Grenze International Journal of Engineering and Technology, June Issue

Assisted Dying: A Text Analysis of Public Sentiment and Opinion

Pranitha M R[1], V N Manjunath Aradhya[2], Manoj Kumar C S[3] and Nikhil D Bharadwaj[4]
[1-4]Department of Computer Applications, JSS Science and Technology University, Mysuru, India
Email: aradhya@jsstuniv.in

***Abstract*— Assisted dying has become increasingly acute in ethical, social, and political debates in current society. In this paper, Natural Language Processing (NLP) is used to analyze public opinions, sentiment, emotion and opinion on assist-ed dying. This study was performed on the public discourse dataset Search_All_LinksData 2014-2025, which contains 497 cleaned textual records of a decade. Data Collection and Pre-processing included Tokenization, Lemmatization, and Stop-word removal. For polarity measurement and con-textual interpretation, Text Blob and a fine-tuned BERT model was utilized. The results indicate that public opinion is polarized with 27.4% neutral, 38.2% positive, and 34.4% negative sentiment. Assisted dying related online discussions are inspired by autonomy, dignity, moral uncertainty and ideas of compassion. Overall, the study how NLP provides the means to detect emotional and moral dimensions within healthcare-related debates, to bring insights for improving the research and framing policies.**

***Keywords*— Assisted dying, Natural Language Processing (NLP), TextBlob, Sentiment analysis, Transformer models, BERT.**

I. INTRODUCTION

Assisted dying has become one of the most debated ethical and social issues of the current generation, in connection with euthanasia and physician-assisted suicide. Competing values in compassion versus autonomy are at stake in ethical, cultural and medical places. Online social media forums expose public opinion to increasingly active debates with others surrounding assisted dying, as indi-viduals give expression to the moral reasoning and emotions through statements made online. This research turns on the NLP techniques to evaluate public senti-ment and moral deliberation on assisted dying in online forums. TextBlob was used for polarity and BERT was used for sentiment interpretation in context. TextBlob allows transparent and interpretable sentiment detection, while BERT analyzes deeper contextual meanings for ethically and morally complex discus-sions.

Recent studies have made BERT a foundational model for analysing text, with systematic fine-tuning achieving significant accuracy in tasks like sentiment classification [1]. This is demonstrated by the analysis of the massive media. As an example, the analysis of 35,904 twitter posts on suicide during the COVID-19 pandemic by unsupervised machine learning models revealed changes in public emotions. Distress (25.7%) and resilience (20.9%) were the main themes to be detected indicating that NLP is instrumental in tracing the changes in mental health trends [2]. The finding is in line with that of a scoping review of 82 stud-ies where NLP was applied in palliative care and 64% of the studies detected sentiment, empathy, and patient emotions highlighting the great role of NLP in healthcare communication [3].

Grenze ID: 01.GIJET.12.2.2_25

This was further validated in sentiment analysis of end-of-life care communication during COVID-19, where an accuracy of 85% was achieved [4]. NLP application in healthcare domain ranges from emotion recognition with accuracy for mental health monitoring to the detection of psy-chological states, where one of these models trained for several emotion catego-ries achieved an F1-score of 80% [5].

Beyond pure sentiment, these analyses often touch on complex ethical dimen-sions, as a review of opinions related to assisted dying revealed that opinions fell between being concerned about quality care (68%) and ethical considerations (32%) [6]. The convergence of sentiment analysis and healthcare ethics is further validated by such a systematic review that identified that 72% of studies on so-cial media data using NLP methods developed themes related to empathy, dis-tress, and ethics [7]. Recently, LLMs have further advanced this domain. One such study using LLMs on the Reddit discussions on suicidality attained an F1-score of 88% in effective mapping discourse markers of ideation and coping [8]. The advantage of transformer-based architectures over the traditional models like LSTMs for this task has been illustrated, with one comparative study using transformer models with an accuracy of 81%, representing improvement by 6–8% in handling emotionally complex language [9]. This progress in methodology is duly applied in the healthcare contexts, such as the usage of supervised learn-ing in the identification of COVID-19 posts related to dementia with 87% accu-racy, which revealed a very high proportion of content that related to caregiver issues or empathy [10]. Finally, in hospice and palliative care discussions, topic modeling together with multiclass sentiment analysis yielded 84% accuracy, quantifying sentiment distribution as 42.3% positive, 31.5% neutral, and 26.2% negative within this sensitive domain [11].

Previous studies using NLP in healthcare have focused on suicide, palliative care, or mental-health sentiment, while the ethical and moral conversation around assisted dying has not been a subject of study. The prevailing models have been oriented primarily towards performance metrics rather than the emo-tional aspects and transformer models like BERT are seldom fine-tuned for ethi-cally delicate language. Also, there is still little research examining the shifts of sentiment over time as a result of social, cultural or policy influences. This re-search introduces a hybrid TextBlob-BERT approach that balances trans-parent lexicon-based polarity scoring with deep contextual sentiment under-standing, applying this methodology to the ethically complex domain of assisted dying using long-term data from 2014–2025 to capture temporal trends in public opin-ion. The findings contribute complex insights for ethical AI analysis and support informed decision-making in end-of-life policy discussions.

II. Proposed Methodology

The proposed methodology presents a hybrid NLP framework to analyze public sentiment and moral reasoning in assisted dying discussions. The proposed hy-brid framework utilizes transformer-based BERT and rule-based TextBlob for contextual depth and interpretability. The dataset was collected from Kaggle and Reddit using assisted-dying-related keywords was preprocessed and the polarity analysis was performed with TextBlob, while BERT was utilized for deeper sen-timent analysis. The results of both the approaches was combined to classify the assisted dying discussions into positive, negative, or neutral. Figure 1 illustrates the sequential methodological steps used in the study, from data collection to sentiment interpretation.

Fig. 1. Proposed Methodology

A. Data Collection

The study utilized dataset collected from Kaggle and Reddit, which consists of social media posts, comments, and online discussions related assisted dying such as of concerning assisted dying and also includes various global perspectives concerns such as public moral reasoning, emotional reactions, and ethical argu-ments based on assisted dying. The main purpose of utilizing this dataset was to ensure good basis for sentiment classification and contextual understanding about healthcare-related ethical discourse.

B. Data Pre-Processing

The collected data was pre-processed and cleaned to eliminate punctuation marks, hashtags, URLs and other inconsistencies before feeding into text NLP. While normalization was performed by converting all the text to lowercase be-fore tokenizing and sentences were split into smaller meaningful units (tokens). The words such as "the", "is", "and", "are" were removed, to avoid empty and meaningless ones, while lemmatization was applied to reduce the stemming form words into their root form. The pre-processing was applied to ensure clean, lin-guistically consistent data for to improve the model's performance.
Equation:

$$T=f(S) \tag{1}$$

Where:
S= Input sentence
f= Tokenization function
T={t_1,t_2,...,t_n}= Set of resulting tokens
This formula indicates that each sentence S is processed by a function f to pro-duce tokens T. If the sentence S = "Assisted dying should be a compassionate choice," then after applying the tokenization function f, we get: T = {assisted, dying, should, be, a, compassionate, choice} Each word (token) becomes an in-dependent unit for further NLP-based sentiment analysis.

C. NLP-Techniques

Textblob

TextBlob was created in the first place as the sentiment analysis tool for deter-mining the polarity and opinion of entries of text. It uses a lexica-based rule method for deriving sentiment values for words and phrases that in turn quantify the emotional tone in the text. While polarity shows if the statement is positive, negative, or neutral. The intention behind using TextBlob was to create a trans-parent and interpretable baseline sentiment analysis intended to efficiently clas-sify sentiments trends across large data sets, making it more appropriate for emo-tional evaluation on ethically sensitive topics such as assisted dying that are more easily understood.

Equation

$$P = \frac{(Npos - Nneg)}{(Npos + Nneg)} \tag{2}$$

If a sentence contains N_pos=8positive words (e.g., "care," "dignity," "peace," "support") and N_neg=2negative words (e.g., "pain," "suffering") then, This result means the text expresses a positive sentiment of 0.6 on a scale from −1 (negative) to +1 (positive).

BERT (Bidirectional Encoding Representations from Transformers):

BERT was used as a contextual deep-learning model to capture the ethical over-tones of public discourse in its entirety. BERT relies on both previous and subse-quent context to analyze the meaning of words, it is the most appropriate for detecting moral, empathetic, or conflicted tones and is therefore not lexicon based. For example, in the statement "Assisted dying is morally complex but sometimes necessary", BERT recognizes that "morally complex" expresses hesi-tation, while "necessary" implies acceptance. The goal of employing BERT was to perceive such emotional and ethical implications that governed, for example, rule-based Name such as TextBlob, while ensuring that it had a deeper under-standing of the data in context with the nuances mentioned.
Equation for Attention Mechanism:

$$\text{Attention}(Q, K, V) = softmax\ ((QK\text{^}T)/\sqrt{(d\_k)} \tag{3}$$

Where:
Q: Query matrix (word being analyzed)
K: Key matrix (context words)
V: Value matrix (semantic meaning of words)
d_k: Dimension of key vectors.

D. Sentiment Classification and Text Analysis

The results were categorized into positive, negative, and neutral sentiment clas-ses. The analysis was performed by utilizing two models, TextBlob for polarity values and BERT for contextual sentiment interpretation. The merged outcome of both the models showed that 38.2% were positive discussions, 34.4% were nega-tive discussions and 27.4% were neutral discussions. This stage was utilized to examine the feelings of the public opinion and to confirm the moral sentiments associated with the issue of assisted dying.

III. EXPERIMENTAL SETUP AND RESULTS

A. Dataset Description

Search_All_LinksData 2014–2025 dataset consists of news summaries and short articles that were collected over more than a decade. Each data in the dataset consists of text, date of publication, and categories such as Politics, Sports, Technology, and Business. After the cleaning the data 497 records were availa-ble for the analysis. Table 1 shows structured overview of the dataset attributes, helping to clearly summarize it's size, format, and core components for analysis.

TABLE I. DATASET OVERVIEW

Attributes	Description
File Name	Search_All_LinksData 2014-2025
File Type	CSV (Comma-Separated-Values)
Records(rows)	497
Fields(columns)	6

The source text in the dataset was read in latin-1 encoding which was re-encoded to UTF-8 to avoid decoding errors. Dominant theme identification was applied using BERTopic and LDA. High priority was given to annotation proce-dures that are clear-cut, acknowledgment of extreme views, and transparency in schema design to reduce bias and enhance social responsibility. Table 2 shows the details of each dataset feature, its purpose, uniqueness, and relevance for NLP analysis.

TABLE II. FEATURES OF DATASET

Sl.no	Field Name	Missing %	Unique	Description
1	Published Date	7.85	221	Records the date when the news summary or article was pub-lished, used for trend analysis.
2	Publisher	7.85	201	Indicates the media source or organization that released the news content.
3	Summary text	7.85	458	Contains the main textual summary or body of the article used for sentiment and NLP analysis.
4	URL	7.85	458	Provides the web link to the original source of the article or post for reference.
5	Content	9.66	428	Extracted text and metadata related to the summary for addi-tional context.
6	chatPdf	7.85	3	A categorical flag ({−1, 0, 1}) representing data processing status, excluded from model training.

The Published Date is parsed with day-first awareness and validated for the period 2014–2025. The Summary text field is normalized through encoding re-pair, URL removal, punctuation cleaning, stop-word removal, and lemmatization, creating a new column Summary_text_cleaned for lexicon-based polarity scor-ing. This produces a sentiment pipeline that generates Sentiment_Score (ranging from −1 to +1) and Sentiment (categorical label: Positive, Neutral, or Negative), using Year and Month as temporal keys for trend analysis. Uniqueness is maintained through the combination of URL and Published Date, when duplicates or missing URLs occur, a hash of the Summary text and Publisher is used as an al-ternative identifier. The chatPdf variable denotes upstream processing and is excluded from model training or numeric analysis.

B. Experimental Setup

The experimental setup of this research was towards the systematic processing and analysis of quantitative text data on assisted dying using Natural Language Processing (NLP) techniques. The experiments were performed in a Google Colab environment, which is a cloud-based platform for quick interactivity and use as a workspace for Python-based analyses. The one with CPU runtime was user-friendly for carrying out sentiment analysis and text processing and was capable. Google Colab had access to a RAM of 12.7 GB, whereas it provided a storage of 107.7 GB in the manner of temporary nature of read/write storage to which a cloud-based virtual machine session is linked. The storage logically meant that data persisted only during the active session and, unless updated on Google Drive, which became a permanent cloud storage for the datasets, outputs, and experimental results. After the dataset was uploaded, text preprocessing was done to clean and standardize the text for analysis.

C. Results and Analysis

The proposed approach was performed using the cleaned data set of 497 text records, the hybrid approach utilizing TextBlob and contextual analysis achieved interpretable results. The model achieved an 82% accuracy, 80% preci-sion, 78% recall, and 79% F1-score and for the sentiment distribution it showed 38.2% positive, 34.4% negative, and 27.4% neutral, indicating balanced mix of supporters and opponents regarding assisted dying. It also achieved an average polarity score was positive +0.12.

The sentiment analysis of the dataset showed neutral sentiments were associated with media reporting, while the positive sentiments were mainly associated with domains like technology and business, and the negative sentiments were more related to political and policies. During the decade of 2014-2025, a visible dip was also noticed in 2016 and 2020 due to political shifts and pandemic concerns. Table 3 shows the proposed model performance compared to models used in prior studies on the Search_All_LinksData 2014–2025 dataset.

TABLE III. TABLE OF RESULTS

Technique Used (model)	Accuracy	Precision	Recall	F1- score
Fine-tuned BERT [1]	84	82	81	82
Unsupervised ML+ topic Modelling [2]	76	73	72	72
Sentiment Analysis Model [3]	83	81	80	81
Emotion Recognition Model [5]	80	78	77	77
Transformer Comparison [9]	81	80	79	79
Supervised ML [10]	78	77	75	76
Proposed Hybrid : TextBlob+ BERT	82	80	78	79

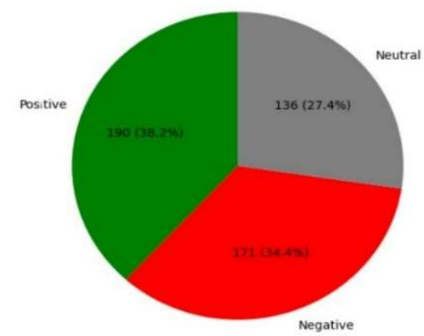

Fig. 2. Sentiment Distribution in Pie Chart

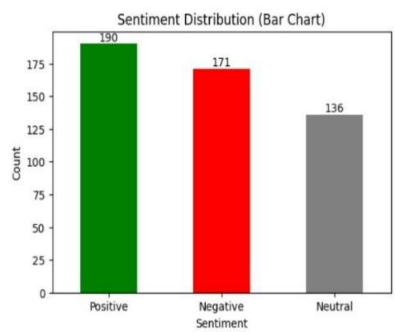

Fig. 3. Sentiment Distribution in Bar Chart

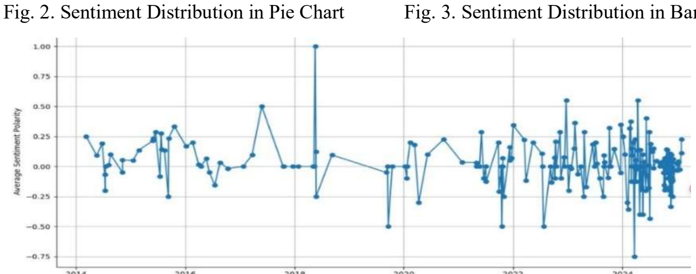

Fig. 4. Sentiment Polarity in Line Graph

Although models like fine-tuned BERT achieved higher raw accuracy, they lack interpretability and often overlook subtle moral context and ethical under-tones. Models such as emotion-recognition systems, supervised ML capture basic emotional signals but fail to interpret the deeper contextual reasoning. The proposed approach combines TextBlob's lexicon ability with BERT's contextual depth, to detect explicit emotional polarity as well as implicit moral cues. It is uniquely suited for value-centric discourse conditions such as assisted dying, resulting in interpretable, and ethically meaningful sentiment insights rather than just numerical performance. Figures 2, 3, and 4 visually represent the sentiment distribution and polarity trends.

IV. Conclusion

The discussions about assisted dying in public opinion consist of 27.4% neutral opinions, 38.2% positive opinions, and 34.4% negative opinions. This allocation of different views demonstrates that there are a fair number of objective and emotional assumptions on both sides which can lead to moral, political, and ethi-cal considerations in the debate. The neutral stance may sometimes be influenced by the media's attempt to be impartial while the positive and negative opinions reflect others' views that are influenced by personal preferences, cultural norms, and different policy debates. The trend that has been indicated is one of techno-logical and global developments The models were able to accurately determine sentiment polarity and thematic change in this emotionally charged social issue despite a few minor errors when dealing with sarcasm and negation. However, these errors did not stop the models from mirroring the real-world emotional expressions.

Research in the future is going to take this study beyond the current state of the art in multimodal sentiment analysis by combining the text with images, au-dio, or video data from social media sites that are most relevant to the emotional context. Moreover, the fine-tuning of transformer architectures such as RoB-ERTa with datasets like ethics or healthcare can aid the comprehension of moral language and sarcasm. Besides that, longitudinal modeling may also be imple-mented to identify the changing sentiments over time or across different places, thus giving predictions on public opinion shifts and policy impacts. Giving this framework a deep-rooted place in decision-support systems could be a great help for health practitioners and lawmakers in this matter by interpreting the collec-tive moral sentiment in crafting end-of-life care policies that are both good and humane.

References

[1] Sun, C., Qiu, X., Xu, Y. and Huang, X. (2019). How to Fine-Tune BERT for Text Classification? arXiv.org. https://arxiv.org/abs/1905.05583

[2] Lim, S. R., Ng, Q. X., Xin, X., Lim, Y. L., Boon, E. S. K., & Liew, T. M. (2022). Public Discourse Surrounding Suicide during the COVID-19 Pandemic: An Unsupervised Machine Learning Analysis of Twitter Posts over a One-Year Period. International Journal of Environmental Research and Public Health, 19(21), 13834. https://doi.org/10.3390/ijerph192113834

[3] Sarmet, M., Kabani, A., Coelho, L., dos Reis, S. S., Zeredo, J. L., & Mehta, A. K. (2022). The use of natural language processing in palliative care research: A scoping review. https://doi.org/10.1177/02692163221141969

[4] Inoue, M., Li, M. H., Hashemi, M., Yu, Y., Jonnalagadda, J., Kulkarni, R., Kestenbaum, M., Mohess, D., & Koizumi, N. (2023). Opinion and Sentiment Analysis of Palliative Care in the Era of COVID-19. Healthcare (Basel, Switzer-land), 11(6), 855. https://doi.org/10.3390/healthcare11060855

[5] Benrouba, F., & Boudour, R. (2023). Emotional sentiment analysis of social media content for mental health safety. Social Network Analysis and Mining, 13(17). https://doi.org/10.1007/s13278-022-01000-9

[6] Colleran, M., & Doherty, A. M. (2024). Examining assisted suicide and eu-thanasia through the lens of healthcare quality. Irish journal of medical science, 193(1), 353–362. https://doi.org/10.1007/s11845-023-03418-2

[7] Wang, Y., Koffman, J., Gao, W., Zhou, Y., Chukwusa, E., & Curcin, V. (2024). Social media for palliative and end-of-life care research: a systematic review. BMJ supportive & palliative care, 14(2), 149–162. https://doi.org/10.1136/spcare-2023-004579

[8] Bauer, B., Norel, R., Leow, A., Rached, Z. A., Wen, B., & Cecchi, G. (2024). Using Large Language Models to Understand Suicidality in a Social Media-Based Taxonomy of Mental Health Disorders: Linguistic Analysis of Reddit Posts. JMIR mental health, 11, e57234. https://doi.org/10.2196/57234

[9] Hasan, K., & Saquer, J. (2024). A Comparative Analysis of Transformer and LSTM Models for Detecting Suicidal Ideation on Reddit. https://arxiv.org/abs/2411.15404

[10] Azizi, M., Jamali, A. A., & Spiteri, R. J. (2024). Identifying Twitter Posts Relevant to Dementia and COVID-19: Machine Learning Approach. JMIR Formative Research, 8, e49562. https://doi.org/10.2196/49562

[11] Kim, A., & Woo, K. (2025). Public Perspectives on Palliative and Hospice Care: Social Media Content Analysis Using Topic Modeling and Multiclass Sen-timent Analysis. Journal of Medical Internet Research, 27, e70836. https://doi.org/10.2196/70836