**Chapter 5:**

**Lexical Study of Emojis**

This chapter aims to evaluate relevance of parameters for inclusion in an emoji lexicon with respect to the most optimal outcomes for later use to detect sarcasm. The most prevalent methodologies in literature are largely unfit for this purpose as sentiment is derived from the corresponding text, however, incongruency in text and emoji sentiment is widely cited as a marker of sarcasm thus there are validity concerns with regards to any works which deploy this strategy.

**Section 5.2.1**

**Dataset selection**

*Basic Emotional Theory*

By the nature of the problem domain, there is a necessity for human-annotated data, where emojis are presented independent of any text which may influence classifications. The EmoTag1200 dataset contains information regarding volunteers’ association of emojis to basic emotions, based on the Plutchik model. The presentation of emojis without any textual prompts makes this dataset uniquely suitable for this task, and methodologies for its generation are robust; multiple annotators were utilised, and agreement was monitored via Pairwise Pearson correlation and Krippendorff’s α. Steps were taken to mitigate fatigue bias through randomisation and reported results represent averages of the reported values.

**Section 5.2.2**

*Dimensional Theory*

The dataset used in this section was the Emoji Sentiment Ranking dataset, containing rankings of the 751 most popular emojis annotated by 83 annotators as positive, negative, or neutral. Annotators were presented with the emoji alongside the text, however, were instructed to rank the emoji sentiment only. While bias associated with the textual component of the content cannot be totally ruled out in this case, no datasets presently exist which are created where emojis were annotated independently of text.

**Sentiment classification:**

The following section aims to evaluate optimal approaches to define sentiment of emoji with regards to the two emotional theory models. Successful methodologies should generate parameters can fingerprint sentiment in terms of a wide array of features to facilitate highly accurate machine learning models to be constructed in subsequent work. Previous works cite challenges regarding emoji lacking characteristics to enable feature extraction. The use of annotated sentiment datasets will bypass this challenge.

*Basic Emotional Theory:*

The EmoTag1200 dataset contains information regarding the affinities of emoji to 8 basic emotions, which enable comparison and acknowledge that emotions are experienced synchronously with varying strengths based on the stimulus in question. Sentiment labels *c*, consist of continuous values representing affinity of the emoji to each emotion within set *e*:

where

Labels are most clearly equated to relative affinities to emotions; however, it is also reasonable to link these values to probabilities that a given emotion is the primary emotion associated with the emoji. Such a definition implies that:

where

The primary basic emotion is , the secondary is and the trend continues with decreasing affinity to the emoji. The former definition aligns more literally with the intention of original authors and provides increasing granularity which may be of value when more nuanced fingerprinting between emoji is necessary. Analysis of patterns discussed in section \*\*\* indicates that this is a highly complex task requiring computationally-expensive models to achieve due to the subtleties in sentiment differences. Such a task is challenging due to limitations associated with feature extraction of emoji. The latter yields reduced granularity however greater accuracy of models generated for prediction of unseen emoji may be possible, and thus the result may be data of overall greater value.

\*\* Add in smoothing optimisation discussion here\*\*

*Dimensional Theory:*

The Emoji Sentiment Ranking dataset contains information regarding the frequency to which emojis are classified as positive, negative, or neutral. These classifications were represented by discrete values which consider logical ordering of categories of equal distance apart to facilitate quantitative analysis:

Using these values, a discrete probability distribution for sentiment label *c*, can be determined where an assumption that the sum of the three respective probability distributions is equal to one is made to normalise data across emojis which appear at varying frequencies:

Where *c* is the sentiment label, and the following abbreviations were used to denote negativity, neutrality, and positivity:

(Negativity)

(Neutrality)

(Positivity)

Due to the low occurrence of annotation (*N* ≥ 5 is true for the lower quartile) presented in the dataset, relative frequency is not an ideal metric for approximation of probabilistic sentiment scores in many cases as will equal zero in events not observed in the data; such situations would result in bias due to the implication that certain events are impossible, whereas a low probability is more likely. For machine learning, contributes to overfitting as data is more sparse. Where non-zero probabilities are assigned to unseen events, some information about their potential occurrence is preserved and thus more robust patterns can be captured. Additionally, where *N* is small averages are increasingly subject to skew. Probability distributions were thus determined using a Laplace estimate:

Where *k* is the cardinality of the class, in this case |c| = 3. The assumption of uniform distribution for the estimator is true for the dataset in question. The smoothing model was selected for its ability to address the issue of zero-probabilities and mitigate the impacts of small *N* where relevant. At larger *N*, the output approaches prior to any transformation. Such a feature of the estimator focuses its effects on cases where it is most necessary. Finally, defining as the mean of the probability distributions weighted against their discrete labels, a sentiment score can be determined:

This approach aims to acknowledge the varying perceptions of emoji sentiment, dependent upon an individuals’ personal usage, while retaining ability to evaluate the sentiment with regards to the consensus classification. An overall sentiment score for sentiment using discrete categorisations often employs a majority decision methodology, however in this context an alternative methodology which acknowledges a degree of subjectivity is preferable. The standard error of the mean is one such methodology which enables a more nuanced classification methodology:

The result of such a methodology enables the identification of a most probable classification and inclusion or exclusion of secondary potential classifications at a 95% confidence.

A diagram of emojis

Description automatically generated

***Figure X*** Sample of result of emoji sentiment classification using dimensional emotional theory. Coloured bars are proportional to *p+, p0* and *p-*. Markers (black) represent . Grey markers represent the limits of the 95% confidence intervals for the . Note weary face (left) has high confidence that a negative classifier is appropriate, however ghost (right) cannot necessarily be classified as positive with 95% confidence.