VizuaMatix

Website: https://www.vizuamatix.com/

NLP Dashboard Guidelines

Implementation of Data Analytical Team

Author: - DA Pandula Pallewatta

TABLE OF CONTENTS

Introduction

What is Natural Language Processing: -

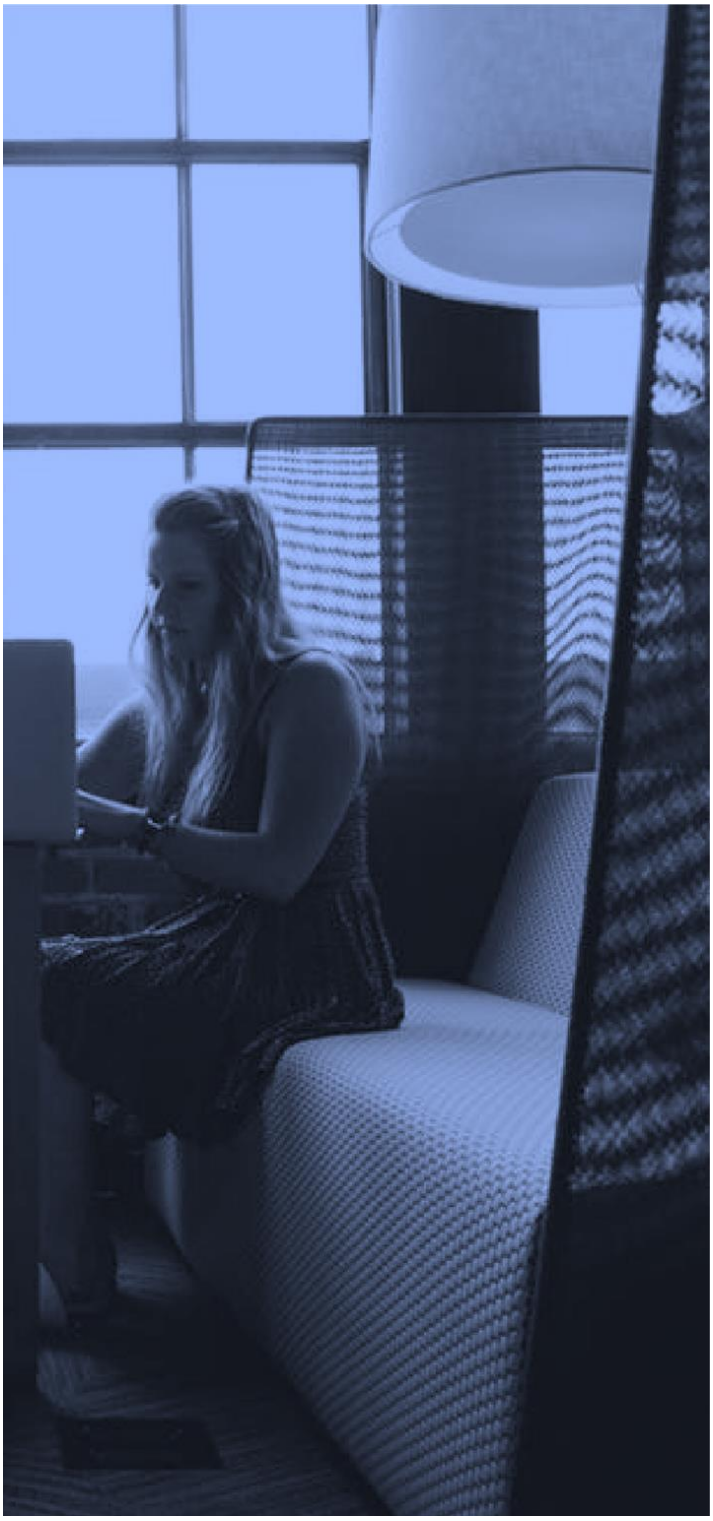
Natural language processing (NLP)
is an artificial intelligence area that aids computers in comprehending, interpreting, and manipulating human language. In order to bridge the gap between human communication and machine comprehension, NLP depends on a variety of fields, including computer science and computational linguistics.

Natural language processing may be used to evaluate massive amounts of text data, such as social media comments, customer service requests, online reviews, news stories, and more, which is one of the key reasons it is so important to organizations.

All of this business data holds a lot of important insights, and **natural language processing (NLP)** can swiftly assist firms in identifying those insights.

NLP Processing Work

Before NLP technologies to understand human language, data scientists must conduct a few basic NLP pre-processing tasks:



Natural Language Processing

- **Tokenization:** breaks down text into smaller semantic units or single clauses
- **Part-of-speech-tagging**: marking up words as nouns, verbs, adjectives, adverbs, pronouns, etc
- **Stemming and lemmatization:** standardizing words by reducing them to their root forms
- **Stop word removal**: filtering out common words that add little or no unique information, for example, prepositions and articles (at, to, a, the).

Natural Language Processing Algorithms

- **Machine learning algorithms.** Machine learning models, on the other hand, are based on statistical methods and learn to perform tasks after being fed examples (training data).

To train machines to create connections between a given input and its related output, machine learning algorithms are fed training data and predicted outputs (tags). Before making predictions for unknown data (new texts), machines utilize statistical analysis methods to create their own "knowledge bank" and determine which qualities best reflect the texts:

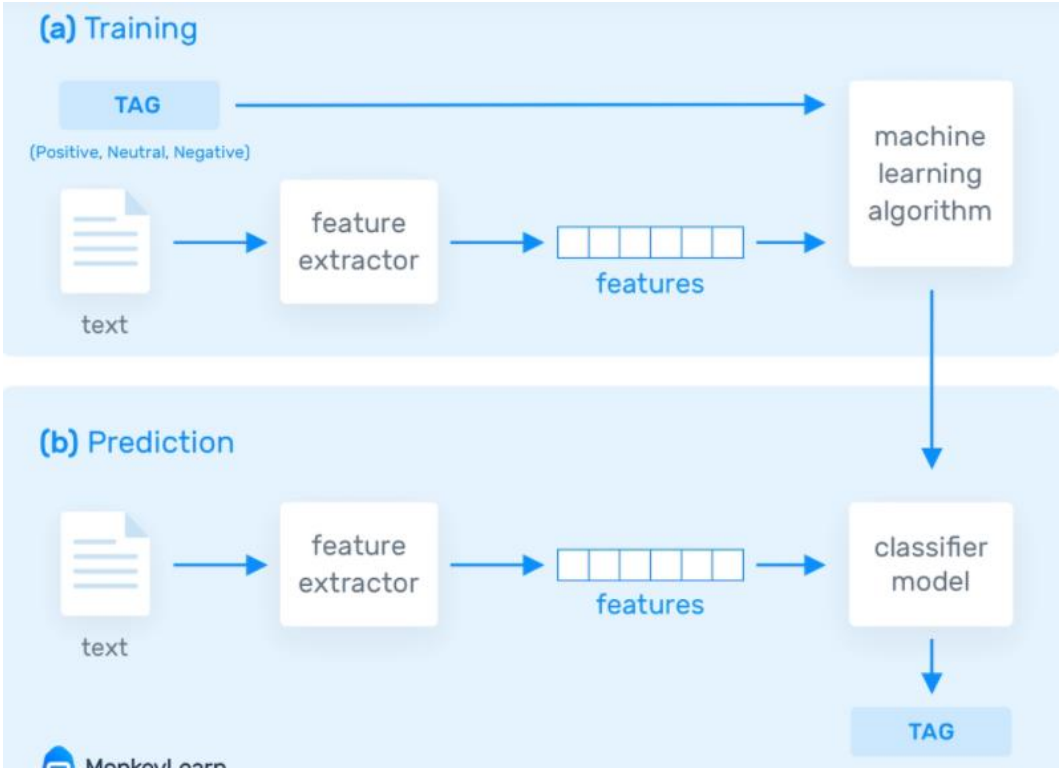


Figure 2: Basic Idea of NLP Machine learning Mechanism

NLP Processing Examples Used in Dashboard

1) Text Classification

Text classification is one of the most fundamental NLP jobs, and it entails categorizing (or tagging) a text based on its content. Models of classification can be used for a variety of objectives, including:

1) Sentiment Analysis

Sentiment analysis is the process of analyzing emotions within a text and classifying them as positive, negative, or neutral. By running sentiment analysis on social media posts, product reviews, NPS surveys, and customer feedback, businesses can gain valuable insights about how customers perceive their brand.Take these Zoom customer and product reviews

2) Topic Classification

Identifying the key themes or subjects inside a document and assigning specified tags is known as topic categorization. You'll need to be familiar with the data you're examining in order to identify appropriate categories while training your topic classifier. If you work for a software firm, for example, you can get a lot of customer support tickets about technical difficulties, usability, and feature requests. In this situation, you may use the tags Bugs, Feature Requests, and UX/IX to organize your tags.

3) Detecting Intent

The objective of intent detection is to figure out what a text's purpose, goal, or intention is. It's a great approach to categorize outbound sales email answers into categories like Interested, Need Information, Unsubscribe, Bounce, and so on. When an email arrives in your inbox, the tag Interested might help you recognize a possible selling opportunity right away!

2) Text Extraction

Text extraction is another type of NLP, which involves extracting certain pieces of data that are already present in a text. It's an excellent technique to quickly summarize material or locate important information. The following are some of the most prevalent extraction models:

1) Keyword Extraction

Keyword extraction extracts the most essential words and expressions from a document automatically. This might give you a taste of the content and its essential points without requiring you to read each section.

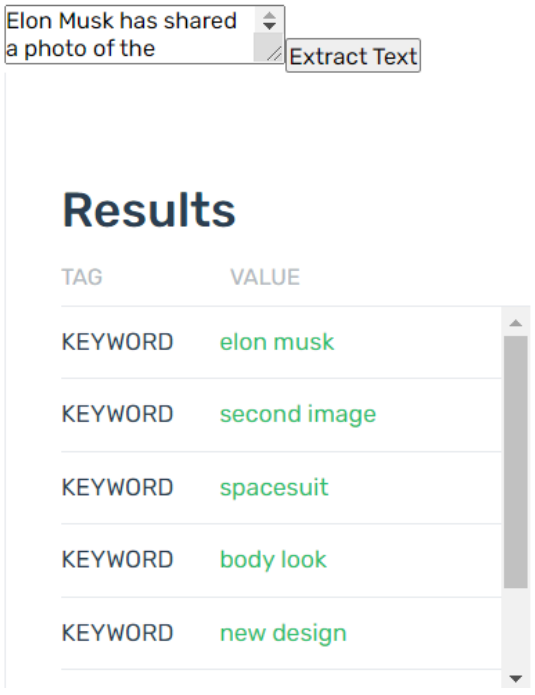


Figure 2: Keyword Identification

2) Named Entity Recognition (NER)

Named Entity Recognition (NER) allows you to extract the names of people, companies, places, etc. from your data.

3) Topic Modeling

Topic modeling and topic categorization are two terms that are often used interchangeably. This example of natural language processing groups texts with similar words and idioms to locate relevant subjects in a document.

It's a fantastic alternative for exploratory research when you don't know your data yet because you don't have to establish a list of preset tags or tag any data.

Dashboard Introduction with Use Cases

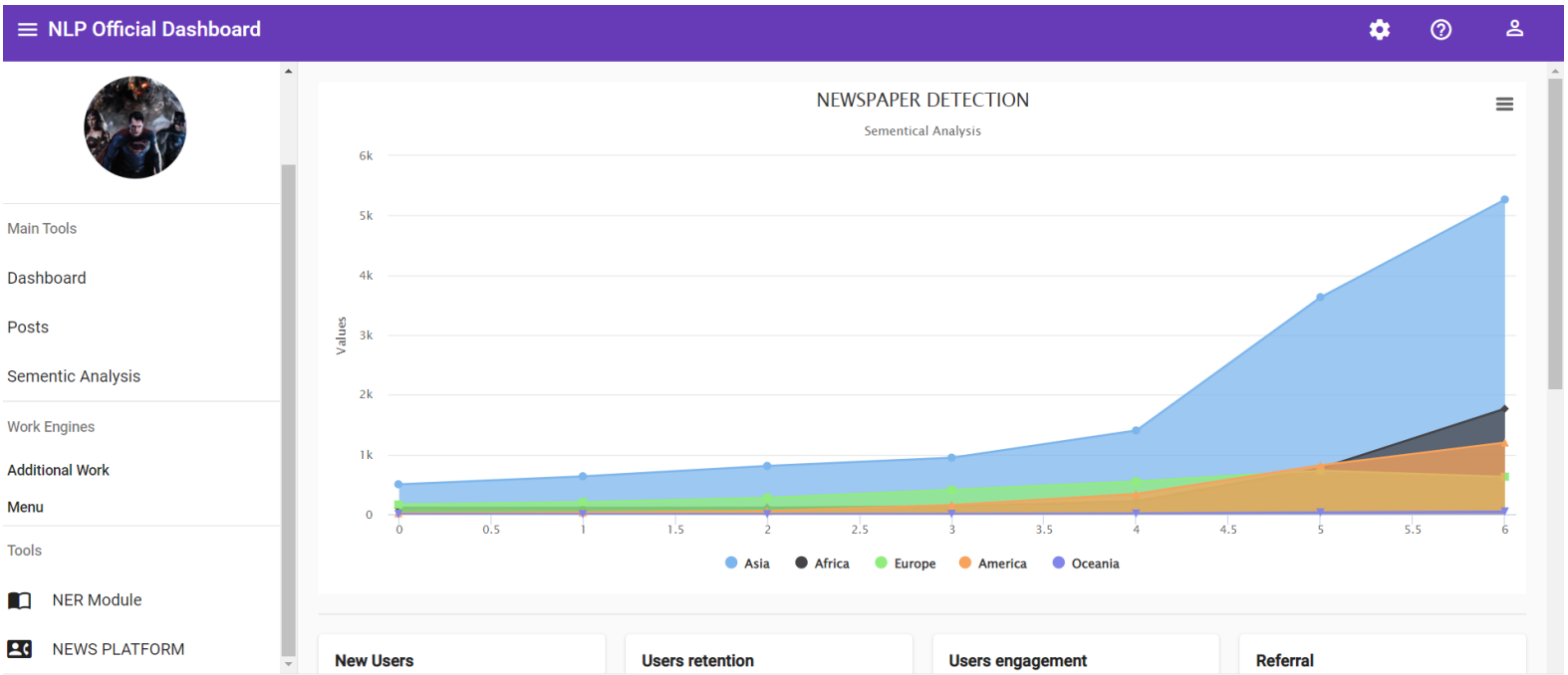


Figure 3: NLP Dashboard Default Page

Prerequisites

Installed Applications

1) **Angular CLI 8.0.X or Higher**
2) **Node Js 14.X or Higher**
3) **Visual Code Up to date Version**

Sidebar Tabs

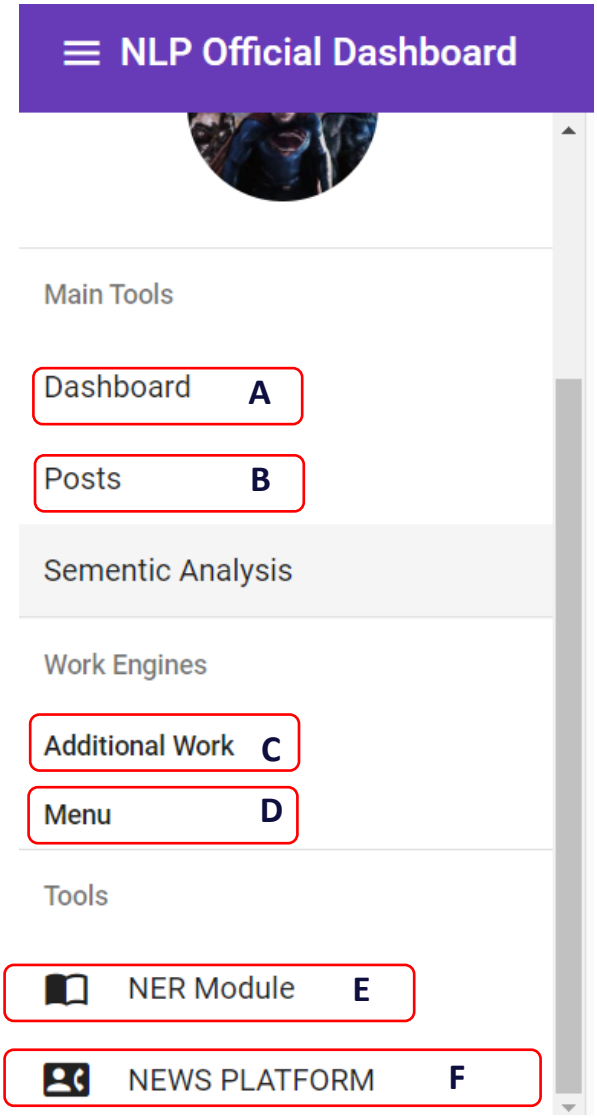


Figure 4 : Dashboard Sidebar

Main Page in NLP Engine Framework

A: Dashboard

This is the root page of the dashboard consisting overview of analyzed newspaper rates, progress up to now & and descriptive conclusion of overall progression.

B: Posts

This page consisting of the main language processing mechanisms in the System.

1) Sentiment Analysis
2) NER Modeling

C: Additional Work

Extra tab created for with sub-menus for future development.

D: Menu

Main 3 Segments of the system included as 3 cards.

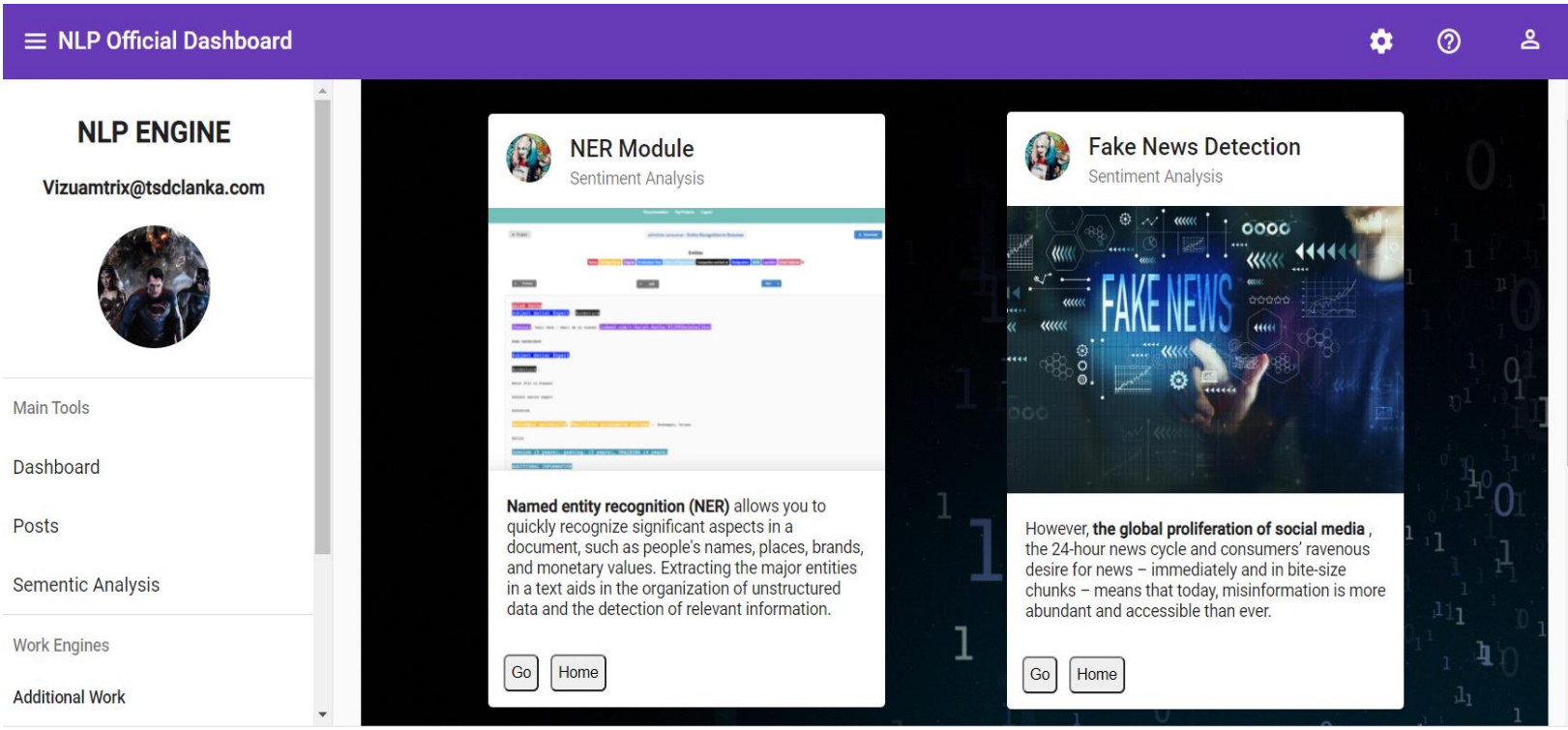
E: NER Module

Consisting of NER modeling resources for needed to do the automation. Currently consists with Stepper, tabs and selective dropdown.

F: News Platform

A main section of the dashboard which provides part of News with Sentiment Analysis results along with it.

1) Posts Page



This page includes main 3 card of the NLP engine which provides each unique service. In each card consists of 2 buttons.

Go: This directs straight to the relevant Page

Home: This direct straight to the Home Page

NER Module Card

In this card consist of the Pathway to NER Page that provides you several mechanisms of NLP functions to work with.

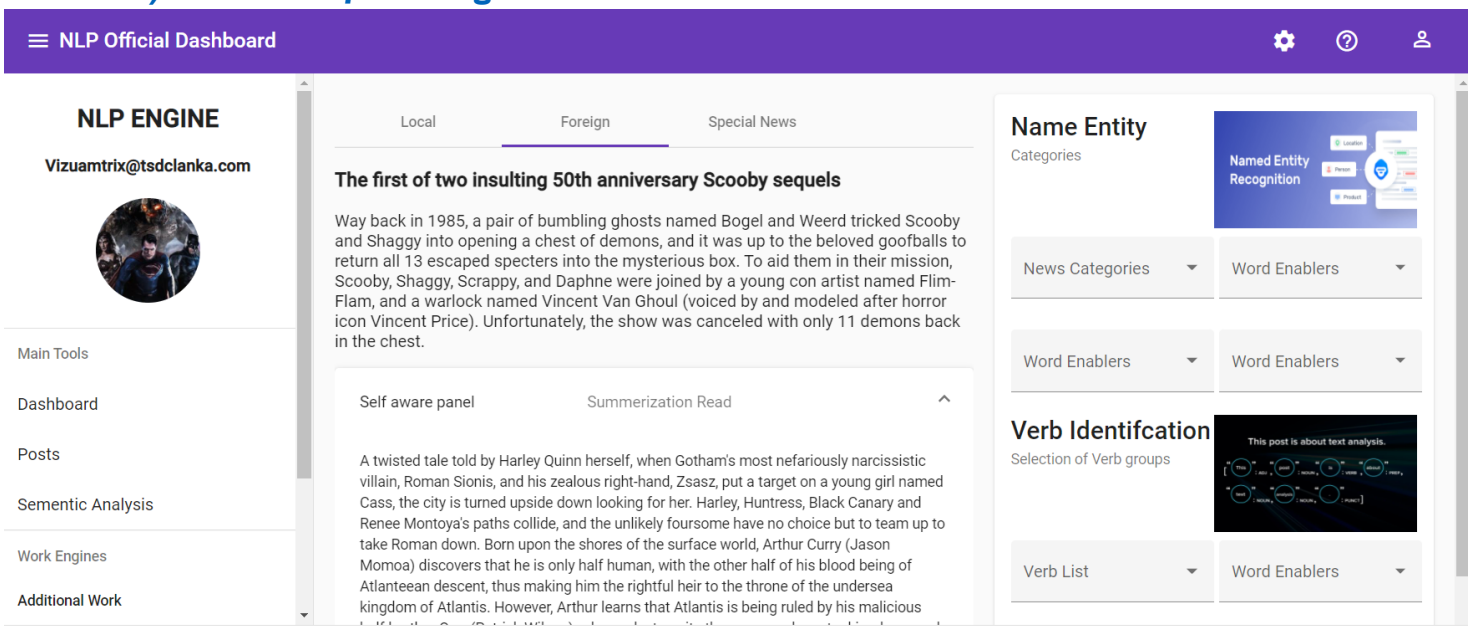
Fake News Detection

In this card directs to the **News Platform Page** that consist of several categorized news with Name Entity Categorization & Verb identification.

Natural Language Processing Card

This card Directs to Analytical Page to train data sets with constraints given by the recipients.

2) News Platform Page



This page showcases the output of the **News captures** that we have analyzed from the dashboard tools. On right side displays 2 sub sectors of the finalized **News Segment**.

A) Name Entity

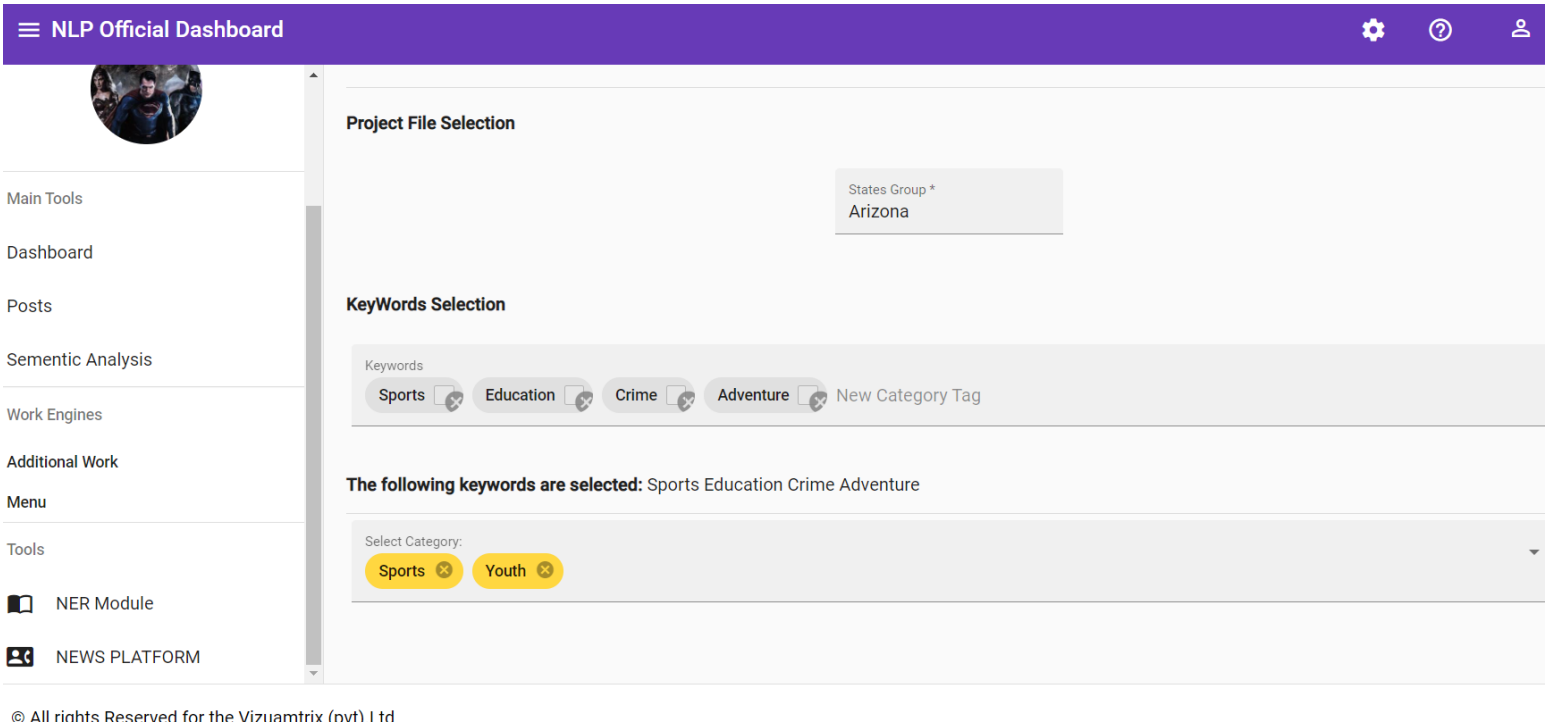
This section provides identified **Name Entities** in each analyzed News segment percentage wise.

B) Verb Identification

This Section identify each news articles frequent verb counts with percentage basis. Also structure based semantic analysis is done.

Also in tabs, every news article headline with summarization will be displayed to provide recipients full analyzation. This will provide a good understanding between authenticity of a news in social media.

3) NER Module Page

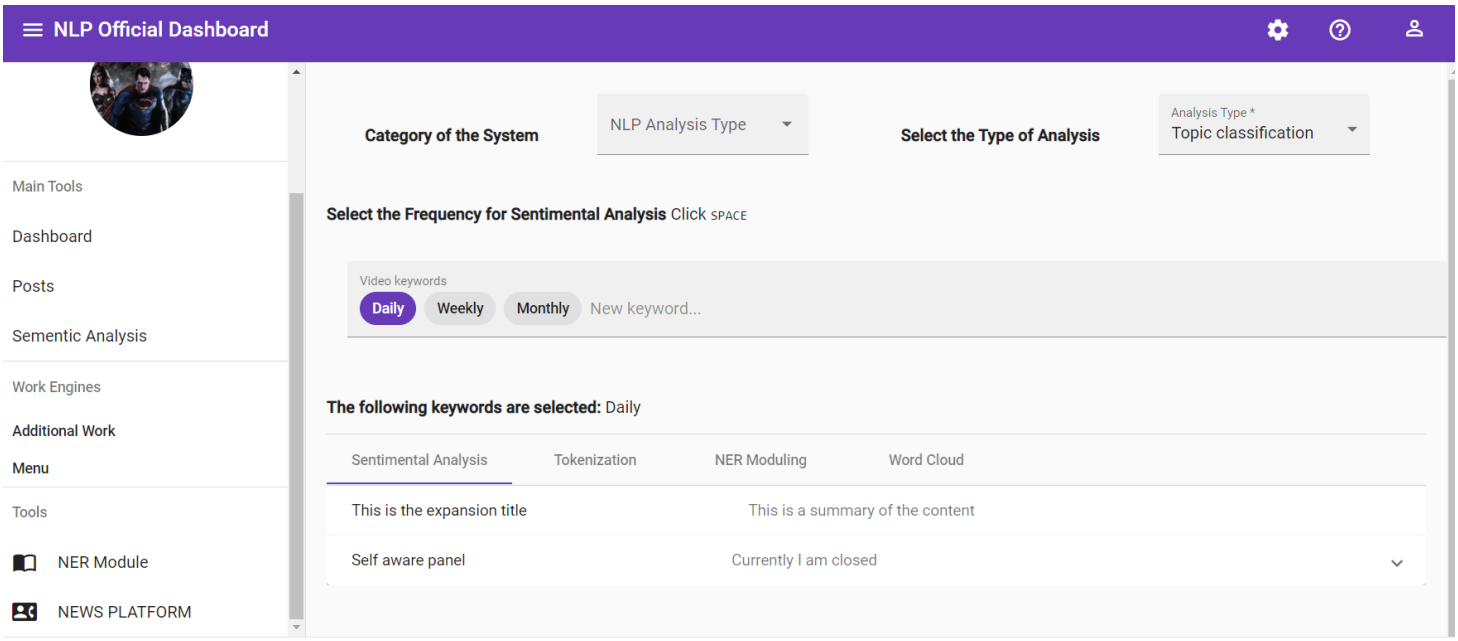


Basically, this page consists of Several Categories select from given csv file and segregate to the given keywords in a news article. Using API integration from News vendor, we can get source files with articles and inputting the data to the Page, can continue the Sentiment Analysis.

```
In [11]: # Process whole documents
text = ("When Sebastian Thrun started working on self-driving cars at "
        "Google in 2007, few people outside of the company took him "
        "seriously. “I can tell you very senior CEOs of major American "
        "car companies would shake my hand and turn away because I wasn’t "
        "worth talking to,” said Thrun, in an interview with Recode earlier "
        "this week.")
doc = nlp(text)

# Analyze syntax
print("Noun phrases:", [chunk.text for chunk in doc.noun_chunks])
print("Verbs:", [token.lemma_ for token in doc if token.pos_ == "VERB"])

Noun phrases: ['Sebastian Thrun', 'self-driving cars', 'Google', 'few people', 'the company', 'him', 'I', 'you', 'very senior C
EOs', 'major American car companies', 'my hand', 'I', 'Thrun', 'an interview', 'Recode']
Verbs: ['start', 'work', 'drive', 'take', 'tell', 'shake', 'turn', 'talk', 'say']

In [12]: # Find named entities, phrases and concepts
for entity in doc.ents:
    print(entity.text, entity.label_)

Sebastian Thrun PERSON
Google ORG
2007 DATE
American NORP
Thrun PERSON
Recode ORG
earlier this week DATE
```

Figure 5: Sentiment Analysis done with Spacy Library

4) Semantic Analysis Page



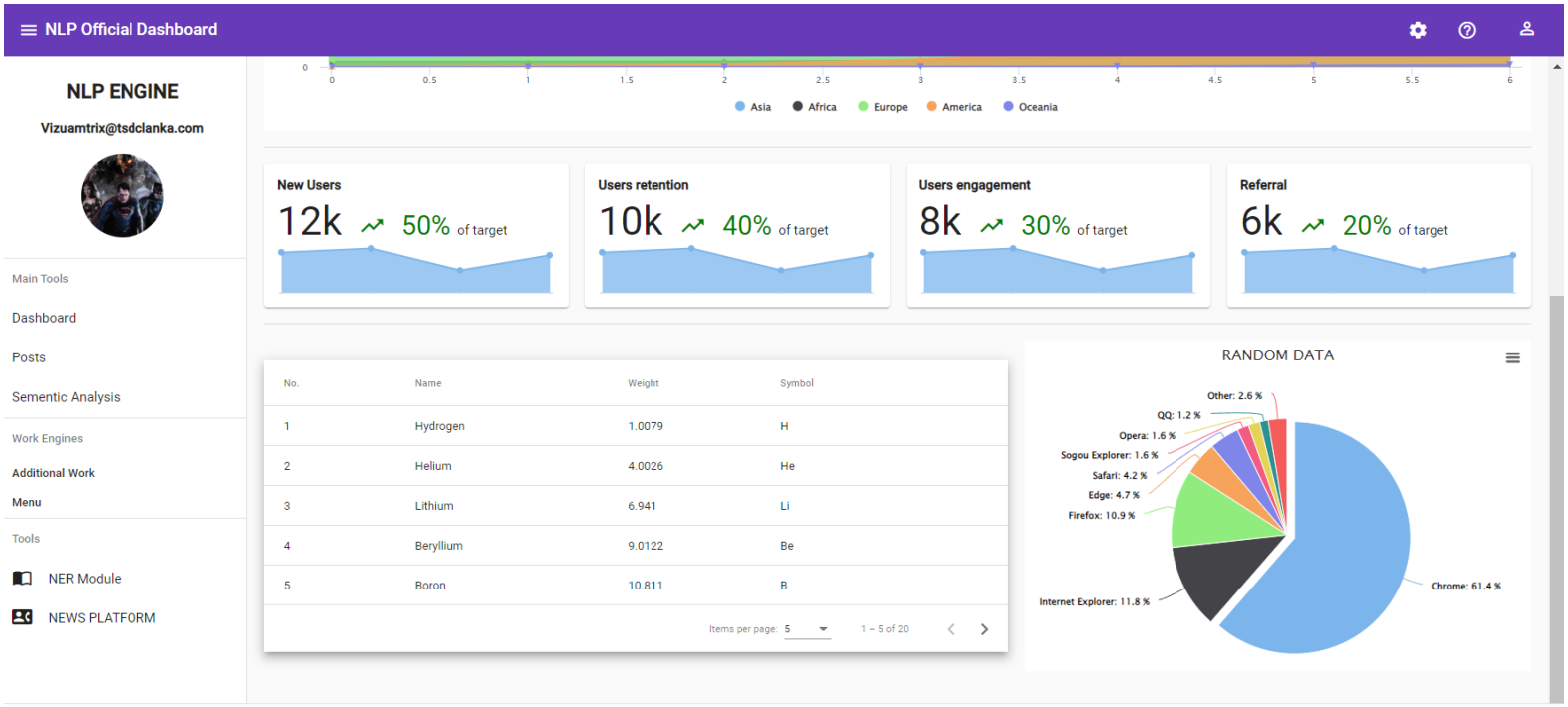
In this page, Semantic Analysis processes will be done here according to some given restrains. Specially,

1) **Topic Classification (Topic Modeling)**
2) **Keyword Extraction**
3) **Word Cloud Creation**

So, these processes mainly done in this page and **Frequency of Sentiment Analysis** can be select according to the user's need.

Also, some Additional tabs like **Sentiment Analysis, Tokenization, NER Modeling and Wordcloud** with additional Angular material in each for future development of the System

5) Home Page



This is the home page of NLP Dashboard consisting progress of the processes done by the system with statistical analysis. Also, some table included with most **frequent keywords** encountered in News articles Analyzation.

Overall Performance be show cased in four cards in the dashboard which is a customer friendly approach for every user.

Also charts and graphs are downloadable for the users in real time updates.

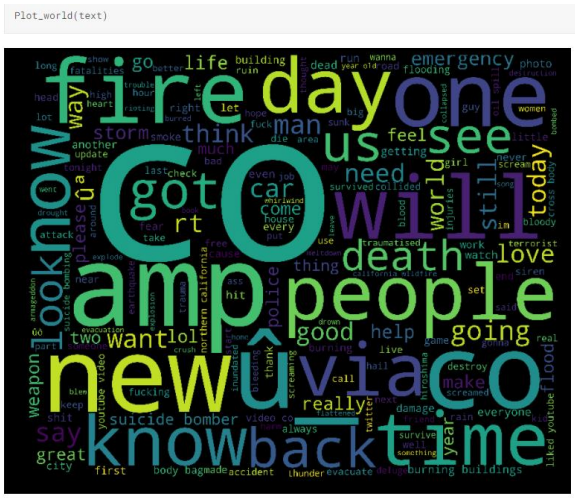


Figure 6: Wordcloud implementation

References

1) **News API Demo. (n.d.). News API Demo** - AYLIEN News API. https://aylien.com/product/news-api/demo. [Accessed 06th February]

2) **Semantic Analysis, Explained.** (2020, August 12). MonkeyLearn Blog. https://monkeylearn.com/blog/semantic-analysis/. [Accessed 09th February]

3) **NLTK :: Natural Language Toolkit. (**n.d.). NLTK :: Natural Language Toolkit. https://www.nltk.org/ [Accessed 28th January]

4) Andrade, F. (2021, June 9). **7 NLP Techniques You Can Easily Implement With Python | By Frank Andrade | Towards Data Science. Medium**. https://towardsdatascience.com/7-nlp-techniques-you-can-easily-implement-with-python-dc0ade1a53c2 [Accessed 24th January]

5) **SpaCy · Industrial-strength Natural Language Processing In Python.** (n.d.). spaCy · Industrial-strength Natural Language Processing in Python. https://spacy.io/ [Accessed 19th January]