Exponential Smoothing: Exponential smoothing methods, like Holt-Winters, are effective for capturing seasonality and trends in time series data.

Regression Models:

- Linear Regression: Linear regression can predict numerical outcomes, making it useful for estimating future traffic metrics or user engagement levels.
- Random Forest Regression: Random forest models can handle complex relationships and provide more accurate predictions, especially when dealing with non-linear data.

CONCLUSION :

Here, we have discussed the project design that we will follow and the machine learning models we will utilize in this project