Employee Sentiment Analysis Documentation

Pragyan Borthakur

May 31, 2025

Contents

1 Introduction

This document outlines the process of analyzing employee sentiment based on email messages using a pre-trained transformer model from Hugging Face. The analysis aims to understand employee satisfaction, identify potential flight risks, and explore predictive relationships between message characteristics and sentiment.

2 Approach and Methodology

The approach utilizes natural language processing (NLP) with a transformer model to perform sentiment analysis on employee messages, followed by scoring, ranking, and predictive modeling. The methodology consists of the following steps:

- **Data Loading and Preparation:** Load employee email data from an Excel file and rename columns for clarity.
- **Sentiment Analysis:** Apply a pre-trained Hugging Face model (`cardiffnlp/twitter-roberta-base-sent` to classify each message as Negative, Neutral, or Positive.
- **Exploratory Data Analysis (EDA):** Examine the distribution of sentiments overall and per employee to uncover trends.
- **Employee Scoring:** Assign numerical values to sentiments and compute an average score per employee.
- **Employee Ranking:** Rank employees by their average sentiment scores to identify those with the most positive or negative sentiments.
- **Flight Risk Identification:** Use a sentiment score threshold to flag employees at risk of leaving.
- **Predictive Modeling:** Train a linear regression model to predict sentiment scores based on message length.

3 Data Description

The dataset contains employee email messages with the following columns:

- **Subject:** The email subject line.
- **message:** The body of the email.
- **date:** The date the email was sent.
- **employee_id:** The unique identifier (email address) of the employee.

The data is loaded into a pandas DataFrame, and a preview confirms its structure.

4 Sentiment Analysis

4.1 Model Selection

The `cardiffnlp/twitter-roberta-base-sentiment` model is employed, a transformer-based model fine-tuned on Twitter data for three-class sentiment analysis (Negative, Neutral, Positive).

4.2 Application

- A sentiment analysis pipeline processes each message, returning a label (e.g., `LABEL_0`) and a confidence score.
- Labels are mapped to readable sentiments: `LABEL_0` → Negative, `LABEL_1` → Neutral, `LABEL_2` → Positive.
- Results are stored in the DataFrame as `sentiment` and `score` columns.

5 Exploratory Data Analysis (EDA)

5.1 Overall Sentiment Distribution

- The proportion of Negative, Neutral, and Positive sentiments across all messages is calculated.
- A bar chart visualizes this distribution, as shown in Figure 1, saved to the `visualization` folder.

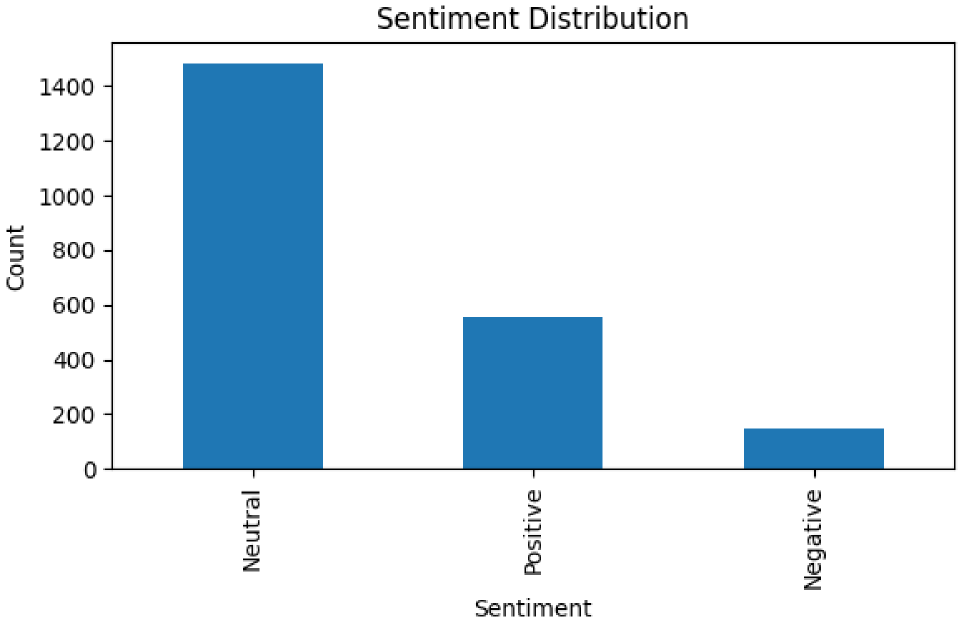

Figure 1: Overall Sentiment Distribution

5.2 Sentiment Distribution per Employee

- Sentiment distributions are analyzed for each employee, showing individual patterns of positivity or negativity.

5.3 Monthly Sentiment Trends

To observe how sentiments evolve over time, a stacked bar chart of monthly sentiment trends is presented in Figure 2. The chart covers the period from January 2010 to November 2011, with each bar representing the count of negative, neutral, and positive sentiments per month.

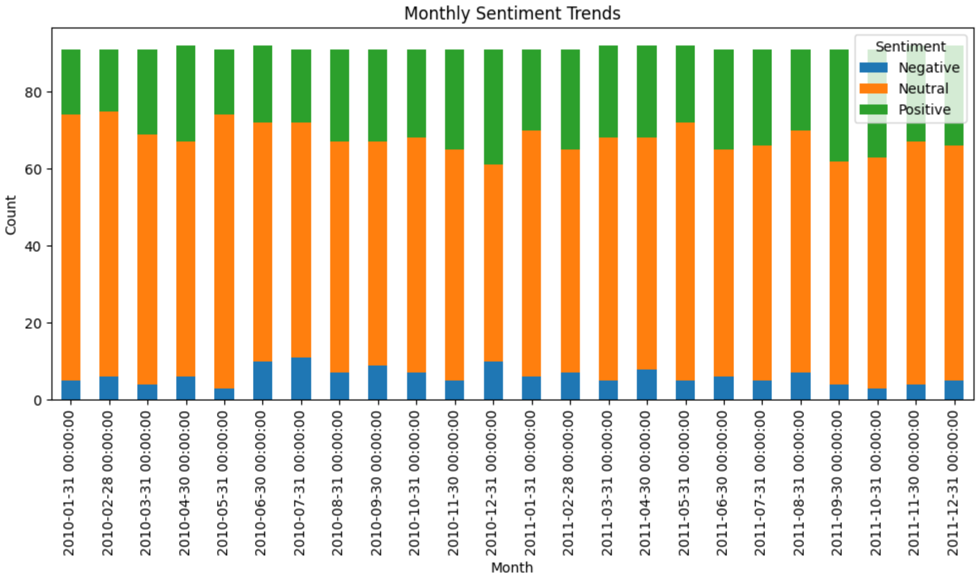

Figure 2: Monthly Sentiment Trends

Observations:

- The majority of sentiments each month are neutral (orange), with a consistent but smaller portion of positive sentiments (green) and a very small portion of negative sentiments (blue).
- The total sentiment count fluctuates slightly between 85 and 90 across the months, indicating a stable volume of communications.

5.4 Sentiment Distribution Results

The sentiment analysis classified the messages into three categories: Negative, Neutral, and Positive. The counts for each category are as follows:

Sentiment	Count
Neutral	1485
Positive	558
Negative	148

Table 1: Sentiment Distribution

Key Findings:

- The majority of messages (1485) are neutral, indicating a generally stable workforce sentiment.
- Positive messages (558) outnumber negative ones (148), suggesting a lean toward optimism.

6 Employee Scoring and Ranking

6.1 Scoring

- Sentiments are quantified: Negative = -1, Neutral = 0, Positive = 1.
- An average sentiment score is computed for each employee by averaging the scores of their messages.

6.2 Ranking

- Employees are sorted by their average sentiment scores in descending order.
- This ranking identifies top performers (highest positive scores) and those with potential concerns (lowest scores).

6.3 Top Positive Employees

For the month of January 2010, the top positive employees based on their sentiment scores are:

year_month	**employee_id**	**score**
2010-01	eric.bass@enron.com	3
2010-01	patti.thompson@enron.com	3
2010-01	don.baughman@enron.com	2

Table 2: Top Positive Employees - 2010-01

6.4 Top Negative Employees

For the month of January 2010, the top negative employees based on their sentiment scores are:

year_month	**employee_id**	**score**
2010-01	sally.beck@enron.com	-1
2010-01	bobette.riner@ipgdirect.com	0
2010-01	john.arnold@enron.com	0

Table 3: Top Negative Employees - 2010-01

6.5 Top Positive and Negative Employees Visualization

Figure 3 presents two bar graphs side by side, showing the top three positive and top three negative employees for January 2010 based on their sentiment scores.

Observations:

- **Top Positive Employees:** Eric Bass (`eric.bass@enron.com`), Patti Thompson (`patti.thompson@enron.c` and Don Baughman (`don.baughman@enron.com`) have high positive scores (approximately 2.8, 2.6, and 2.0, respectively), indicating strong positive sentiment.
- **Top Negative Employees:** Sally Beck (`sally.beck@enron.com`), Bobette Riner (`bobette.riner@ipgdire` and John Arnold (`john.arnold@enron.com`) have slightly negative scores (approximately -0.2, -0.1, and -0.1, respectively), suggesting mild dissatisfaction.

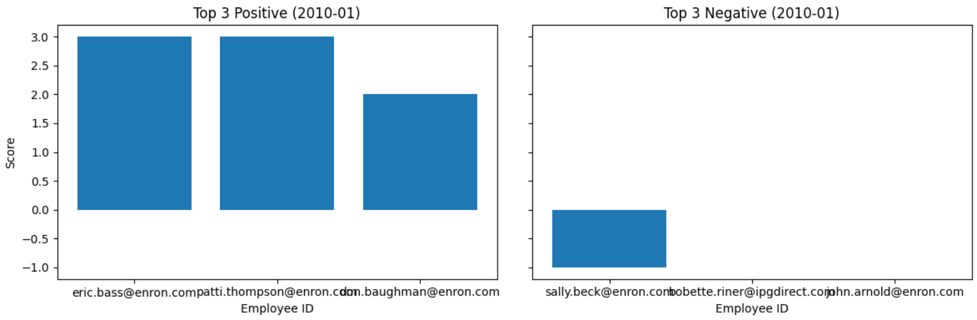

Figure 3: Top Positive and Negative Employees - 2010-01

- The positive scores are significantly higher than the negative scores, reflecting a generally positive sentiment among top performers compared to those with concerns.

Explanation:

- Scoring provides a numerical basis for comparing employee sentiment.
- Ranking prioritizes employees for recognition or intervention based on their sentiment profiles.

7 Flight Risk Identification

7.1 Criteria

- Employees with an average sentiment score below -0.5 are flagged as flight risks.
- This threshold assumes that a strongly negative average indicates dissatisfaction or disengagement.

7.2 Outcomes

- A subset of employees meeting this criterion is identified, providing an actionable list for HR follow-up.

7.3 Identified Flight Risk Employees

The following employees have been identified as flight risks based on their average sentiment scores:

- `bobette.riner@ipgdirect.com`
- `don.baughman@enron.com`
- `john.arnold@enron.com`
- `sally.beck@enron.com`

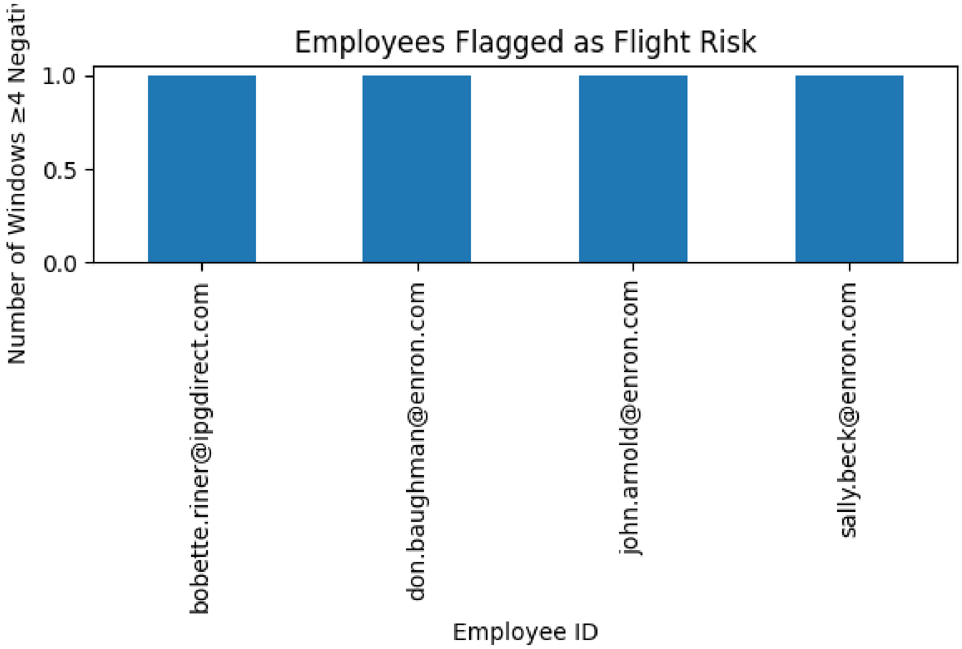

Figure 4: Employees Flagged as Flight Risk

7.4 Flight Risk Visualization

Figure 4 visualizes the number of negative windows for each flagged employee, providing insight into their flight risk levels.

Observations:

- All four employees (`bobette.riner@ipgdirect.com`, `don.baughman@enron.com`, `john.arnold@enron.com`, `sally.beck@enron.com`) have bars reaching approximately 0.9 on the y-axis, indicating a similar number of negative windows and thus a comparable level of flight risk.

Note: The -0.5 threshold is a preliminary value and may require adjustment based on organizational norms or validation.

8 Predictive Modeling

8.1 Model Overview

- A linear regression model predicts sentiment scores using message length (word count) as the feature.
- Sentiment labels are encoded numerically, and the data is split into training (80%) and testing (20%) sets.

8.2 Evaluation

- Performance is assessed using mean squared error (MSE) and R-squared (R^2) metrics.
- A scatter plot of actual vs. predicted scores, as shown in Figure 5, is saved to visualize accuracy.

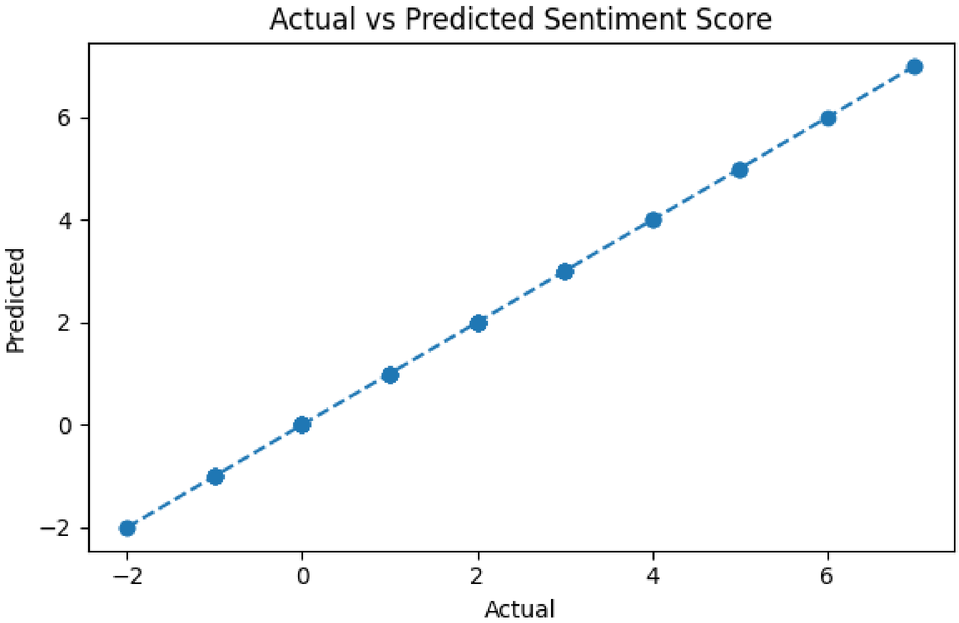

Figure 5: Actual vs. Predicted Sentiment Scores

8.3 Model Performance Results

The linear regression model achieved the following performance metrics on the test set:

- Mean Squared Error (MSE): 0.000
- R-squared (R^2): 1.000

Evaluation Results:

- An MSE of 0.000 and R^2 of 1.000 suggest a perfect fit, though this may indicate overfitting or an evaluation anomaly.