ABSTRACT:

Web servers generate large volumes of raw log data that record every request made to a website. Analyzing these logs can provide valuable insights into traffic patterns and unusual network behavior. This project focuses on parsing and preprocessing raw web server access logs to extract meaningful features such as timestamps, request counts, response sizes, and user information. The unstructured log data is transformed into structured, time-based traffic data for analysis. Machine learning techniques are then applied to predict traffic levels during different time intervals and to detect anomalies in network activity. Additionally, the system distinguishes between bot and human traffic using user-agent information to ensure accurate traffic analysis. This project demonstrates how data preprocessing and machine learning can be effectively used to analyze real-world server logs for traffic monitoring and anomaly detection.

Domain:

Machine Learning, Data Analytics

Keywords:

Web Server Logs, Anomaly Detection, Unsupervised Machine Learning, Traffic Analysis, Feature Engineering, Network Security