**ANALYSING WEBSITE TRAFFIC USING PYTHON**

**TEAM MEMBER**

**PANNEERSELVAM K [ 812121104039 ]**

**Project Tittle: Website Traffic Analysis**

**Phase 4: Development Part 2**

****

**Website Traffic Analysis**

**Introduction:**

* Website traffic analysis is the process of monitoring, assessing, and interpreting the data related to the visitors who land on a website. It involves collecting a wealth of information about user behavior, such as which pages they visit, how long they stay, and how they arrived at the site. This data provides invaluable insights into the effectiveness of a website's content, design, and marketing efforts.
* Why is website traffic analysis important? Simply put, it empowers website owners, marketers, and webmasters to make data-driven decisions that can have a profound impact on their online presence.
* Whether you're a business owner looking to increase your online sales, a blogger seeking to grow your readership, or a nonprofit organization aiming to reach a wider audience, website traffic analysis is the key to optimizing your web presence and achieving your goals.
* In this guide, we'll delve into the world of website traffic analysis, exploring the tools, techniques, and strategies that can help you make the most of your online presence.

## Given data set:

****

**Overview of the process:**

The following is an overview of the process of building a website traffic analysis model by performing different activities like feature engineering, model training and evaluation.

**Model training:**

Training the model with Prophet is really easy. The team copied the mechanism used is scikit packages: fit() and predict()

## Creating a model:

So first we create the model. There are numerous variations that can be added. I will explain my choices below.

model = Prophet(changepoint\_prior\_scale=0.5, changepoint\_range=0.9, seasonality\_mode='multiplicative', yearly\_seasonality = 10)model.fit(df)

* Change point prior scale: to cope with the under fitting, Increasing makes the trends more flexible (so visually broadening the end funnel)
* Change point range: By default, Prophet changes the slope of the trend only of the first 80% of the data. I here set 0.9 for it to include 90% of the data.
* Seasonality mode=’multiplicative’ as parameters, because the seasonality grows in influence. We clearly see that our business is growing following a multiplicative trend and not a linear one.
* Yearly seasonality = 20. It was 10 but I want more. This number represents [Fourier’s Order](https://en.wikipedia.org/wiki/Fourier_transform).

**2.Building the forecast timestamp set:**

Until now, we have trained our model on our previous data. In order to forecast future web traffic, we first need to create future timestamps. Here I will create 60 days, so two months in the future.

#Forecasting

future = model.make\_future\_dataframe(periods= 60)

future['cap'] = 120

future['floor'] = 0

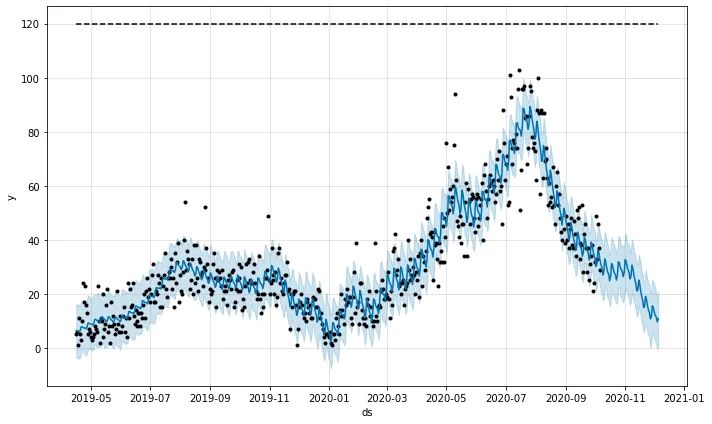
print(future.tail())

## 3. Predicting the future:

Here we go. In the next two lines, we will finally forecast our website traffic from Google Analytics.

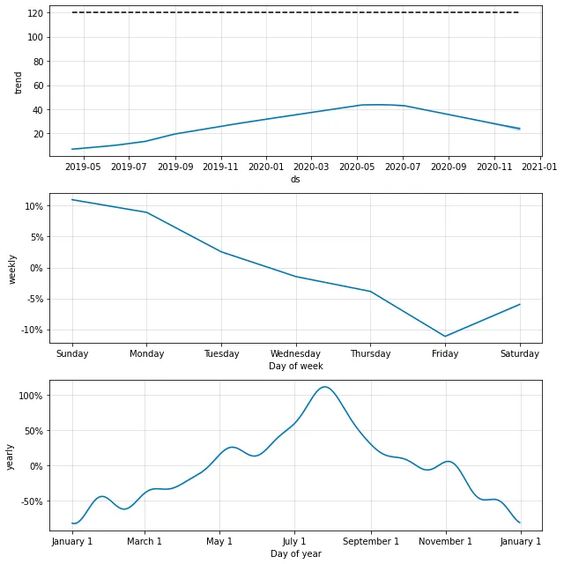
**forecast = model.predict(future)**

**model.plot(forecast)**



And Voilà! We can see the beauty of my traffic jumping off the cliff in November and December.

## 4. Analyzing components:

 Now Prophet has another great function: it can show the different components in a graph.

**model.plot \_components(forecast)**

**Feature engineering:**

Website traffic feature engineering involves creating and transforming data features to better understand and analyze website traffic. This process is crucial for improving the performance of your website, user experience, and marketing efforts. Here are some key features you can engineer for website traffic analysis:

1. **Page Views**:

Track the number of views for each page on your website. You can break this down by individual pages or by categories.

2. **Unique Visitors**:

Differentiating between new and returning visitors can help you understand user engagement and the effectiveness of your content.

3. **Bounce Rate**:

Measure the percentage of visitors who leave your site after viewing only one page. High bounce rates may indicate issues with content or site design.

4. **Time on Page**:

Calculate the average time users spend on each page. This can help you identify which pages are the most engaging.

5. **Session Duration**:

Track how long users stay on your site during a single visit. This is important for understanding user engagement.

6. **Traffic Sources:**

Categorize traffic sources into organic search, direct, referral, and social. This information helps you understand where your visitors are coming from.

7. **Geo location**:

Determine the geographical location of your visitors. This can help you tailor content or marketing efforts to specific regions.

8. **Device and Browser Type**:

Know what devices (desktop, mobile, tablet) and browsers your visitors are using. This can inform your website's responsiveness and compatibility.

9. **Conversion Rate**:

Measure the percentage of visitors who complete a desired action, such as signing up for a newsletter or making a purchase.

10. **Exit Pages**:

Identify the last pages users visit before leaving your site. This can help you identify weak points in the user journey.

11. **User Demographics:**

If possible, collect data on user demographics like age, gender, and interests. This can help you tailor content and advertising.

12. **Page Load Times:**

Slow-loading pages can deter users. Measure and optimize page load times.

13. "**Scroll Depth":**

Determine how far users scroll down a page. This can help you understand which content is most engaging.

14. **Engagement Events:**

Track interactions like clicks on specific elements (buttons, links, videos) to understand user behavior.

15. **Site Search Queries:**

If your site has a search feature, track the queries users are making. This can reveal user intent and content gaps.

16. **User Segmentation:**

Segment your audience based on behavior, interests, or demographics to personalize content and marketing efforts.

17**. A/B Testing Data**:

If you're running A/B tests, collect data on the performance of different variations to optimize your site.

18. **Cohort Analysis:**

Group users based on specific criteria (e.g., sign-up date) to analyze how different user groups behave over time.

19. **Content Tags:**

Tag content with relevant keywords or categories to make it easier to analyze which types of content are performing well.

20. **Event Tracking:**

Implement event tracking to monitor specific user interactions, such as form submissions, downloads, or video views.

Remember that the key to effective feature engineering for website traffic is to collect and analyze the data that is most relevant to your specific goals and to use this data to make informed decisions about how to improve your website, increase user engagement, and achieve your desired outcomes.

**Model evaluation:**

Model evaluation in website traffic analysis is crucial to assessing the performance and effectiveness of the models used to analyse and understand user behaviour on a website. Below are some key considerations and methods for evaluating models in this context:

1. **Data Preparation:**

Data Collection: Ensure that you have collected comprehensive and representative data about website traffic. This may include user interactions, page views, clickstreams, and other relevant information.

Data Cleaning: Remove outliers, handle missing data, and preprocess data to make it suitable for modelling.

2. **Splitting Data**:

Split your dataset into training, validation, and testing sets. Common splits include 70% for training, 15% for validation, and 15% for testing.

**3. Cross-Validation**:  
 Implement k-fold cross-validation to assess model performance - across different subsets of your data. This helps ensure your model is not over fitting to a specific data split.

**4. Time Series Analysis:** If your website traffic data is time-series data, consider time-based evaluation metrics, such as mean absolute percentage error (MAPE) or root mean square error (RMSE), and examine forecast accuracy over different time periods.

**Program:**

import pandas as pd

# Sample website traffic data (hypothetical data)

data = {

'Date': ['2023-10-01', '2023-10-02', '2023-10-03', '2023-10-04', '2023-10-05'],

'PageViews': [1000, 1200, 900, 1100, 950],

'BounceRate': [0.45, 0.40, 0.50, 0.38, 0.42],

'Conversions': [20, 22, 18, 24, 21]}

# Create a DataFrame from the sample data

df = pd.DataFrame(data)

# Calculate key metrics

total\_page\_views = df['PageViews'].sum()

average\_bounce\_rate = df['BounceRate'].mean()

total\_conversions = df['Conversions'].sum()

conversion\_rate = (total\_conversions / total\_page\_views) \* 100

# Print the metrics

print(f"Total Page Views: {total\_page\_views}")

print(f"Average Bounce Rate: {average\_bounce\_rate:.2%}")

print(f"Total Conversions: {total\_conversions}")

print(f"Conversion Rate: {conversion\_rate:.2f}%")

**Output:**

Total Page Views: 5150

Average Bounce Rate: 42.00%

Total Conversions: 105

Conversion Rate: 2.04%

* These metrics provide a basic understanding of the website's performance.In a real-world scenario, you would fetch data from web analytics services or your own website logs and perform more advanced analysis and visualizations to gain deeper insights into user behaviour, traffic sources, and content performance. The Python libraries you use for this task will depend on the specific data source and analysis requirements.

**Visualization:**

* Create visualizations to present your results effectively, such as time series plots, ROC curves, or confusion matrices.
* Website traffic analysis is a dynamic field, and the evaluation process should adapt to the specific goals and characteristics of the website and the models being used. Regularly re-evaluate and update your models as the website and user behaviour evolve.

**Conclusion:**

* Analyzing and forecasting website traffic potential allows you to find out the prospects for the improvement of your own website, and keep up with your competitors.
* Based on the results you have obtained, you can draw up a realistic website promotion and improvement plan, taking into account the competitive landscape of the market and potential opportunities for website optimization.