(AI4001/CS4063 – NLP) Report: Assignment 3

From: Muhammad Faheem

To: Ma’am Laraib Afzal

Table of Contents

## Overview

## Implementation Process

## Challenges Faced

## Performance of the Model

## Insights

## Conclusion

# Overview

This assignment focused on performing sentiment classification and temporal sentiment trend analysis on the Daily Routine Diaries dataset. The objective was to classify diary entries as Positive, Negative, or Neutral and to analyze how sentiments change over time to identify periods of stress (e.g., exams, deadlines) and happiness (e.g., holidays, special events).

# Implementation Process

The solution was developed in a modular fashion, with key components including:

* **Utility Functions:** Implemented helper functions to label sentiment using a lexicon-based approach and compute numeric sentiment scores.
* **Implementation Functions:** Developed functions to process diary entries, compute moving averages for trend analysis, and generate visualizations using Matplotlib.
* **Data Loader:** Utilized the provided data loader script to extract diary entries from the dataset.
* **Keyword Expansion:** Expanded the keyword database to better capture the nuances of Roman Urdu diary entries.
* **Visualization Enhancements:** Added a specialized plot to highlight stress periods (in red) and happy periods (in green), in addition to the overall sentiment trends.

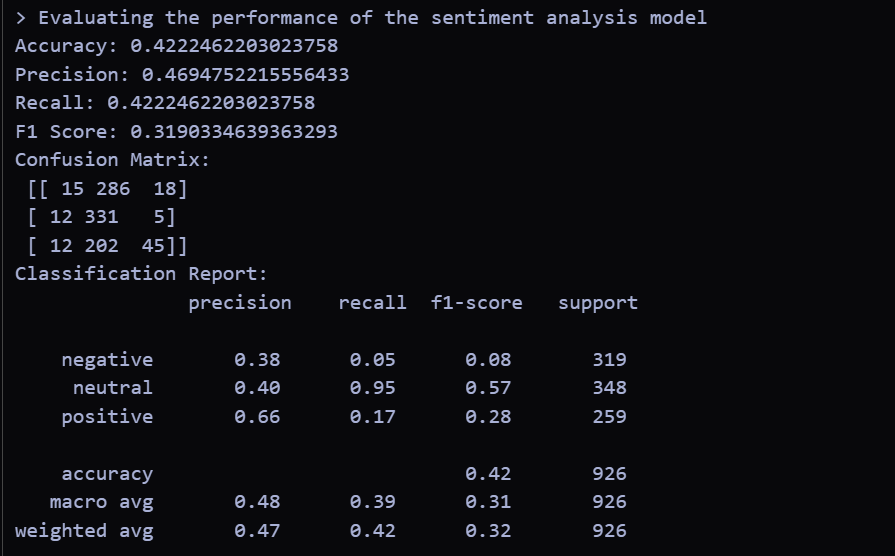
# Challenges Faced

* **Limited Labeled Data:** The absence of a fully annotated dataset required the use of a lexicon-based approach which, while effective, may not capture all nuances of sentiment.
* **Language Nuances:** Roman Urdu expressions are highly colloquial and context-dependent, making it challenging to design a keyword list that comprehensively covers all sentiment expressions.
* **Trend Validation:** Establishing concrete evaluation criteria for temporal trends was difficult because trends are subjective and require external validation, such as matching academic or social calendars.
* **Visualization Alignment:** Aligning the moving average with actual diary entry dates needed fine-tuning to accurately reflect temporal sentiment changes.

# Performance of the Model

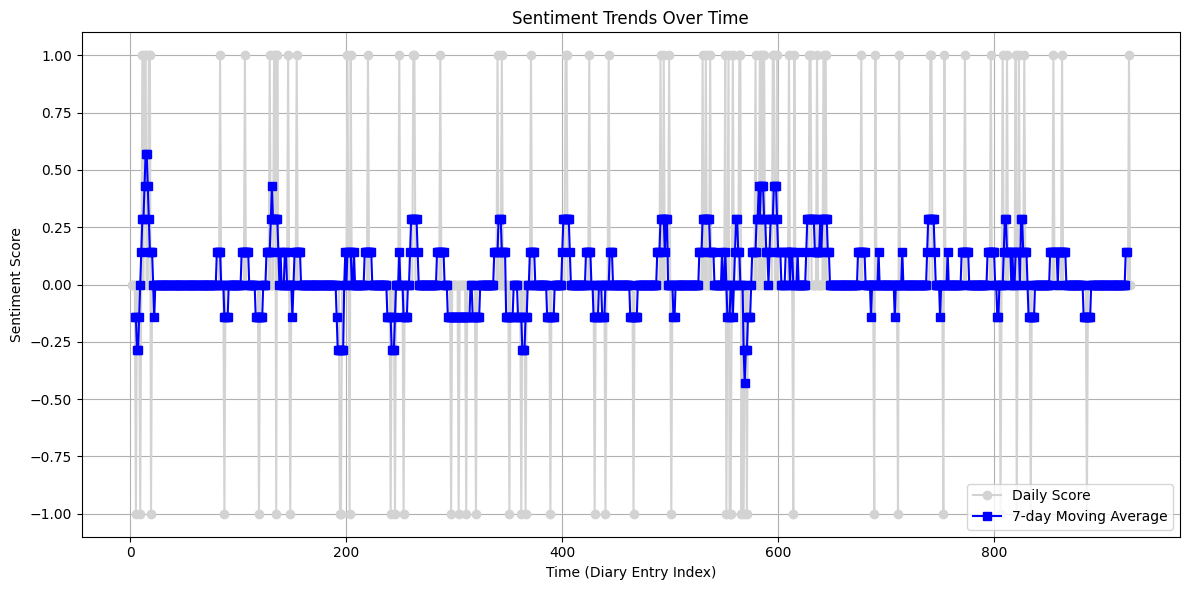
The model’s performance can be evaluated using both quantitative and qualitative methods:

* **Quantitative:** If a manually labeled subset is available, metrics such as accuracy, precision, recall, and F1-score can be computed. The sample model across which the ground-truth labels were generated and compared with the custom developed model is: cardiffnlp/twitter-xlm-roberta-base-sentiment.
* **Qualitative:** Expert evaluations, visual consistency with known events (such as exam schedules and holidays), and user feedback are crucial in validating the effectiveness of both the sentiment classification and the temporal trend visualizations.



# Insights

* **Emotional Trends:** The moving average visualization successfully highlighted fluctuations in student emotions over time, identifying periods of stress and happiness.
* **Lexicon Limitations:** While the lexicon-based approach provided a straightforward implementation, it highlighted the need for a more robust model (e.g., a supervised machine learning or transformer-based model) when scaling up the analysis.
* **Data Annotation:** Manual annotation of a sample subset could further improve performance evaluation, paving the way for training more advanced models in the future.
* **Visualization Utility:** The additional plot highlighting stress and happy periods proved valuable for stakeholders to quickly grasp key emotional patterns.



A graph with colored lines

AI-generated content may be incorrect.

# Conclusion

The assignment provided valuable insights into both sentiment classification and temporal sentiment analysis in Roman Urdu diary entries. Despite challenges such as limited labeled data and the subtleties of language, the implemented solution demonstrated that even simple methods can yield useful insights into emotional trends. Future work will involve refining the sentiment detection approach and exploring more sophisticated machine learning techniques to improve performance and reliability.