

Sentiment Analysis on Windows 11 Product Launch Video through Web Scraping



Windows 11

MDA 720 Capstone Project

Project By: Naisarg Bhavsar

Table of Contents

- Background
- Goals of the Project
- Information and link for the API Key
- Link and details of the YouTube Video used to extract the data
- Web Scrapping and Data Extraction
- Data Exploration
- Data Analysis/ Sentiment Analysis
- Data Visualization/ Conclusion
- References

Background

- In this project we will try to extract the comments of the YouTube video of the Product Launch of Windows 11.
- As we are trying do the sentiment analysis based on the web scrapping to know the public opinions and gestures on the launch of windows 11.
- This will help us to know windows 11 fits good or not for our OS Development start-up.

We will do the following steps for the analysis:

- Get Authenticated API for the Web Scrapping
- Data extraction from YouTube video through the web scrapping techniques
- Pre-process that data for the further analysis
- Sentiment Analysis using TextBlob library
- Defining the polarity of the sentiment based on comments using NLTK libraries
- Generating WordCloud based on the results data for each polarity
- Generation of Confusion Metrics for the final result and comparison of all three sentiment polarities

Goals of the Project

- There are various platform available in fast changing tech world like Apple, Microsoft, etc.
- Developers and IT background people are being highly impacted based on their product popularity and trends.
- We aim here to check on the public response on the launch of [Windows 11](#) YouTube video.
- Through this we will be able to predict if the product is good choice for our start-up of OS Development.
- We will use the Web Scrapping and Sentiment Analysis to predict the product launch response.

Information and link for the API Key

- Google Cloud API Key Link: <https://console.cloud.google.com/>
- API Key used for this project is ‘YouTube Data API v3’ which is authenticated and free-of cost for the one month period.

The screenshot shows the 'API/Service Details' page for the YouTube Data API v3. At the top, there's a navigation bar with a back arrow, the text 'API/Service Details', and a 'DISABLE API' button. Below this, there's a large red play button icon next to the service name 'YouTube Data API v3'. A brief description follows: 'The YouTube Data API v3 is an API that provides access to YouTube data, such as videos, playlists, and channels.' It's noted as being 'By Google'. The service details table includes columns for 'Service name' (youtube.googleapis.com), 'Type' (Public API), and 'Status' (Enabled). To the right of the table are three links: 'LEARN MORE', 'TRY IN API EXPLORER', and 'MAINTENANCE & SUPPORT'. At the bottom left, a small note states 'This Photo by Unknown Author is licensed under CC BY-NC-ND'.

Service name	Type	Status
youtube.googleapis.com	Public API	Enabled

[LEARN MORE](#) [TRY IN API EXPLORER](#) [MAINTENANCE & SUPPORT](#)

This Photo by Unknown Author is licensed under [CC BY-NC-ND](#)

Basic details of the Launch Video

- **Upload Date:** Jun 24, 2021
- **YouTube Account:** Windows
- **Subscribers of the account:** 1.08 Million
- **Views:** 10 Million
- **Total Comments:** 40,928
- **Likes:** 382K
- **Dislikes:** Zero
- **Video Title:** Introducing Windows 11
- **New Changes:** Look, Widgets, Snap Layouts, Settings App,
- **Tablet Mode and Docking**

Below is the list of packages used for the authentication and Web Scrapping:

- google-auth-oauthlib (library for Google Authentication with OAuth 2.0 credentials)
- demoji (library for removing emojis from text)
- langdetect (language detection library that detects the language of a text)
- mlxtend (library for machine learning algorithms and tools, including ensemble methods, clustering and feature selection)
- google-api-python-client (library for accessing Google APIs using Python, including YouTube API, Google Drive API, etc.)
- import nltk (libraries for NLTK (National Language Tool Kit) package)
- googleapiclient (Retrieving comments from a YouTube video using the Google API client for YouTube.)

Web Scrapping and Data Extraction

- Here we use YouTube API Key v3 by Google Cloud services for the Web Scrapping.
- This is private and free-trial based key for the 30 Days time frame.
- We extracted the comment of the Product Launch video of Window 11 with the video ID: Uh9643c2P6k
- Here is the link of the Video: [Introducing Windows 11](#)
- Library used for the data extraction: [googleapiclient](#)
- The extracted data than transformed to csv file for sentiment Analysis.

Data Exploration

Then we imported the extracted data into a csv file using the codes below with the result.

Creating the csv file by importing the comments data

```
#importing the data in the list to a csv file for further analysis

import csv

header = ['Comment'] # Change this to the desired header text

with open('comments.csv', 'w', newline='', encoding='utf-8') as file:
    writer = csv.writer(file)
    writer.writerow(header) # Write the header row
    for comment in list1:
        writer.writerow([comment])
```

```
#Reading the csv file to make sure with the comments imported
df = pd.read_csv('comments.csv')
```

```
print(df)

      Comment sentiment
0      This new os really great aesthetic compared 10!  positive
1      compared release - already quite stable system.  neutral
2      I Windows 11 laptop, I started watching master...  positive
3                      So gorg 💙  neutral
4      Honestly, I love everything Windows 11. The an...  positive
...
4155                  ...  ...
4156          Win 11 ❤️ Win 10X  positive
4156  Windows version cant even make taskbar smaller...  neutral
4157                  I LOVE WINDOWS  positive
4158                  We need KDE animations smooth!  positive
4159  windows 11 &lt;3 absolutley<br> love recent u...  positive
[4160 rows x 2 columns]
```

Data Analysis / Sentiment Analysis

- We used here TextBlob library function for the Analysis of the sentiments after the Data Manipulation.
- First, we imported the data into csv file and then made cleaning and preprocessfor the data analysis.
- In this process we check for the null values define the polarity based on the sentiment through specific key words.
- We defined three types of sentiment here as follows:
- Positive, Negative and Neutral
- Then we generated wordcloud for the visualization of our data results.

Data Visualization

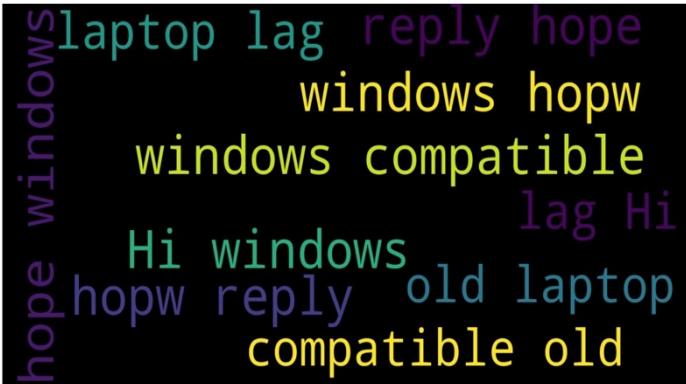
General



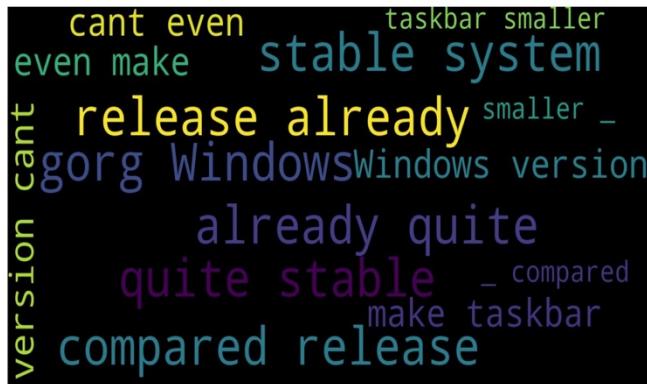
Positive Sentiments



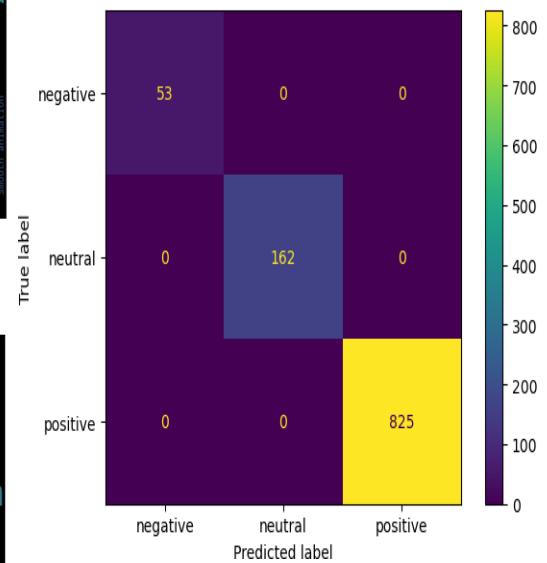
Negative Sentiments



Neutral Sentiments



Confusion Metrics



References and links:

- <https://console.cloud.google.com/welcome?project=innate-empire-384403>
- Video Link: [Introducing Windows 11](#)
- Examples for Data Extraction Codes were taken from the Abhi Goya's blog available on geeksforgeeks:
- <https://www.geeksforgeeks.org/how-to-extract-youtube-comments-using-youtube-api-python/>
- Important Libraries and Packages were referenced from StackOverFlow
- Images for the ppt are taken from Google Images.

