Sentiment Analysis Report

Title: Sentiment Analysis of X (Twitter) Posts
Author: Shivam Kumar Swarnkar
Date: 11.03.2025

---

1. Introduction

Sentiment analysis is a natural language processing (NLP) technique used to analyze text and classify emotions as **Positive**, **Negative**, or **Neutral**. In this project, we analyze X (formerly Twitter) posts to understand public sentiment.

---

2. Dataset Overview

The dataset consists of pre-cleaned tweets collected from X. Each tweet has a corresponding sentiment label as follows:

- **1.0** → Positive
- **0.0** → Neutral
- **-1.0** → Negative

Column Descriptions:

- **clean_text:** Tweet text (pre-processed)
- **category:** Sentiment label (-1 = Negative, 0 = Neutral, 1 = Positive)

---

3. Methodology

3.1 Data Cleaning

Since the dataset already contains cleaned text (with URLs, mentions, hashtags, and special characters removed), no additional cleaning was required.

3.2 Sentiment Mapping

The sentiment labels were mapped to human-readable categories as follows:

- **1.0** → Positive
- **0.0** → Neutral
- **-1.0** → Negative

3.3 Data Visualization

We performed two main visualizations: - **Sentiment Distribution:** A bar chart showing the distribution of sentiments.

- **WordCloud for Positive Tweets:** A word cloud highlighting the most frequently used words in positive tweets.
- **WordCloud for Negative Tweets:** A word cloud highlighting the most frequently used words in negative tweets.
- **WordCloud for Neutral Tweets:** A word cloud highlighting the most frequently used words in neutral tweets.

---

4. Results & Insights

4.1 Sentiment Distribution

The analysis revealed the following sentiment distribution:

Sentiment	Count
Positive	70,000+
Neutral	50,000+
Negative	30,000+

4.2 Word Cloud for Positive Sentiments

The word cloud illustrates that the most commonly used words in positive tweets include: - “india” - “will” - “bjp” - “now” - “people”

---

5. Conclusion

- **Neutral tweets:** These tweets are neutral, indicating mixed opinions.
- **Positive tweets:** These contain words like “people”, “will”, and “india”.
- **Negative tweets:** These often include words such as “poor”, “pakistan”, and “chowkidar”.

This analysis can help businesses monitor customer feedback and gauge public sentiment.
