Problem Statement

Analysis of general opinion of people about NFT in Twitter through Natural Language Processing (NLP) method (text mining technique to be more specific)

Text Mining is the technique of exploring large amounts of unstructured text data and analyzing it in order to extract patterns from the text data

I have manually created dataset named “Tweets.csv” which has tweets of six people along with their sentiments: positive, negative, and neutral. The tweets are present in the ‘Text’ column and sentiments in ‘Sentiment’ column

Overall Opinion of People about NFT

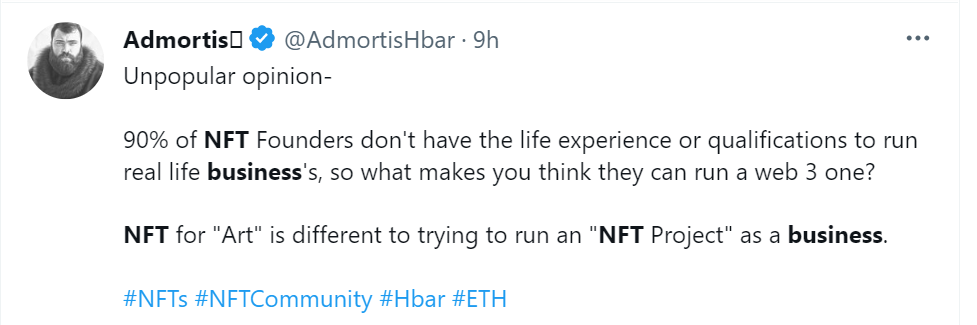
Step 1: Extracting NFT-related text in Twitter

(Unstructured and Noisy Data)

$Unpopular opinion-

90% of NFT Founders don’t have the life experience or qualifications to run real life business’s, so what makes you think they can run a web 3 one?

NFT for “Art” is different to trying to run an “NFT Project” as a business.



@from a business perspective, having these huge conferences back to back like this isn’t smart in my opinion.



Reminder!

As an artist, it is essential that you have the following secured:

1. Your own smart contract
2. Your ENS name
3. Your own website
4. Your Business Email
5. Ledger

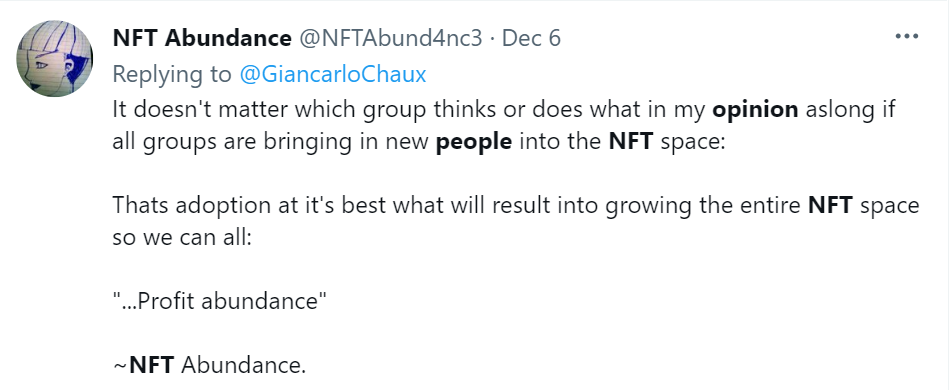


It doesn’t matter which group thinks or does what in my opinion as long if all groups are bringing in new people into the NFT space:

That adoption at it’s best what will result into growing the entire NFT space so we can all:

“…Profit abundance”

~NFT Abundance.



unpopular opinion

web3/nft/crypto people ruined “OG” for me..

-bruh, that was just last year.. that not really ‘og’



Unpopular opinion:

TRON is the best chain for small, medium and large scale NFT projects of all kinds..



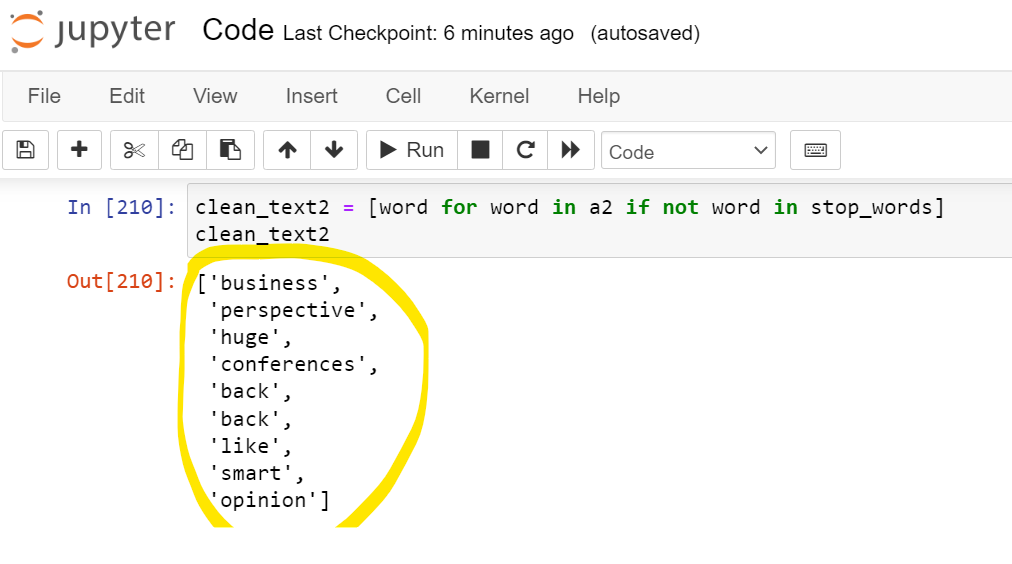
Step 2: Refining the text

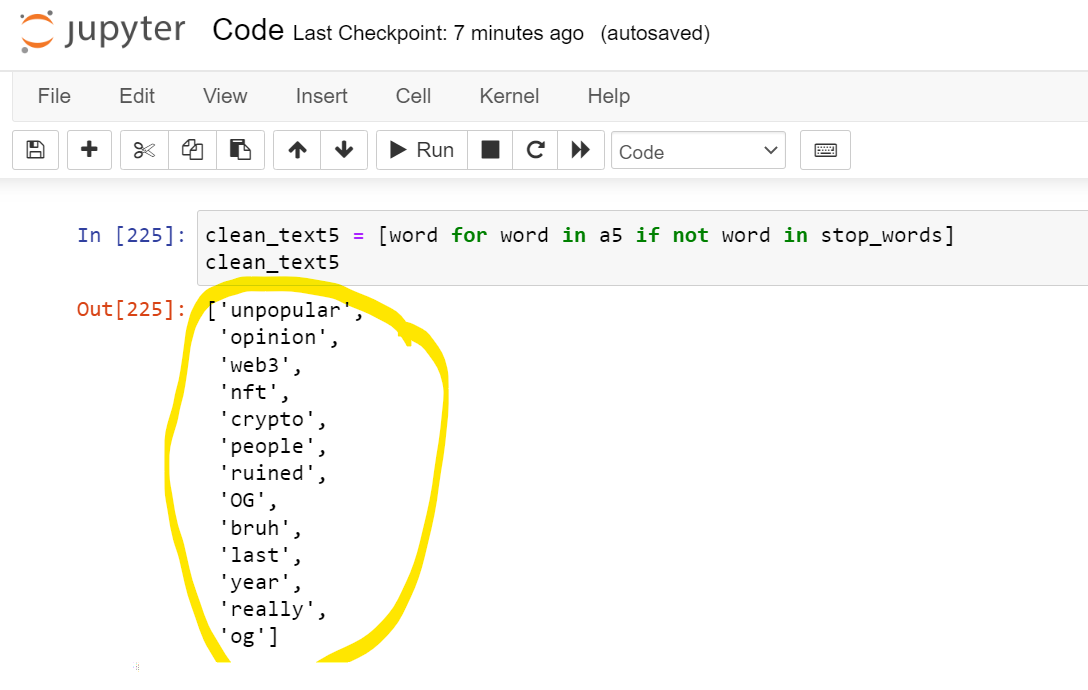
NLP Workflow in text refining:

* Tokenization – removing all punctuations and splitting the text into pieces (tokens)
* Stopword Removal – removes commonly used words like “the”, “and” which are not relevant to analysis
* Stemming and Lemmatization – reduces words to base form to be analyzed as a single item
* Information Retrieval – Extracts relevant information from source

This is how text looks like after Refining

* Structured and Clean Data





Step 3: Word Cloud Before and After Refining

Step 4: Overall Analysis

*Table 1*

|  |  |  |
| --- | --- | --- |
| Number of Words in a Raw Data | Number of Words of Cleaned Data | Improvement Rate% |
| 45 | 25 | 55.55555556 |
| 18 | 9 | 50 |
| 32 | 17 | 53.125 |
| 48 | 22 | 45.83333333 |
| 18 | 13 | 72.22222222 |
| 18 | 12 | 66.66666667 |

As you can see from the Table 1 above the unnecessary and useless data was reduced by 50% minimum and 72.2% maximum and the accuracy was not lost.

*Figure 1*

In the Figure 1 represented above blue columns demonstrate “Number of Words in Raw Data” in Unstructured dataset while orange columns demonstrate reduced “Number of Words of Refined Data”.

As a final result, I can surely say that through text mining technique in Natural Language Processing (NLP) the dataset that was extracted from internet is now more accurate, more understandable and most importantly easy to read.