# An Analysis of Emotion Detection Using Fine-Tuned BERT

**D. Curley** **J. Sarabia** **T. Hajny** **D. Jarrell** **B. Bob**

## Abstract

Sentiment Analysis (SA) has proven to be a highly effective way to gain an understanding of a users' positive/negative sentiment. More recently, researchers have been expanding the scope of SA in order to detect emotions. There is also an increasing recognition that many editorial task can be accomplished more efficiently using computer software. An attempt was made to apply Emotion Detection (ED) to the field of Bible translation with the hope that—the Bible's fixed corpus, narrative accounts with relatively clear boundaries, and many parallel English translations—the results would indicate major emotions present (or absent) in the original account in ancient Greek. This could then be utilized for comparative analysis. Unfortunately, the results were far from promising and, after much testing, proved to be insufficient for industrial applications and potentially called into question the overall effectiveness of any State of the Art system for Emotion Detection.

## 1 Introduction

Sentiment Analysis (SA) can be described as one of the core capabilities of artificial intelligence (AI) systems. It is being used to do things from understanding trends in customer opinion on products or services to political opinions from social media. In its basic form, SA outputs a scalar value. It is scalar because it only gives an amplitude, the amount some value is present in a text. It is not even capable of giving a single, binary distinction. Take a product review. The review can be thought of in a very basic form as expressing either positive or negative feelings about the product, to some degree. Basic SA needs to break this into two questions: (1) How much are positive emotions expressed in this text, and (2) How much are negative emotions expressed in this text. To get an overall understanding of the text, the individual scores can then be combined to form a vector with the amplitude being how strong the emotional content is and the direction being either positive or negative (quite literally). This analysis has been incredibly effective in helping companies to understand trends in customer opinion on products or services as well as helping people understand political opinions from social media. In Natural Language Processing (NLP), it can help researchers discover different elements of style.

Sentiment Analysis as a field has continued to grow and develop. New avenues of research has been pursued in attempt to broaden the application of SA. There has been a shift from SA to Emotional Analysis (EA). This is not simply an iteration for SA, but a paradigm shift if it can be accomplished successfully. This is because of the incredible complexity of analyzing human emotions. Elements affecting the emotional content of a text can be explicit, but more often they rely on subtle implications and implicatures. The original goal of this research was to determine whether EA could be effectively done on a narrative passage and compared to a parallel text. The choice was made to compare different English translations of the narrative account of the Prodigal Son in the Christian Bible. The hope was that translators would be able to use EA techniques to compare the emotional content of newly drafted Bible translations with existing Bible translations to improve translation accuracy. The source data was ideal as all the translations utilized would be expected to have the same general emotional content expressed in similar ways. Unfortunately, if EA is unable to produce sufficient quality in its analysis, there will not be much usage for it in industry.

## 2 Background

Sentiment Analysis (SA) has undergone an incredible amount of development in the last 25 years. Gupta et al. (2024) in their historical analysis of the field, follow SA from the early 2000s with rules-based statistical machine translation to today's deep learning techniques for neural machine translation. Although not technically a solved-problem, there are certainly contexts and scenarios where SA is highly effective. Perhaps the most notable implementation of SA is in industrial application telling the owner of a product the general view people have of their product based on a set of reviews.

While generalized SA is binary, emotions are much more complex. If SA can be thought of as a Boolean problem, perhaps emotions could be thought of as a multi-variable calculus problem. Not only this, emotions can appear both implicitly and explicitly in a text. When emotions are present explicitly, they can be directly detected in the text (joy is found indicating the presence of joy). When emotions are present implicitly, the are indirectly observable through inference (joy can be present because the main character was given a great gift) or through implicature (when the author wants you to feel a sense of joy). They can be inferred from implicit markers (like a wink) or through related terms (rejoice is found indicating the presence of joy). Because of this, Emotional Analysis (EA), also called Emotion Detection (ED), can be seen as more than just another developmental iteration of SA. EA is a whole new generation and with that it needs to be tested and understood before it can be implemented. So while EA may be considered promising to researchers, it is not considered to be a "solved-problem" in Natural Language Processing (NLP).

**Emotional Analysis**

Emotion detection of direct words is rather straightforward, leaving the true difficulty in identifying the strength and presence of an emotion when it is communicated indirectly in some way. Some have attempted to connect different lexemes to a given emotion(s) in the NRC Emotion Lexicon (cf. Mohammad & Turney, 2011), the WordNet-Affect Lexicon (cf. Strapparava & Valitutti, 2004), LIWC (cf. Tausczik & Pennebaker, 2010), SenticNet (cf. Cambria, et al., 2016), DepecheMood++ (cf. Araque, et al., 2018), Anew (cf. Bradley & Lang, 1999), VADER (cf. Hutto & Gilbert, 2014), SentiWordNEt (cf. Esuli & Sebastiani, 2006), and LexAT (cf. Neviarouskaya, 2007). A variety of models have been created without a clear standard emerging.

This begs the question of what is current State of the Art (SotA) for both SA and ED. It appears that the research in SA has shifted towards Multimodal Sentiment Analysis (MSA), incorporating information from a multitude of sources beyond just the text. This is certainly a positive development as human communication is often affected by historical context, socio-cultural factors, etc. The meaning of the words themselves are great affected when they are spoken and a whole host of elements (e.g. tone, volume, etc) are added to the words. This then begs the question, if the field is headed in that direction, are people using SA or ED for help with editorial decision-making? While written research is scant, it is likely that writing software companies do SA/ED work for features like Grammarly's AI-powered tone analysis.

When conducting this research, it was necessary to utilize an Large Language Model (LLM). The BERT family was chosen due to its open nature and the presence of a variety fine-tuned on 20 "Go" emotions (cf. Google Research, 2021; Demszky, 2020). Unfortunately, the list of emotions is far from standardized. After initial testing, other models were added to testing in an to attempt to improve results. This is discussed further below.

**Biblical Sentiment Analysis and Emotion Detection**

There are limited important studies in biblical NLP having to do with SA or EA. Perhaps thje most significant is Vora, et al (2024), who performed sentiment analysis of the Sermon on the Mount using BERT. In their experiment, they reviewed major sentiments expressed and found that the vocabulary of the respective translations is significantly different. Interestingly, they detected different levels of humor, optimism, and empathy. One issue with their research was in the choice of text. The Sermon on the Mount lacks a clear interpretation. It is a classic problem in NT Studies and there are many opinions about the function and audience of this sermon. The result is a passage to be analyzed that lacks certainty in its meaning which will make research near impossible. There are two other

works of note. The first is viz.Bible's character-based SA (cf. Rouse, 2023)). Interestingly, there results were rather ironic due to the fact that their results showed a higher overall positive sentiment in the Bible for certain people (notably Cyrus and Mary) than Jesus, whom Christians view as divine. The second work was done by http://openbible.info/OpenBible.info who examined the trends in sentiment (positive/negative) by studying the biblical text chronologically. This resulted in a general understanding of the good times and bad times throughout history (as recorded in the Bible)

## 3 Method

The initial iteration of testing was done on a narrow selection of texts on all 20 of the "Go" emotions present in BERT model utilized as the benchmark for all experiments (joeddav/distilbert-base-uncased-go-emotions-student). The test chosen for examination were a cross-section of popular Bibles used by evangelicals in the United States (NLT, ESV, KJV, NKJV, NASB(95), NIV(84), CSB, LSB). Each translation was run independently. This results in 8 normalized outputs which can be examined for mean and standard deviation. The case study chose was the narrative account known as the prodigal son from the gospel of Luke (15:11-32). From there other accounts with a sudden but vital element of relief were examined. These included the sacrifice of Isaac (Gen. 22:1-19) and David and Goliath (1 Samuel 17).

During the second iteration, The text's chosen were intended as prototypical examples of a given emotion other than relief. This attempted to built on the idea that the story of the prodigal son (Luke 15:11-32) has a very dramatic sense of relief that is central to the narrative that was not being detected by the LLM. Other accounts included: the death and resurrection of Jesus (Mark 14:1-6) which ends with tremendous joy. The story of Jonah (1:1-17) is one where Jonah's anger plays a central part in the story. With David and Bathsheba (11:1-12:17), their story ends with great sadness at the loss of their child. The intent behind these choices was to eliminate as many variables as possible and attempt to give the model narrative, which it was presumed was more normative for the model than more complex genre's (e.g. prophecy, wisdom, etc.). The narrative accounts have, as much as possible, certain emotional content clearly present. The goal was to avoid testing how word frequency impacts emotional analysis. To keep this data from interfering with our results, climatic emotions were examined. These emotions may be strong/intense and may even be a central element to the story, but these emotions are only present in the very climax of the story. This meant that we were better able to isolate other variables independent of the frequency of emotional terms.

The next iteration involved a switch to poetry. It was hypothesized that LLM performance could improve if the thing it is attempting to identify (an emotion) occurs repeatedly in a given passage. It could also hypothetically improve if a passage uses stronger words when describing emotions (e.g. dislike v. hate). Because of this, it was presumed that the LLM would have a greater chance of success examining Hebrew poetry because it often repeats ideas and themes within a psalm and it also uses highly emotionally-charged language. Emotions were also limited to Ekman's (1981) list with the exception of disgust which was not included as it is difficult to identify specific instances in the Bible. At first, a psalm was chosen to represent each of these emotions in a concentrated or very prototypical manner: anger (Ps. 7), joy (Ps. 110), Fear (Ps. 23), Sadness (Ps. 42), and Surprise (Ps. 126). The the variable of text length was examined. Instead of a whole psalm, a single verse was chosen as representative of each emotion: anger (4:4), joy (16:10), fear (56.3), sadness (42:11), and surprise (126:1).

From here, emotions were examined to determine whether different perspectives altered the results in a significant way. Conflicting emotions were isolated by viewing the narrative through the lens of each character. The story of the Prodigal Son was chosen as a case study (Luke 15:11-32). To isolate emotions, ChatGPT 4.o was used to summarize the narrative from three distinct perspectives: the Father, the Older Son, and the Younger Son. For each perspective, summaries of varying lengths—short and long—were generated. This method aimed to capture the unique emotional journey of each character and eliminate potential emotional overlaps that could skew the model's detection, such as the father's relief, older son's disgust, and younger son's guilt when his brother returns home.

The final test was completed by combining a perspective-based analysis with sliding win-

dow chunking. This iteration of testing sought to refine the results from the previous iteration by enhancing the model's ability to track emotional trajectories throughout the narrative. The previous method, which relied on sentence-by-sentence chunking, was replaced by a sliding window approach. In this method, the text is divided into overlapping segments, where each segment "slides" forward by a sentence, capturing multiple emotional states across continuous text. The hypothesis being that this allows the model to better capture emotional transitions and identify climactic moments, where emotional intensity builds over time, rather than isolating emotions within individual sentence boundaries.

## 4 Results

Since this initial iteration of testing, the results have been rather curious. For the Prodigal Son account, the highest scores, indicating the strongest correlation between the emotion and the passage were anger, confusion, curiosity, desire, excitement, grief, and remorse. The emotions not detected as being very present were amusement, joy, love, optimism, and relief. The highest scores would all fall into the category: expected results. This is because the prodigal son account (see: Figure 1) gave many high results that can reasonably be argued from the text.

The second iteration produced results for Jonah (Figure 2), David/Bathsheba (Figure 3), and death/resurrection of Jesus (Figure 4). For Jonah, anger was moderately present (0.55). For David and Bathsheba, sadness was moderate (0.48). For the death and resurrection, joy was very low (0.10). Interestingly, the results for Jonah were more concerning when compared to the highest scoring emotion, caring (0.79). Caring is the last thing someone would call Jonah. Likewise with David and Bathsheba, the high score of embarrassment (0.68) and confusion (0.55) are concerning. At best, this indicates a lack of certainty or trustworthiness in the model's results.

The results of the third iteration involving a switch to poetry (see: Figure 5) were inconclusive. The LLM rarely saw a score greater than 0.5 and never one greater than 0.67 even in these prototypical examples. Neither is this a normalizing issue as each emotion was never the most detected emotion. For "anger psalm," approval and realization were the most detected emotion (0.72 v. 0.25 for anger). For the "anger verse," relief was the most detected emotion (0.64 v. 0.52 for anger). For "fear psalm," realization was the most detected emotion (0.70 v. 0.19 for fear). For the "fear verse," relief was the most detected emotion (0.73 v. 0.19 for fear). For "joy psalm," embarrassment was the most detected emotion (0.69 v. 0.20 for joy). For the "joy verse," embarrassment was also a highly detected emotion (0.80 v. 0.66 for joy). For "sadness psalm," confusion and relief were the most detected emotions (0.70 v. 0.40 for sadness). For the "sadness verse," confusion was the most detected emotion (0.74 v. 0.5 for sadness). For "surprise psalm," approval was the most detected emotion (0.77 v. 0.35 for surprise). For the "surprise verse," nervousness was the most detected emotion (0.64 v. 0.29 for anger). This SotA model is really unable to consistently identify these types of emotions even in the best circumstances.

When examining a narrative account's emotional content from different perspectives, the results were not very promising. For the short summary from the older brother's viewpoint, the most prominent emotions detected were anger (0.72), remorse (0.68), and annoyance (0.67). These results are in line with expected emotional responses, given the nature of the narrative. Three emotions that displayed the greatest variation: sadness (0.42), disappointment (0.42), and joy (0.36), which suggests a wider emotional range in this perspective.

In the long summary from the older brother's perspective, the top emotions shifted slightly, with realization (0.63), remorse (0.62), and disapproval (0.57) emerging as the dominant emotions. Once again, the emotions with the most variation were sadness (0.28), disappointment (0.27), and grief (0.22). These findings reflect a deeper internal conflict.

From the younger brother's perspective, the short summary revealed realization (0.71), surprise (0.67), and remorse (0.66) as the top emotions. The most varied emotions in this case were disappointment (0.43), sadness (0.43), and relief (0.39), showing a blend of both negative and hopeful sentiments.

For the longer summary, the dominant emotions were realization (0.66), desire (0.62), and surprise (0.61). The emotions with the greatest variation were disappointment (0.34), sadness (0.33),

and joy (0.30), suggesting that while realization and desire are primary emotional drivers, there remains a level of internal emotional turmoil.

The father's perspective in the short summary provided a different emotional landscape, with the highest scores being caring (0.63), approval (0.57), and remorse (0.56). Emotions with the greatest variability, including sadness (0.45), disappointment (0.44), and grief (0.41).

In the long summary, the primary emotions detected were realization (0.60), desire (0.60), and remorse (0.59). Emotions with the most variation included disappointment (0.31), sadness (0.31), and annoyance (0.27), which demonstrates a nuanced blend of feelings, as the father reflects on the situation from a place of both understanding and internal conflict.

The results of the perspective-based analysis combined with sliding window chunking were as follows. In the short summary from the older Brother's viewpoint, the most prominent emotions were anger (0.80), annoyance (0.75), and remorse (0.74). The emotions displaying the greatest variation were sadness (0.40), disappointment (0.36), and excitement (0.28), suggesting a complex mix of dissatisfaction and occasional hope.

In the long summary, the top emotions were remorse (0.68), disappointment (0.62), and disapproval (0.61). The emotions that varied the most were disappointment (0.26), sadness (0.26), and anger (0.25), indicating persistent negative feelings over time.

With the cross-sectional approach, the results can best be described as being inconsistent to the point of being truly erratic with significant hallucination undermining the trustworthiness of the results (see: figure TODO). Perhaps most startling were the conclusions that the biblical account of the resurrection of Jesus, a central tenant or Christian faith, came back as high on embarrassment as well as the conclusion that the prophecy of God's final judgment on the Earth was high on "disappointment."

## 5 Conclusion

One is left with one rather unfortunate conclusion, current SotA for emotion detection has yet to reach a point of being accurate enough for use in industry, let alone one like Bible translation where trustworthiness and reliability of responses is at a premium. It is believed that the reason may be due to the very nature of sentimental analysis in comparison to emotion detection. The added dimensionality required for emotion detection to be successful appears to exceed current capabilities. Put simply, it appears the transition from SA to ED reflects a difference that may not be solvable through fine-tuning alone, but would require a whole separate approach like in math when moving from linear to non-linear system. There are a few possible paths for the development of future research. The development of logical thinking in AI systems, which is currently being researched extensively, is one obvious area that could have an impact on this task in the future. EmoLLMs have been proposed and show promise (Liu, et al., 2024). However, these researchers also recognize the lack of usable data, especially tagged data with clear emotional content. This does beg the question about the nature of emotions and our ability to quantify them appropriately. This certain would be worth investigating further. Other researchers have shown promise by rethinking ED as a multi-modal problem by researching dialogue exchanges. It would be interesting to investigate ED in narrower contexts, for instances, performing ED on different examples within a singular type of speech act.

## 6 References


- F. A. Acheampong, H. Nunoo-Mensah, W. Chen, Transformer models for text-based emotion detection: a review of BERT-based approaches, *Artificial Intelligence Review* (2021) 1–41.

- Araque, O., Gatti, L., Staiano, J., Guerini, M. (2018). *DepecheMood++: A Bilingual Emotion Lexicon Built Through Simple Yet Powerful Techniques*. *IEEE Transactions on Affective Computing*, 10(3), 471–483.

- Bradley, M. M., Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical report C-1, The Center for Research in Psychophysiology, University of Florida.

- Cambria, E., Poria, S., Bajpai, R., Schuller, B. (2016). *SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives*. *COLING* (pp. 2666–2677).

- Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., Ravi, S. (2020). *GoEmotions: A dataset of fine-grained emotions*. *arXiv*. https://arxiv.org/pdf/2005.00547https://arxiv.org/pdf/2005.00547.

- Esuli, A., Sebastiani, F. (2006). *SentiWordNet: A publicly available lexical resource for opinion mining*. *Proceedings of LREC*, 417–422.

- Google Research. (2021, May 27). *GoEmotions: A dataset for fine-grained emotion classification*. Google Research Blog. https://research.google/blog/goemotions-a-dataset-for-fine-grained-emotion-classification/https://research.google/blog/goemotions-a-dataset-for-fine-grained-emotion-classification/.

- Hutto, C., Gilbert, E. (2014). *VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text*. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media (ICWSM)*.

- Mohammad, S. M., Turney, P. D. (2013). *Crowdsourcing a word-emotion association lexicon*. *Computational Intelligence*, 29(3), 436–465.

- Neviarouskaya, A., Prendinger, H., Ishizuka, M. (2007). *Textual affect sensing for sociable and expressive online communication*. *Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction*, 218–229.

- Perry, P. O. (2012). *Affect WordNet*. http://corpustext.com/reference/affect$_{w}ordnet.html$.

- Rouse, R. (2023). *Sentiment analysis of biblical people*. Viz.Bible. https://viz.bible/sentiment-analysis-of-biblical-people/https://viz.bible/sentiment-analysis-of-biblical-people/.

- Strapparava, C., Valitutti, A. (2004). *WordNet-Affect: An affective extension of WordNet*. In *LREC* (pp. 1083–1086).


## 7 Figures

Figure 1: Prodigal Son

| **Emotion** | **Mean** | **Std Dev** |
|---|---|---|
| Admiration | 0.22 | 0.02 |
| Amusement | 0.19 | 0.03 |
| Anger | 0.65 | 0.03 |
| Annoyance | 0.55 | 0.05 |
| Anticipation | 0.54 | 0.04 |
| Approval | 0.50 | 0.05 |
| Caring | 0.73 | 0.07 |
| Confusion | 0.64 | 0.03 |
| Curiosity | 0.70 | 0.03 |
| Desire | 0.57 | 0.03 |
| Disappointment | 0.63 | 0.03 |
| Disapproval | 0.43 | 0.07 |
| Disgust | 0.54 | 0.04 |
| Embarrassment | 0.69 | 0.02 |
| Excitement | 0.32 | 0.02 |
| Fear | 0.22 | 0.02 |
| Gratitude | 0.59 | 0.02 |
| Grief | 0.13 | 0.02 |
| Joy | 0.18 | 0.02 |
| Love | 0.42 | 0.03 |
| Nervousness | 0.18 | 0.01 |
| Optimism | 0.24 | 0.02 |
| Pride | 0.45 | 0.04 |
| Realization | 0.17 | 0.02 |
| Relief | 0.66 | 0.03 |
| Remorse | 0.61 | 0.03 |
| Sadness | 0.51 | 0.01 |
| Surprise | 0.25 | 0.02 |

Figure 2: Jonah Emotion Results

| Emotion | Mean | Std Dev |
|---|---|---|
| Admiration | 0.19 | 0.02 |
| Amusement | 0.10 | 0.01 |
| Anger | 0.55 | 0.06 |
| Annoyance | 0.36 | 0.04 |
| Anticipation | 0.56 | 0.02 |
| Approval | 0.44 | 0.03 |
| Caring | 0.68 | 0.05 |
| Confusion | 0.47 | 0.04 |
| Curiosity | 0.52 | 0.03 |
| Desire | 0.45 | 0.03 |
| Disappointment | 0.51 | 0.04 |
| Disapproval | 0.38 | 0.06 |
| Disgust | 0.41 | 0.05 |
| Embarrassment | 0.49 | 0.02 |
| Excitement | 0.29 | 0.02 |
| Fear | 0.21 | 0.03 |
| Gratitude | 0.50 | 0.02 |
| Grief | 0.18 | 0.02 |
| Joy | 0.24 | 0.01 |
| Love | 0.42 | 0.03 |
| Nervousness | 0.17 | 0.02 |
| Optimism | 0.27 | 0.02 |
| Pride | 0.39 | 0.04 |
| Realization | 0.22 | 0.02 |
| Relief | 0.54 | 0.03 |
| Remorse | 0.48 | 0.03 |
| Sadness | 0.37 | 0.02 |
| Surprise | 0.26 | 0.02 |

Figure 3: David Sadness Results - Mean and Standard Deviation

| Emotion | Mean | Std Dev |
|---|---|---|
| Admiration | 0.21 | 0.04 |
| Amusement | 0.14 | 0.03 |
| Anger | 0.76 | 0.07 |
| Annoyance | 0.39 | 0.06 |
| Anticipation | 0.55 | 0.06 |
| Approval | 0.47 | 0.05 |
| Caring | 0.72 | 0.06 |
| Confusion | 0.61 | 0.05 |
| Curiosity | 0.59 | 0.04 |
| Desire | 0.51 | 0.04 |
| Disappointment | 0.65 | 0.04 |
| Disapproval | 0.46 | 0.07 |
| Disgust | 0.63 | 0.06 |
| Embarrassment | 0.72 | 0.03 |
| Excitement | 0.33 | 0.03 |
| Fear | 0.25 | 0.04 |
| Gratitude | 0.55 | 0.03 |
| Grief | 0.15 | 0.02 |
| Joy | 0.20 | 0.03 |
| Love | 0.46 | 0.04 |
| Nervousness | 0.19 | 0.02 |
| Optimism | 0.26 | 0.03 |
| Pride | 0.43 | 0.05 |
| Realization | 0.20 | 0.03 |
| Relief | 0.58 | 0.04 |
| Remorse | 0.63 | 0.04 |
| Sadness | 0.55 | 0.02 |
| Surprise | 0.28 | 0.03 |

Figure 4: Resurrection Results - Mean and Standard Deviation

| Emotion | Mean | Std Dev |
|---|---|---|
| Admiration | 0.17 | 0.01 |
| Amusement | 0.19 | 0.02 |
| Anger | 0.77 | 0.03 |
| Annoyance | 0.72 | 0.02 |
| Anticipation | 0.37 | 0.03 |
| Approval | 0.44 | 0.02 |
| Caring | 0.69 | 0.04 |
| Confusion | 0.54 | 0.03 |
| Curiosity | 0.60 | 0.03 |
| Desire | 0.51 | 0.02 |
| Disappointment | 0.66 | 0.02 |
| Disapproval | 0.42 | 0.03 |
| Disgust | 0.61 | 0.03 |
| Embarrassment | 0.65 | 0.02 |
| Excitement | 0.29 | 0.02 |
| Fear | 0.20 | 0.02 |
| Gratitude | 0.58 | 0.02 |
| Grief | 0.12 | 0.01 |
| Joy | 0.21 | 0.01 |
| Love | 0.47 | 0.02 |
| Nervousness | 0.22 | 0.01 |
| Optimism | 0.25 | 0.02 |
| Pride | 0.40 | 0.03 |
| Realization | 0.18 | 0.02 |
| Relief | 0.63 | 0.02 |
| Remorse | 0.59 | 0.02 |
| Sadness | 0.56 | 0.01 |
| Surprise | 0.26 | 0.02 |

Figure 5: Poetry Results

| Emotion | Anger(Chapter) | Anger(Verse) |
|---|---|---|
| Fear(Chapter) | Fear(Verse) | |
| Admiration | 0.17 | 0.01 |
| Amusement | 0.19 | 0.02 |
| Anger | 0.77 | 0.03 |
| Annoyance | 0.72 | 0.02 |
| Anticipation | 0.37 | 0.03 |
| Approval | 0.44 | 0.02 |
| Caring | 0.69 | 0.04 |
| Confusion | 0.54 | 0.03 |
| Curiosity | 0.60 | 0.03 |
| Desire | 0.51 | 0.02 |
| Disappointment | 0.66 | 0.02 |
| Disapproval | 0.42 | 0.03 |
| Disgust | 0.61 | 0.03 |
| Embarrassment | 0.65 | 0.02 |
| Excitement | 0.29 | 0.02 |
| Fear | 0.20 | 0.02 |
| Gratitude | 0.58 | 0.02 |
| Grief | 0.12 | 0.01 |
| Joy | 0.21 | 0.01 |
| Love | 0.47 | 0.02 |
| Nervousness | 0.22 | 0.01 |
| Optimism | 0.25 | 0.02 |
| Pride | 0.40 | 0.03 |
| Realization | 0.18 | 0.02 |
| Relief | 0.63 | 0.02 |
| Remorse | 0.59 | 0.02 |
| Sadness | 0.56 | 0.01 |
| Surprise | 0.26 | 0.02 |