Sentiment Analysis - Twitter

DISSANAYAKE MUDIYANSELAGE NELUM CHATHURANI
- DATA SCIENCE / DATA ANALYTICS INTERN -

Table of Contents

1. Introduction

This report presents the results of a sentiment analysis conducted on a specific tweet using a pre-trained transformer model, specifically the **Cardiff NLP Twitter RoBERTa Base Sentiment Model**. The goal was to assess the sentiment expressed in the tweet.

2. Methodology

Libraries Used

1. Transformers: For loading the pre-trained RoBERTa model and tokenizer.
2. Scipy: For applying the softmax function to convert model outputs into probabilities.
3. Emoji: For handling emoji representation in the text.

Steps

- Tweet Preparation
- Tokenization

  The processed tweet was tokenized using the RoBERTa tokenizer, which converts the text into a format suitable for model input.
- Model Loading

  The Cardiff NLP Twitter RoBERTa Base Sentiment Model was loaded for sentiment classification.
- Sentiment Analysis

  The tokenized input was passed through the model to obtain sentiment scores.
- Score Normalization

  The raw model outputs were normalized using the softmax function to obtain probabilities for each sentiment label.

3. Results

Sentiment Scores

The sentiment analysis produced the following scores:

Sentiment Label	Score
Negative	0.0014318319
Neutral	0.018191433
Positive	0.9803767

Interpretation

- **Negative Sentiment**: This score indicates how likely the tweet expresses negative sentiment.
- **Neutral Sentiment**: This score reflects how neutral the sentiment is perceived.
- **Positive Sentiment**: This score shows the likelihood of a positive sentiment.

4. Conclusion

The sentiment analysis indicates that the tweet expresses positive sentiment, as suggested by the highest score in the analysis. The inclusion of a happy emoji further emphasizes a positive tone.