Why Sentiment Analysis is a Joke with JOKER data? Word-Level and Interpretation Analysis (CLEF 2023 JOKER Open task)

Tremaine Thomas-Young [1]

[1] *HCTI, Université de Bretagne Occidentale, 20 Duquesne Street, Brest, 29490, France*

Abstract

Sentiment analysis, a subfield of natural language processing, aims to determine the emotional tone conveyed by text. While sentiment analysis has been extensively explored in various domains, the analysis of jokes poses unique challenges due to their inherent humor and often ambiguous nature. These working notes present an overview of sentiment analysis applied specifically to JOKER data, one word-level analysis and in application of wordplay interpretation from the JOKER corpus for sentiment analysis.

Keywords

Sentiment Analysis, Natural Language Processing, Word-Level & Interpretation Analysis, Humor

1. Introduction

The primary objective of sentiment analysis in jokes is to assess the emotional polarity associated with the words and phrases used within the joke. By examining the sentiment expressed in each component of a joke, researchers can gain insights into the intended humor and overall sentiment conveyed. This analysis can assist in automating the categorization of jokes based on their sentiment, which has implications for joke recommendation systems, content filtering, and targeted humor generation.

To conduct sentiment analysis of jokes, researchers leverage various techniques and approaches. Firstly, the identification and extraction of relevant textual features, such as humor-related words, puns, or sarcasm, play a crucial role. These features are then processed using sentiment lexicons, machine learning algorithms, or deep learning models, which assign sentiment scores to each word or phrase [9]. While word-level analysis offers benefits, it encounters obstacles including word ambiguity, sarcasm, and cultural subtleties, as words can exhibit varying meanings depending on context, sarcasm can convey sentiments contrary to their literal definitions, and cultural influences shape the sentiment attached to specific words, emphasizing the importance of considering the cultural background of the text's recipients.

Interpreting the sentiment analysis results in the context of jokes requires additional considerations. Humor is subjective, and the same joke may elicit different emotional responses from different individuals. Consequently, sentiment analysis of jokes must account for the nuances and cultural factors that influence the perception of humor. Furthermore, contextual information, such as the setup and punchline of a joke, must be considered to comprehend the overall sentiment [4].

[1]CLEF 2023: Conference and Labs of the Evaluation Forum, September 18–21, 2023, Thessaloniki, Greece
EMAIL: tthomas@univ-brest.fr

(Written with Open AI technology)

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

This paper will analyze JOKER Track data at the word-level and the interpretation of JOKER data with the entire puns using sentiment analysis to gauge the emotional tone conveyed by text. By analyzing JOKER Track data at the word-level and interpretation using sentiment analysis, it is possible to gauge the efficacy of emotional tone conveyed by the text. It is expected that sentiment analysis techniques, including word-level sentiment analysis and interpretation-level sentiment analysis, will capture and classify the sentiments expressed in the JOKER Track data. The hypothesis assumes that sentiment analysis can contribute to a deeper understanding of the sentiment conveyed in JOKER Track data.

2. Open Task

Comprehending and translating humorous wordplay presents a challenge for both humans and computers due to the need for recognizing implicit cultural references, understanding word formation processes, and discerning double meanings. These factors introduce complexities that hinder the accurate interpretation and translation of humorous language, requiring a deep understanding of linguistic nuances and cultural context. The JOKER-2023 track aims at an interdisciplinary approach to the automatic processing of wordplay. The JOKER track is a part of the CLEF initiative which promotes the systematic evaluation of information access systems, primarily through experimentation on shared tasks [5]. This paper will explore tasks around Wordplay and Interpretation of puns through evaluation of humor in English.

Task 2 of the JOKER Track Data involves the interpretation of puns in English. Puns are a form of wordplay that rely on the use of words or phrases with multiple meanings or similar sounds to create humor or a play on words. In this task, participants are provided with a dataset containing various puns in English and are required to analyze and interpret the intended meaning or humor behind each pun. The goal is to assess the participants' ability to understand the nuanced linguistic elements and contextual cues necessary to comprehend and appreciate puns in the English language. This task aims to advance the understanding of computational models and algorithms in capturing the intricacies of pun interpretation, contributing to the broader field of natural language processing and humor comprehension.

3. Discussion: Humor and Sentiment

Humor and sentiment are two fascinating aspects of human communication and emotions. While humor is often associated with laughter and amusement, sentiment refers to the underlying emotional tone expressed in a message. The exploration of humor and sentiment through joint studies offers valuable insights into the complex interplay between emotions and humor, leading to a deeper understanding of human psychology and communication dynamics.

Liana Ermakova, Anne Gwenn-Bosser, Adam Jatowt, and Tristan Miller, have conducted numerous studies to investigate The JOKER Corpus: English–French parallel data for multilingual wordplay recognition[3]. Miller in particular has conducted studies and developed computational approaches that aim to understand the structure and mechanisms of humor, as well as its application in natural language processing and artificial intelligence. His work often involves leveraging linguistic and semantic analysis techniques to detect and generate humorous content. One notable contribution by Tristan Miller is the development of the "Humor Computation" framework, which combines linguistic analysis, knowledge representation, and reasoning to computationally model various aspects of humor[11]. This framework provides a foundation for computational systems to analyze and generate humor in a structured and systematic manner.

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

One area of exploration involves examining the impact of humor on sentiment. Martin L. Martin and Herbert M. Lefcourt (1983): In their study, Martin and Lefcourt examined the effects of humor on emotional responses. They found that humorous material led to more positive emotions and enhanced positive affect in participants.The study aimed to examine how exposure to humorous material influences emotional states and affective experiences. Participants were exposed to various forms of humorous content, such as jokes, cartoons, and comedic videos. The researchers then measured the participants' emotional responses and assessed changes in their affective states. The findings of the study revealed that exposure to humorous material had a positive impact on emotional responses. Participants reported experiencing more positive emotions, such as joy, amusement, and happiness, after engaging with the humor stimuli. This suggests that humor can enhance positive affect and contribute to positive emotional experiences[7]. Lefcourt and Martin's study provided empirical evidence for the beneficial effects of humor on emotional well-being. Their research supported the idea that humor can serve as a mood enhancer and elicit positive emotional responses. These findings have implications for understanding the role of humor in promoting well-being, stress reduction, and coping mechanisms.

Conversely, researchers like Rada Mihalcea at the University of Michigan have also delved into how sentiment affects the perception and reception of humor. Different emotional states can influence individuals' receptiveness to various types of humor. For example, individuals in a positive mood may find a wider range of humor enjoyable, while those experiencing negative emotions may have a narrower appreciation for specific comedic styles. Sentiment also plays a role in determining the appropriateness and effectiveness of humor in different contexts. Joint studies have explored the nuanced relationship between humor and sentiment in specific domains [10]. Mihalcea's research focuses on leveraging sentiment analysis techniques to enhance computational models for detecting and generating humorous content. She investigates how sentiment, both positive and negative, influences the perception and interpretation of humor. By analyzing the emotional aspects of text, Mihalcea aims to improve computational models' ability to identify and generate humorous content. Her work includes developing sentiment-aware algorithms and models that can capture the affective dimensions of humor. These models consider sentiment-related cues in language, such as emotion words, sarcasm, and irony, to enhance the accuracy of humor detection and generation systems. Her research also explores the intersection of sentiment and humor in different domains, including social media and online platforms. By analyzing user-generated content and online conversations, she investigates how sentiment interacts with humor in these contexts.

In recent years, computational approaches have gained prominence in studying humor and sentiment jointly. Natural Language Processing (NLP) techniques have been employed to analyze large-scale datasets, social media content, and online forums to extract humor and sentiment-related information. These computational methods allow for quantitative analysis, pattern recognition, and the identification of underlying factors contributing to the relationship between humor and sentiment. The findings from joint studies exploring humor and sentiment have practical implications across various domains. They can inform the development of more engaging and persuasive communication strategies, improve the design of humor-based interventions in fields like healthcare and education, and facilitate the creation of emotionally intelligent conversational agents and chatbots [6]. But none of these studies focus on applied sentiment analysis at the word-level to identify and / or justify proper interpretation of the sentiment. Applied sentiment analysis at the word-level is crucial because it allows for precise identification and justification of proper interpretation, filling the gap in existing research that often focuses on higher-level emotional responses and overall sentiment without examining fine-grained nuances within the text.

The exploration of humor and sentiment through joint studies offers valuable insights into the complex interplay between emotions and humor. Yet understanding how humor influences sentiment

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

and vice versa within the digital space, we must unravel the mechanisms behind human communication, emotional responses, and the accuracy of sentiment analysis tools.

4. Overview of Word-Level Analysis in Sentiment Analysis

Word-level analysis is a fundamental component of sentiment analysis and natural language processing (NLP) techniques. It involves examining individual words within a text to determine their sentiment or emotional connotation. This approach provides valuable insights into the emotional tone, subjective meaning, and overall sentiment expressed within a given piece of text.

In sentiment analysis, word-level analysis involves assigning sentiment scores to individual words based on their polarity, which can be positive, negative, or neutral. Sentiment lexicons or dictionaries are commonly used resources that contain predefined sentiment scores associated with specific words[6]. These lexicons can be manually curated or generated using automated methods. By matching words in a text to sentiment lexicons, sentiment analysis algorithms can assign sentiment scores to each word, thereby quantifying the overall sentiment of the text [2].

Various techniques are employed in word-level analysis, ranging from rule-based approaches to machine learning and deep learning methods. Rule-based approaches utilize predefined rules or patterns to determine the sentiment of words. Machine learning techniques involve training models on labeled data, where words are associated with sentiment labels, to learn patterns and make predictions on new, unseen data. Deep learning models, such as recurrent neural networks (RNNs) and transformer models, can capture complex relationships between words and their context, enabling more accurate sentiment analysis[12].

Despite the advantages of word-level analysis, it faces challenges such as word ambiguity, sarcasm, and cultural nuances. Words can have different meanings based on context, and sarcasm may convey sentiments opposite to their literal meanings. Cultural factors influence the sentiment associated with certain words, making it necessary to consider the cultural background of the text's audience.

In summary, word-level analysis is a critical aspect of sentiment analysis and NLP. It involves examining individual words within a text to determine their sentiment or emotional connotation. By assigning sentiment scores to words, researchers can quantify the sentiment expressed in a text, enabling a deeper understanding of subjective meaning and facilitating various applications in sentiment analysis, text classification, and opinion mining.

Below in Table 1 is an example of the Word-Level Sentiment Analysis and Interpretation sentiment scores, in addition to charts (1, 2, and 3).

The word-level analysis Table 1 reveals the sentiment scores assigned to specific words. In this case, the word "Fall" is classified as having a positive sentiment with a score of 0.643, indicating that it is associated with positive connotations. The word "Work" is considered neutral with a score of 0.579, suggesting that it doesn't convey a strongly positive or negative sentiment. Lastly, the word "Callous" is categorized as having a positive sentiment with a score of 0.661, indicating that it is associated with positive sentiments despite its typically negative connotation. These sentiment scores provide insights into the emotional tone attributed to these words within the given context of analysis.

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

Table 1
Word-Level Analysis

Word	Sentiment	Score
Fall (en_157)	Positive	0.643
Work (en_591)	Neutral	0.579
Callous (en_5969)	Positive	0.661

Chart 1

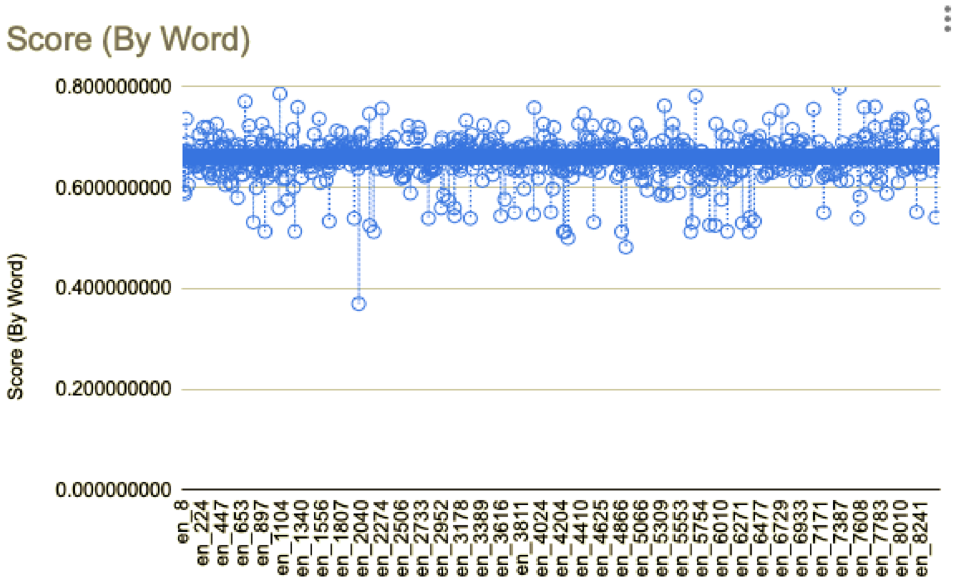


Figure 1: Word-level Sentiment Analysis Scores

The chart has condensed all JOKER Track Task 2 data into sentiment scores assigned to specific "words". The Y axis reflects sentiment score (0 -1) by word and the X axis reflects the complete words list (en_1 through en_8241). Including the words used in Table 1 - "Fall" (en_157) with a positive sentiment score of 0.643, "Work" (en_591) with a neutral sentiment score of 0.579, and "Callous" (en_5969) with a positive sentiment score of 0.661. The scores reflect the emotional connotations associated with these words, providing valuable insights into their sentiment within the analyzed context.

5. Wordplay interpretation for Sentiment Analysis

Interpreting humor data with the entire puns can play a vital role in the continued development of sentiment analysis and natural language processing (NLP) research by providing insights into the meaning and implications of sentiment analysis results. It involves understanding and analyzing the context, nuances, and cultural factors that influence the sentiment expressed in text, thereby enabling a deeper understanding of the overall message conveyed [9].

Interpretation in sentiment analysis encompasses several aspects :

1. Contextual Analysis: To accurately interpret sentiment, it is crucial to consider the broader context in which the text is presented. This includes analyzing the surrounding sentences, paragraphs, or even the entire document to capture the full meaning and intended sentiment.

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

Contextual analysis helps avoid misinterpretation that may arise from considering individual words or phrases in isolation.

2. Tone and Intensity: Sentiment analysis typically provides polarity labels such as positive, negative, or neutral. However, interpreting sentiment goes beyond polarity. It involves assessing the tone and intensity of sentiment expressed, which can range from mild to strong. For example, a statement may be slightly positive or extremely negative, conveying different levels of sentiment impact.

3. Sarcasm and Irony: Sentiment analysis faces challenges when dealing with sarcastic or ironic expressions, where the sentiment conveyed may be opposite to the literal meaning of the words. Interpreting such instances requires understanding the context, linguistic cues, and cultural knowledge to identify the underlying sentiment correctly.

4. Subjectivity and Cultural Nuances: Sentiment analysis should consider the subjectivity of emotions and the cultural influences that shape sentiment expression. Different cultures and communities may associate different sentiments with particular words or expressions. To ensure accurate interpretation, it is essential to account for cultural nuances and context-specific interpretations of sentiment.

5. Domain-Specific Interpretation: Sentiment analysis is often applied in specific domains, such as product reviews, social media, or financial markets. Interpretation requires domain knowledge and understanding of the specific terminology and jargon used within that domain. This domain-specific interpretation enhances the accuracy and relevance of sentiment analysis results.

6. Visualizations and Explanations: To facilitate understanding and interpretation, visualizations and explanations of sentiment analysis results can be valuable. Visual representations, such as word clouds, bar charts, or heatmaps, can highlight the distribution and intensity of sentiment in a text. Additionally, providing explanations or highlighting influential words or phrases can enhance the transparency and trustworthiness of sentiment analysis results.

Interpretation in sentiment analysis and NLP is an iterative process that requires human judgment and domain expertise. While automated algorithms can provide sentiment scores, understanding the true meaning and implications of sentiment within a given context still heavily relies on human interpretation [1]. Through continuous research and refinement, the interpretation of sentiment analysis results can be further improved, leading to more accurate and insightful analyses of text sentiment.

Below in Table 2 is an example of the same words as in Table 1. However, Interpretation Analysis of the sentiment and scores were applied.

The interpretation level analysis Table 2 presents sentiment scores assigned to specific interpretations of words. In this case, the interpretation "autumn / drop" is categorized as having a positive sentiment with a score of 0.649, indicating a favorable emotional tone associated with the concept of autumn or dropping. The interpretation "work out; exercise / work out" is also considered positive with a higher score of 0.781, suggesting a strong positive sentiment linked to physical exercise and workouts. Conversely, the interpretation "indurate; pachydermatous / callus; callosity" receives a negative sentiment score of 0.236, indicating a negative emotional connotation associated with the notion of calluses or hardened skin. These sentiment scores provide insights into the emotional responses and perceived sentiments related to specific interpretations of the given words within the analyzed context.

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

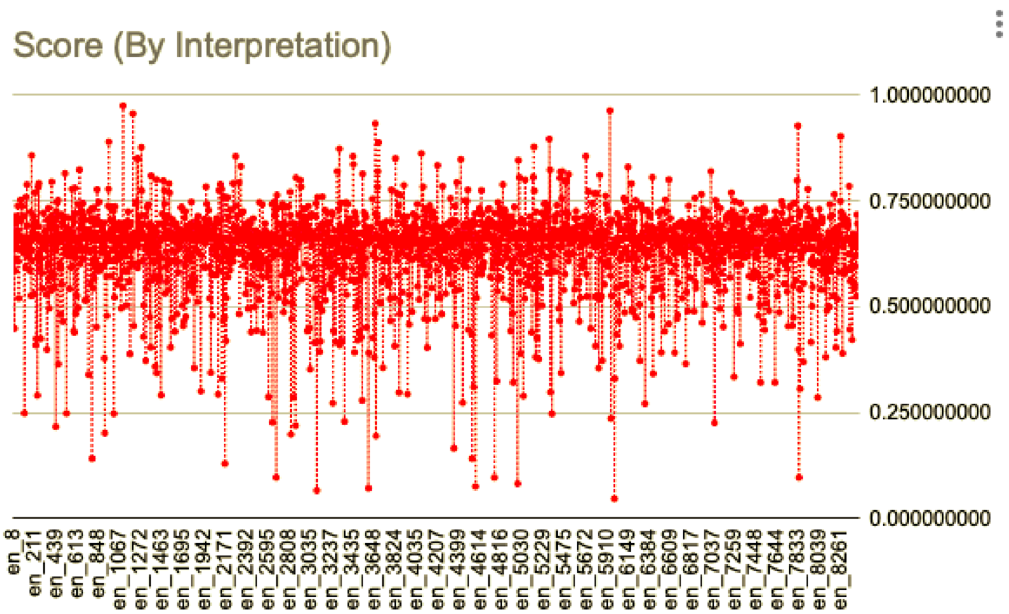
Table 2

Interpretation-level Sentiment Analysis Scores

Word	Sentiment	Score
autumn / drop (en_157)	Positive	0.649
work out; exercise / work out (en_591)	Positive	0.781
indurate; pachydermatous / callus; callosity (en_5969)	Negative	0.236

(Plan to experiment more in the future)

Chart 2



Caption for Chart 2: Interpretation-level Sentiment Analysis Scores

The chart has condensed all JOKER Track Task 2 data into sentiment scores assigned to specific "words". The Y axis reflects sentiment score (0 -1) by word and the X axis reflects the complete words list (en_1 through en_8241). Including the words used in Table 2 sentiment scores assigned to different interpretations of words. The interpretations "autumn / drop" (en_157) and "work out; exercise / work out" (en_591)are both associated with positive sentiments, scoring 0.649 and 0.781, respectively. In contrast, the interpretation "indurate; pachydermatous / callus; callosity" (en_5969) receives a negative sentiment score of 0.236. These sentiment scores reveal the emotional connotations attributed to specific interpretations of the words, providing valuable insights into their perceived sentiment within the analyzed context.

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

Chart 3

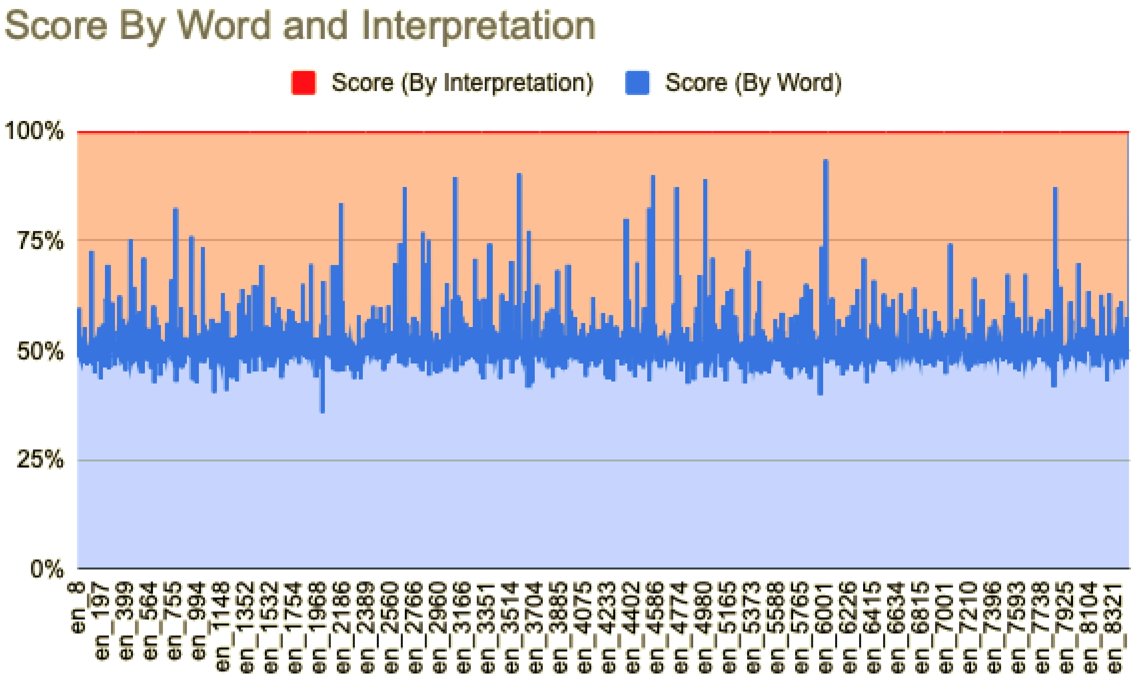


Figure 2: Interpretation-level Sentiment Analysis Scores

The chart has condensed all JOKER Track Task 2 data into sentiment scores assigned to specific "words". The Y axis reflects sentiment score (0 -1) by word and the X axis reflects the complete words list (en_1 through en_8241). Including the words used in Figures 1 & 2 sentiment scores.

6. Methodology

The goal of the methodology is to perform sentiment analysis at the word-level and interpretation using the JOKER Track data task 2. By analyzing sentiment at the word-level and interpretation, we can gain deeper insights into the efficacy of emotional tone conveyed by text and improve the accuracy of sentiment analysis compared to traditional document-level approaches.

The methodology aims to evaluate the effectiveness of sentiment analysis techniques applied to the JOKER Track data by examining the word-level sentiment and interpretation of words. By comparing the results with existing sentiment analysis methods, the goal is to determine if analyzing sentiment at a more granular level provides more accurate and nuanced insights into the emotional content of the text.

The sentiment analysis feature provides sentiment labels (such as "negative", "neutral" and "positive") based on the highest confidence score found by the service at a sentence and document-level. This feature also returns confidence scores between 0 and 1 for each document & sentences within it for positive, neutral and negative sentiment. This process collects a representative dataset that includes a range of sentiments (positive, negative, neutral) related to the target domain. Preprocessing all of the Task 2 of the JOKER involves the interpretation of puns in English which was performed by focusing on word-level and interpretation of those words to be analyzed.

*The Microsoft Azure Cognitive Service for Language was used to perform the task.

7. Results & Conclusion

The final results at the word level and the interpretation of JOKER data with the entire puns are inconclusive. The programming language, libraries, and frameworks are inadequate for processing word-level and interpretation analysis of sentiment. Future research should focus on the use of sentiment analysis at the word level as well as contextual analysis that require designed models

In conclusion, sentiment analysis of jokes at the word level and the interpretation of JOKER data with the entire puns cannot provide a valuable tool for understanding the emotional dynamics of humorous content. By employing sentiment analysis techniques, researchers may gain insights into the sentiment conveyed by groups of words and phrases within jokes. However, leading to improved comprehension of a joke or humor and potential applications in humor-related technologies are far from having the capacity to definitively predict sentiment analysis at the word-level and interpretation. Further research is needed to address the challenges posed by the subjectivity and ambiguity at the word-level before the context of jokes can be interpreted as useful sentiment analysis results within the context of humor.

8. References

[1] Balahur, A. and Steinberger, R., 2009. Rethinking Sentiment Analysis in the News: from Theory to Practice and back. Proceeding of WOMSA, 9, pp.1-12.

[2] Chen, M., Wang, S., Liang, P.P., Baltrušaitis, T., Zadeh, A. and Morency, L.P., 2017, November. Multimodal sentiment analysis with word-level fusion and reinforcement learning. In Proceedings of the 19th ACM international conference on multimodal interaction (pp. 163-171).

[3] Ermakova, Liana, et al. "Overview of joker@ clef 2022: Automatic wordplay and humor translation workshop." International Conference of the Cross-Language Evaluation Forum for European Languages. Cham: Springer International Publishing, 2022.

[4] Kuipers, G., 2015. Good humor, bad taste. In Good Humor, Bad Taste. De Gruyter Mouton.

[5] Liana Ermakova, Tristan Miller, Anne-Gwenn Bosser, Victor Manuel Palma Preciado, Grigori Sidorov, and Adam Jatowt. 2023. Overview of JOKER - CLEF-2023 track on Automatic Wordplay Analysis. In Avi Arampatzis, Evangelos Kanoulas, Theodora Tsikrika, Stefanos Vrochidis, Anastasia Giachanou, Dan Li, Mohammad Aliannejadi, Michalis Vlachos, Guglielmo Faggioli, Nicola Ferro (Eds.) Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Fourteenth International Conference of the CLEF Association (CLEF 2023)

[6] Liu, Bing, and Bing Liu. "Opinion mining and sentiment analysis." Web data mining: exploring hyperlinks, contents, and usage data (2011): 459-526.

[7] Martin, R.A. and Ford, T., 2018. The psychology of humor: An integrative approach. Academic press.

[8] Martin, Rod A., and Herbert M. Lefcourt. "Sense of humor as a moderator of the relation between stressors and moods." Journal of personality and social psychology 45.6 (1983): 1313.

[9] Mejova, Y., 2009. Sentiment analysis: An overview. University of Iowa, Computer Science Department.

[10] Mihalcea, Poria, Soujanya, et al. "Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research." IEEE Transactions on Affective Computing (2020).

[11] Miller, Tristan, and Iryna Gurevych. "Automatic disambiguation of English puns." Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2015.

[12] Pang, Bo. "Lee, l.(2008)." Trends in Information Retrieval 2.

[13] Weinberger, Marc G., and Charles S. Gulas. "The impact of humor in advertising: A review." Journal of advertising 21.4 (1992): 35-59.

*The Microsoft Azure Cognitive Service for Language was used to perform the task.