

Problem statement information: (SMART)

Which single class represents the sentiment of the majority of tweets on Twitter?

Context:

Twitter, now X, is a social media giant that allows users to create posts, or tweets, and share them with others through their platform. These tweets can contain a wide range of words. For humans, the task of determining the attitude (sentiment) of tweets is a relatively easy one. Computers do not have the same luxury. However, with advancements in natural language processing, computers are beginning to bridge the gap and the results are promising. Given a tweet and an entity (company), the task is to judge the sentiment of the tweet about the entity.

Criteria for Success:

Determine 1 class as the majority sentiment of tweets on Twitter

Scope of solution space:

Three classes: Positive, Negative and Neutral. Irrelevant messages are regarded as Neutral.

Constraints:

Some tweets may contain instances of more than one sentiment, potentially leading to misclassifications. The hope is that such a case will be assigned as Neutral.

Stakeholders:

Social media companies

Data Sources:

Two CSV files containing the Twitter Sentiment Analysis Dataset from Kaggle

Citation:

passionate-nlp. (January 2021). Twitter Sentiment Analysis Dataset. Retrieved [3/21/2025] from <https://www.kaggle.com/datasets/jp797498e/twitter-entity-sentiment-analysis>