NLP FUNCTIONALITY APPLICATION:

A Project Report Submitted

to

MANIPAL ACADEMY OF HIGHER EDUCATION

For Partial Fulfilment of the Requirement for the

Award of the Degree

Of

Bachelor of Technology

in

Computer and Communication Engineering

by

Snigdha Shambhavi Jha, Jinay Patel, Anusha Arra

210953082, 210953236, 210953092

Under the guidance of

Dr. Manjula C. Belavagi
Assistant Professor- Selection Grade
Department of I&CT
Manipal Institute of Technology
Manipal, Karnataka, India

Dr. Chetana Pujari
Associate Professor- Senior Scale
Department of I&CT
Manipal Institute of Technology
Manipal, Karnataka, India

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
A Constituent Unit of MAHE, Manipal

November 2023

INDEX

ABSTRACT

The Natural Language Processing (NLP) Functionality app is a dynamic and user-friendly web-based application built using Streamlit and Python libraries and other APIs. The app integrates a suite of NLP functionalities, including sentiment analysis, keyword extraction, similarity checks, summarization, translation, text prediction, paraphrasing, and spell checking.

The app showcases advanced language processing, enabling tasks such as text prediction and paraphrasing. The integration of cutting-edge models enhances the app's versatility and performance, making it a powerful tool for both novice and experienced users. Incorporating CSS animations for background effects adds a visually engaging element to the user interface.

In summary, the NLP Functionality app bridges the gap between sophisticated NLP capabilities and user accessibility. Its intuitive design and comprehensive feature set make it a valuable resource for individuals seeking to explore, analyse, and gain meaningful insights from textual data using state-of-the-art natural language processing techniques.

INTRODUCTION

The NLP (Natural Language Processing) Functionality app is a web-based application designed to provide users with a user-friendly interface for performing various natural language processing tasks as per their whims and requirements. Developed using Streamlit, a Python library for creating web applications, the app offers a range of functionalities to analyse and manipulate textual data.

The NLP web application features a visually appealing design, making it accessible to users with varying levels of technical expertise.

The application's NLP capabilities and functionalities include tasks such as sentiment analysis, keyword extraction, similarity checks, summarization, translation, text prediction, paraphrasing, and spell checking.

The web application is a consolidated version of the services currently on the market that offer the same functionalities, but separately, and it also serves as a versatile tool for text analysis and manipulation, making NLP accessible to a broader audience.

Overall, the NLP Functionality app combines powerful natural language processing capabilities with a user-friendly interface, making it a valuable resource for individuals looking to interact with and gain insights from textual data.

Sustainable Development Goals:

1. SDG 4: Quality Education- The app can be used as an educational tool to help individuals learn about natural language processing concepts and techniques.

2. SDG 10: Reduced Inequality- The app's user-friendly interface and accessibility can contribute to reducing the inequality in access to and understanding of advanced NLP technologies.

3. SDG 16: Peace, Justice, and Strong Institutions- Natural language processing can be applied in legal contexts for document summarization and analysis, contributing to legal processes.

BACKGROUND INFORMATION

The NLP Functionality app emerged to address a growing need for accessible and versatile natural language processing (NLP) tools in an era where NLP plays a crucial role in various domains, from sentiment analysis to text summarization, and where the demand for user-friendly interfaces to interact with these technologies has become the need of the hour.

The motivation behind the app stems from the complexities of NLP and the desire implement a user-friendly interface to have users interact with the technology irrespective of their level of programming knowledge. As part of the Advanced Programming lab project, the development team recognized the potential impact of creating a platform that not only showcased the capabilities of these NLP models but also made them accessible to a broader audience.

The app provides a seamless interface for users to explore different NLP functionalities, such as sentiment analysis, keyword extraction, text summarization, translation, and more. Through collaborative efforts and a commitment to enhancing programming skills, the development team conceptualized, designed, and implemented the NLP Functionality app. The project not only serves as a practical application of advanced programming concepts but also aligns with the broader goal of fostering innovation and accessibility in the field of natural language processing.

PYTHON CONCEPTS USED

Our Project is a Python programmed graphical user interface using Streamlit to implement our NLP functionality web application, wherein users can input a text and get a variety of language processing option to choose from as per their whims and requirements.

Below are the Python concepts used in the code, along with explanations:

1. Importing Libraries:

- The import statements at the beginning of the code are used to include external Python libraries and modules that provide specific functionalities. Some of the libraries used include:

- streamlit: Used for capturing images from the user's camera.

- textblob: A GUI library for creating the graphical user interface.

- translator (from googletrans): A library for steganography, allowing you to hide text

within images.

- heapq: Used for handling images.

- rake: Provides functions for interacting with the operating system,

such as file and directory operations.

- openai: Utilized for obtaining the user's location based on their

IP address.

- sentencepiece: unsupervised text tokenizer/detokenizer used for text generation systems.

- nltk: to build python programs that work with human language data for application in NLP functions.

- Tokenizer: is in charge of preparing the inputs for a model

- Lsasummarizer: helps in summarization of texts without losing vital information.

- plaintextparser: a module in the Sumy library that's used for parsing plaintext documents

- cosine_similarity: helps measure how similar two text inputs are (used in text similarity function on the website)

2. User Defined Functions:

The code defines several functions, each responsible for a specific

task. These functions include:

- def set_app_style():Function to set a common style for the app.

- def extract_keywords_from_text(text, numOfKeywords): This function implements the NLP functionality of displaying a list of keywords in a given text entered by the user.

- def perform_sentiment_analysis(text): Returns the sentiment analysis, resulting in either a "positive", "negative", or "neutral" sentiment.

- def similarity_check(text1, text2): Gives the level of similarity between two input texts.

- def paraphrase_text(text, api_key): Rephrases a certain input text and provides it as an output.

- def summarise_text(text, api_key): Takes some text as an input, and provides a crisp summary of the same.

- def predict_text(text, api_key): Predicts text based on the input.

- def spell_check(text): Corrects any incorrect spellings in the input text.

- def translate_text(text, target_language): Uses the google translate API to translate to and from any English-based script.

- def main(): Main function.

3. Event Handling:

- The code associates specific functions with GUI buttons, so when a button is clicked, the corresponding function is executed.

5. User Input and Dialog Boxes:

- The website has text boxes for all the functionalities, allowing users to enter text messages for encoding.

- The website also has a menu button at the very left, whose function is used to open a specific functionality and execute it.

6. Displaying Text in a well-designed GUI:

- The Streamlit-based code is designed in such a way that it not only imparts efficient NLP functionalities to the best possible extent, but it also has a sleek and professional graphics interface that makes it very user-friendly and attractive.

7. GUI Elements and Layout:

- The code defines the graphical user interface elements, including

buttons, textboxes, and a menu-driven code with sleek and user-friendly graphics and specifies their layout within the main application window.

METHODOLOGY

As our project, the NLP Functionality App has been made as a platform that aims to provide users with an expansive array of Natural Language Processing (NLP) functionalities and capabilities, finally incorporated into a user-friendly web interface via the usage of Python-based Streamlit.

Each distinct functionality that is carried out within this application addresses nuanced aspects of text analysis, catering to a broad spectrum of user requirements and facilitating a comprehensive understanding of textual data. The functionalities include keyword extraction, summarizer, sentiment analysis, similarity check, translation, spell check and correction, text prediction and paraphrasing of the given text.

IMPLEMENTATION

Starting off the program with the "Set App Style" function, to carry out a sleek and professional graphic user interface, its primary objective is to not merely define a visually appealing aesthetic but to immerse users in a consistent and engaging interface that is easy to navigate. By incorporating a professional-looking background and animated elements, this function ensures an aesthetically pleasing design, setting the stage for an immersive NLP experience.

Next, we have the function, "Extract Keywords from Text" which provides the keyword extraction functionality. To carry out the functionality, the app utilizes the YAKE library. This empowers users with the ability to discern and extract pertinent keywords from the provided text, thereby enhancing their capacity to discern key themes and topics. The user-centric design allows for customization, letting users specify the desired number of keywords for a more tailored and insightful analysis.

Moving on to the "Perform Sentiment Analysis" feature, the TextBlob library is harnessed to delve into the emotional nuances of provided text. By calculating a sentiment score between 1 and 0, and provide the necessary label, users gain deeper insights into the emotional tone conveyed—whether it's positive, negative, or neutral. This nuanced analysis holds immense value for applications ranging from social media sentiment tracking to brand sentiment assessment.

The "Similarity Check" functionality further enriches the app's capabilities by employing TF-IDF vectorization and cosine similarity metrics. This approach quantifies the likeness between two distinct pieces of text, contributing to a nuanced understanding of textual relationships. Users benefit from a more insightful analysis of textual similarities, which further aids in tasks such as plagiarism detection and document matching.

Within the realm of summarization, the "Summarizer" function harnesses the power of OpenAI's GPT-3 engine. This advanced functionality goes beyond mere condensation, generating concise and coherent summaries that encapsulate the essence of the provided text. As a result, users gain a valuable tool for distilling complex information into digestible yet informative summaries.

The global reach of language diversity is addressed through the "Translator" functionality, which is the need of the hour by way of increasing globalization and the moving abroad of students in multiple countries, which, more often than not, pose a language barrier. By leveraging the Googletrans library, users can seamlessly translate text into a specified target language. This functionality significantly broadens the application's utility, making it a versatile tool for users operating in multilingual environments and engaging in diverse language processing tasks.

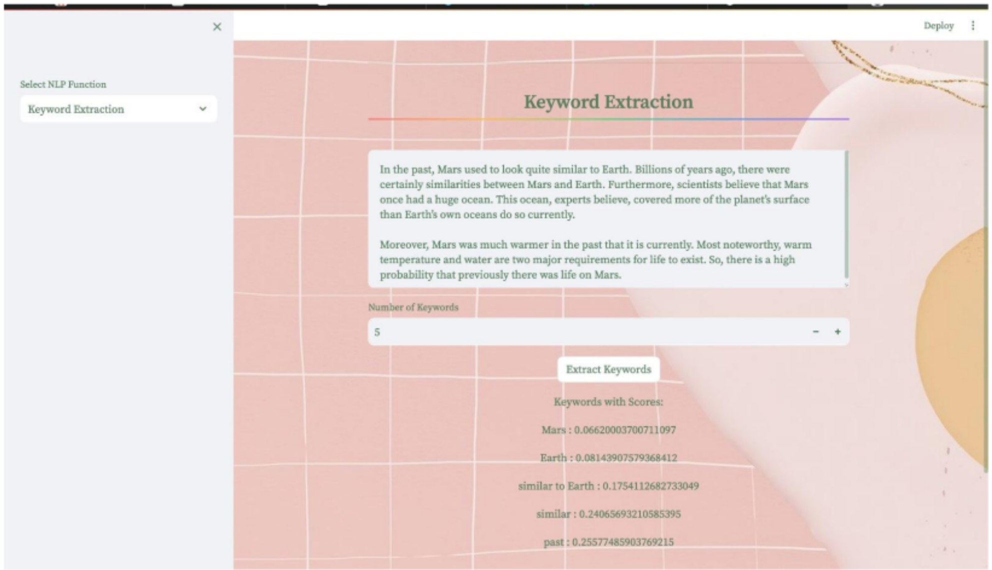
The "Text Prediction" feature extends the app's capabilities by allowing users to generate predictions for the next word in each sentence. Utilizing OpenAI's GPT-3 engine, this function empowers users to complete sentences or generate coherent text. This predictive text generation mechanism is a powerful asset for content creators and those looking to streamline their writing process.

In the realm of linguistic precision, the "Spell Check and Corrector" functionality, driven by the TextBlob library, equips users with the means to identify and rectify spelling errors within a given text. This meticulous spell-checking mechanism ensures the accuracy and readability of text, enhancing overall communication effectiveness.

Lastly, the "Paraphrasing" function integrates the capabilities of OpenAI's GPT-3 engine to rephrase a given text while preserving its original meaning. This advanced paraphrasing tool empowers users to effortlessly generate alternative versions of text, enhancing content versatility for applications such as content rewriting and diversifying textual outputs.

The orchestration of these diverse functionalities into a cohesive and user-friendly Streamlit-based web application signifies a holistic approach. This integration ensures that users, irrespective of their proficiency in NLP, can access and leverage these powerful tools for tasks ranging from sentiment analysis to language translation and beyond.

RESULTS



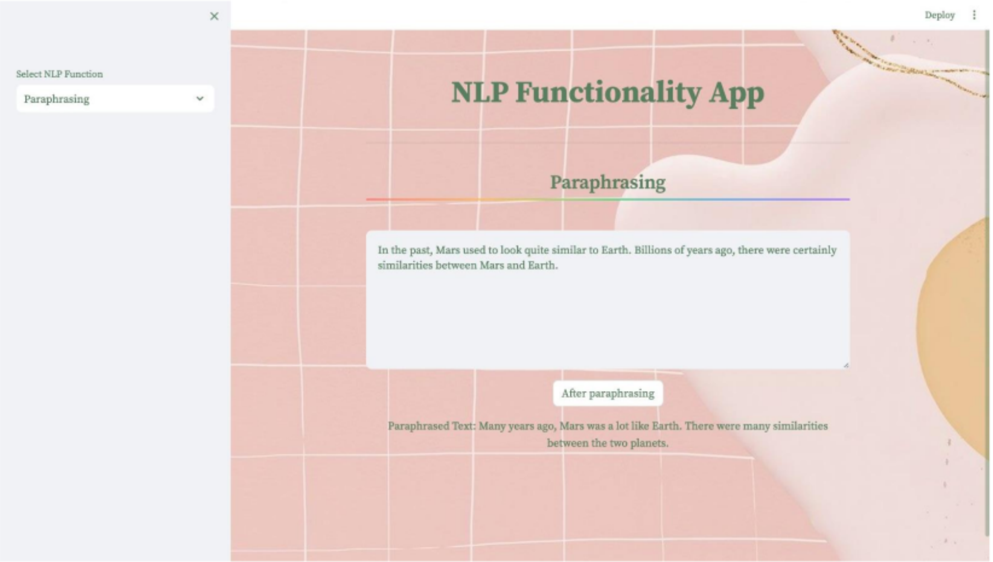
Keyword Extraction

In the past, Mars used to look quite similar to Earth. Billions of years ago, there were certainly similarities between Mars and Earth. Furthermore, scientists believe that Mars once had a huge ocean. This ocean, experts believe, covered more of the planet's surface than Earth's own oceans do so currently.

Moreover, Mars was much warmer in the past that it is currently. Most noteworthy, warm temperature and water are two major requirements for life to exist. So, there is a high probability that previously there was life on Mars.

Number of Keywords: 5

```
Keywords with Scores:

Mars : 0.06620003700711097

Earth : 0.08143907579368412

similar to Earth : 0.1754112682733049

similar : 0.24065693210585395

past : 0.25577485903769215
```

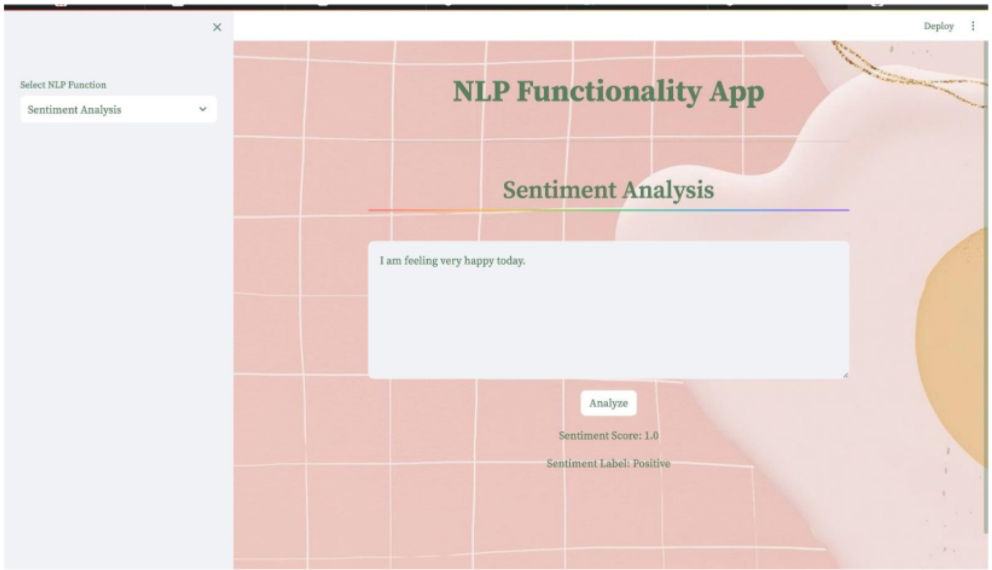


NLP Functionality App

Paraphrasing

In the past, Mars used to look quite similar to Earth. Billions of years ago, there were certainly similarities between Mars and Earth.

Paraphrased Text: Many years ago, Mars was a lot like Earth. There were many similarities between the two planets.



Select NLP Function: Sentiment Analysis

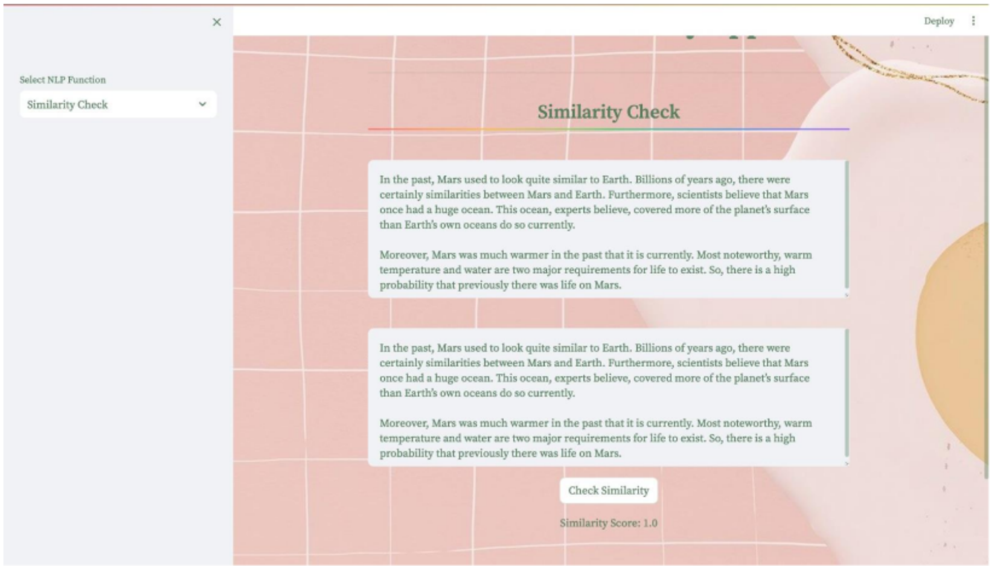
NLP Functionality App

Sentiment Analysis

I am feeling very happy today.

Sentiment Score: 1.0

Sentiment Label: Positive



Select NLP Function: Similarity Check

Similarity Check

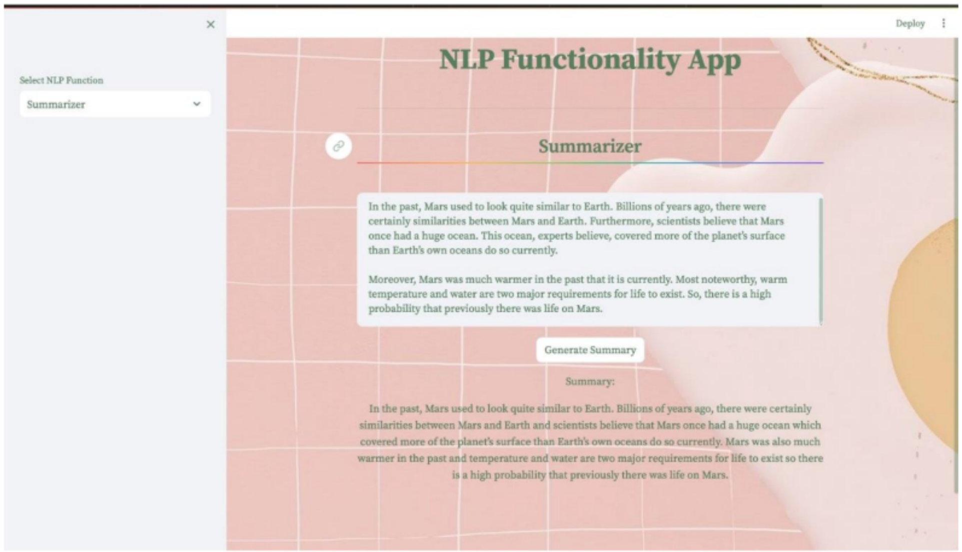
In the past, Mars used to look quite similar to Earth. Billions of years ago, there were certainly similarities between Mars and Earth. Furthermore, scientists believe that Mars once had a huge ocean. This ocean, experts believe, covered more of the planet's surface than Earth's own oceans do so currently.

Moreover, Mars was much warmer in the past that it is currently. Most noteworthy, warm temperature and water are two major requirements for life to exist. So, there is a high probability that previously there was life on Mars.

In the past, Mars used to look quite similar to Earth. Billions of years ago, there were certainly similarities between Mars and Earth. Furthermore, scientists believe that Mars once had a huge ocean. This ocean, experts believe, covered more of the planet's surface than Earth's own oceans do so currently.

Moreover, Mars was much warmer in the past that it is currently. Most noteworthy, warm temperature and water are two major requirements for life to exist. So, there is a high probability that previously there was life on Mars.

Similarity Score: 1.0

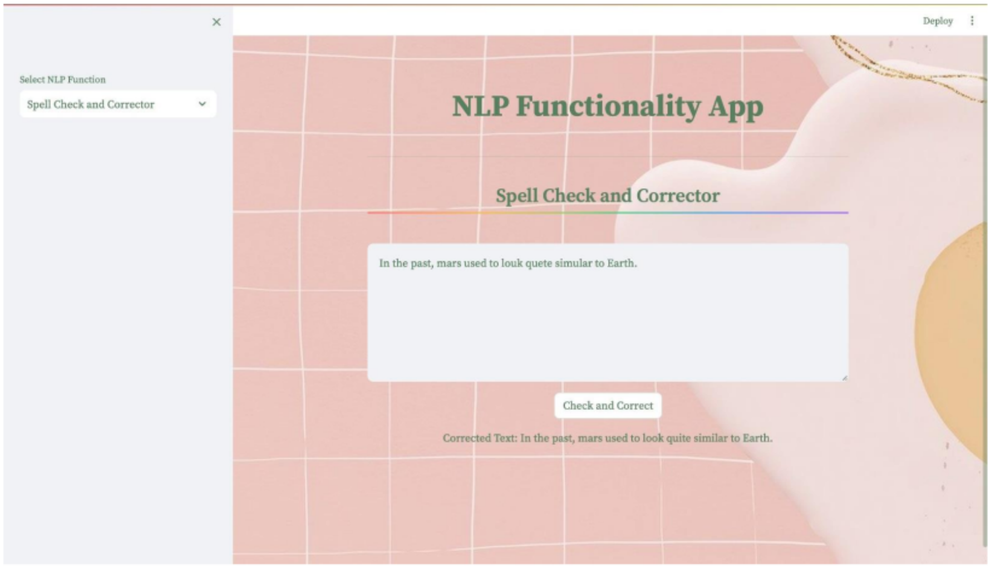


In the past, Mars used to look quite similar to Earth. Billions of years ago, there were certainly similarities between Mars and Earth. Furthermore, scientists believe that Mars once had a huge ocean. This ocean, experts believe, covered more of the planet's surface than Earth's own oceans do so currently.

Moreover, Mars was much warmer in the past that it is currently. Most noteworthy, warm temperature and water are two major requirements for life to exist. So, there is a high probability that previously there was life on Mars.

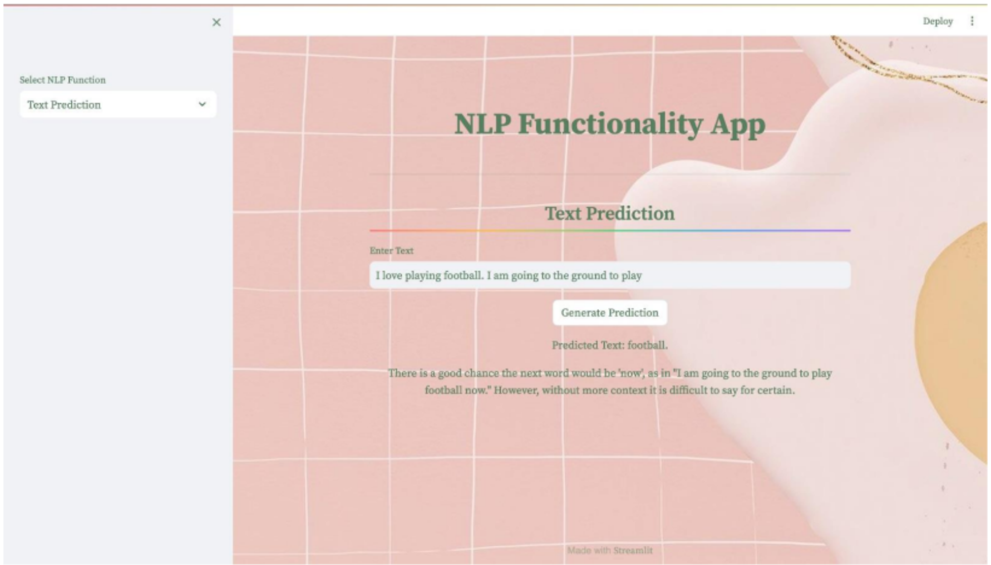
Summary:

In the past, Mars used to look quite similar to Earth. Billions of years ago, there were certainly similarities between Mars and Earth and scientists believe that Mars once had a huge ocean which covered more of the planet's surface than Earth's own oceans do so currently. Mars was also much warmer in the past and temperature and water are two major requirements for life to exist so there is a high probability that previously there was life on Mars.



In the past, mars used to louk quete simular to Earth.

Corrected Text: In the past, mars used to look quite similar to Earth.



Select NLP Function

Text Prediction

NLP Functionality App

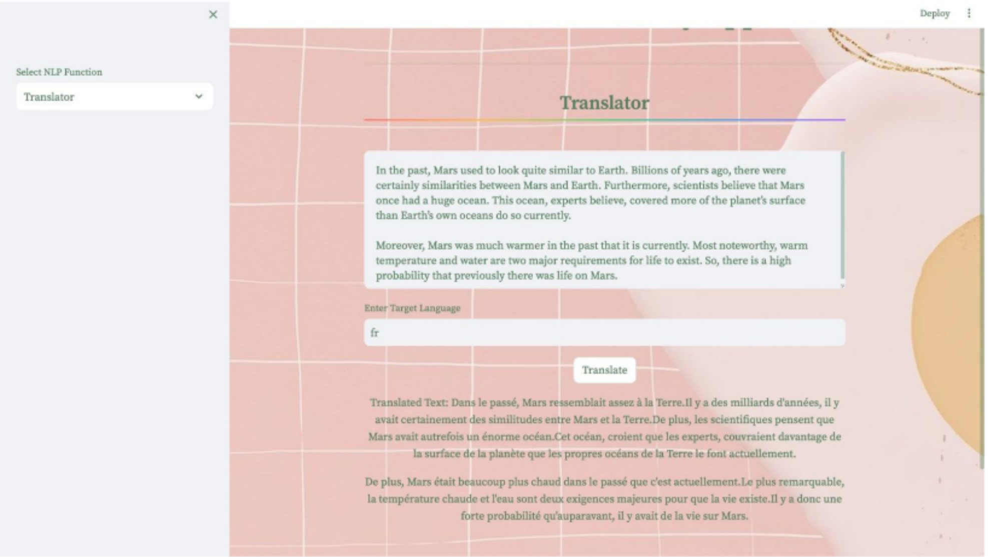
Text Prediction

Enter Text

I love playing football. I am going to the ground to play

Predicted Text: football.

There is a good chance the next word would be 'now', as in "I am going to the ground to play football now." However, without more context it is difficult to say for certain.



Select NLP Function

Translator

Translator

In the past, Mars used to look quite similar to Earth. Billions of years ago, there were certainly similarities between Mars and Earth. Furthermore, scientists believe that Mars once had a huge ocean. This ocean, experts believe, covered more of the planet's surface than Earth's own oceans do so currently.

Moreover, Mars was much warmer in the past that it is currently. Most noteworthy, warm temperature and water are two major requirements for life to exist. So, there is a high probability that previously there was life on Mars.

Enter Target Language

fr

Translated Text: Dans le passé, Mars ressemblait assez à la Terre.Il y a des milliards d'années, il y avait certainement des similitudes entre Mars et la Terre.De plus, les scientifiques pensent que Mars avait autrefois un énorme océan.Cet océan, croient que les experts, couvraient davantage de la surface de la planète que les propres océans de la Terre le font actuellement.

De plus, Mars était beaucoup plus chaud dans le passé que c'est actuellement.Le plus remarquable, la température chaude et l'eau sont deux exigences majeures pour que la vie existe.Il y a donc une forte probabilité qu'auparavant, il y avait de la vie sur Mars.

CONCLUSION

In a nutshell, the NLP Functionality App is like your all-in-one wizard for playing with writings and implementing text functionalities. Multiple functionalities have been implemented, everything from figuring out if your words sound happy, sad, or neutral (sentiment analysis), to picking out the most important keywords in a long and boring text. And when it comes to spinning words into something new entirely, that’s also something that can be done.

Irrespective of the user being a professional in the field or just new to the the NLP scenes entirely, this app handles it all; and breaks down fancy language stuff into easy-to-use tools. But it's not just about playing around with words, because this app also offers to help you understand sentiments, create snappy summaries, and even fix any typos in an input text.

And as tech keeps pushing the boundaries, this app grows with it.

REFERENCES

https://www.nltk.org/

https://spacy.io/

https://nlp.stanford.edu/

http://www.nltk.org/book/

https://www.aclweb.org/anthology/

https://www.kaggle.com/competitions?sortBy=prize&group=general&page=1&pageSize=20&category=natural-language-processing

NLP Functionality Application

ORIGINALITY REPORT

4%	4%	1%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	shreyarajagopal13.files.wordpress.com Internet Source	1%
2	impressions.manipal.edu Internet Source	1%
3	doaj.org Internet Source	1%
4	watersolutions.us Internet Source	<1%
5	dokumen.pub Internet Source	<1%
6	en.ird.fr Internet Source	<1%
7	www.coursehero.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 3 words
Exclude bibliography	On		