

Multi-Class Text Classification of COVID-19 Tweets

ARYAMAN PATEL

Performed Cross Validation analysis, Feature importance analysis and Error analysis on a covid-19 tweets dataset where the model was trained to analyze sentiments.

1. Cross-Validation (CV) Performance Analysis:

In assessing the model's performance, three CV strategies were employed: **2-fold**, **10-fold**, and **20-fold**.

2-fold CV: It was found to be less stable, likely due to the higher variance inherent in fewer data splits.

10-fold CV: This is commonly used in practice and struck a balance between computation power and its performance estimation.

20-fold CV: While providing a thorough evaluation, it did not significantly enhance performance over the 10-fold CV, suggesting a point of diminishing returns with respect to the number of folds.

2. Feature Importance Analysis and Model Interpretation:

Feature weights assigned by the logistic regression model were analyzed to pinpoint the most and least influential features for sentiment classification.

Positive Sentiment Class:

Top Features:

Hopeful: Strong positive connotation, often present in optimistic statements.

Thankful: Expressions of gratitude, typically reflecting positive sentiment.

Support: Indicative of positive social interactions and solidarity.

Bottom Features:

Cancelled: Although negative, its relevance to sentiment is less direct.

Delayed: Suggests inconvenience but not a strong sentiment indicator.

Problem: Varied usage across contexts dilutes its sentiment indication.

Negative Sentiment Class:

Top Features:

Worried: Directly correlates with negative sentiment.

Scared: Explicit expression of fear, commonly associated with negativity.

Death: Heavily negative and emotionally potent.

Bottom Features:

Statistics: Typically neutral, serving an informative purpose.

Report: Often used in a factual context, not inherently negative.

Study: Generally neutral, associated with academic or research contexts.

The model's weighting appears to favor terms with clear emotional associations for their respective sentiment classes, while terms with neutral or varying implications are deemed less predictive.

3. Error Analysis:

Analysis of misclassified tweets revealed several patterns:

Sentiment Ambiguity: Tweets with mixed sentiments or subtle emotional cues often led to classification errors.

Sarcasm and Irony: The model's limitations in detecting sarcasm and irony resulted in misclassifications, which is a common and big challenge in sentiment analysis.

Evolving Language: Terms that have taken on new emotional meanings in the context of the pandemic were sometimes misinterpreted, indicating a need for the model to adapt to language evolution.

For instance, in the **Positive class**, the most important features included words like **"thankful"** and **"support"** which are intuitively associated with positive sentiments. Conversely, the least important features included neutral terms like **"report"** and **"case"** which are less indicative of sentiment.

For the **Negative class**, terms like **"worry"** and **"problem"** were among the most important, reflecting the negative connotations. Least important features were more neutral or context-dependent, such as **"update"** or **"response"**.

These findings suggest that the model is effective at recognizing clear sentiment expressions but requires further refinement to handle the complexities of language, including context, sarcasm, and evolving usage. The model assigns higher weights to features that are strong indicators of the sentiment class based on their frequency and distinctiveness in the training data, whereas features with low weights are those that do not contribute significantly to class differentiation, possibly due to their common occurrence across multiple classes/their rarity.