

Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:
2026-10-09

Deposited version:
Accepted Version

Peer-review status of attached file:
Peer-reviewed

Citation for published item:

Barbosa, L. & Coelho, J. V. (2025). Sentiment analysis of patient complaints in healthcare systems using VADER: Can it contribute to a better service?. In Mateus Mendes, José Torres Farinha, Ana Rita Malta, Hugo Raposo (Ed.), Proceedings of PAMDAS 2025 (pp. 443?453). RCM2+ — Research Centre for Asset Management and Systems Engineering.

Further information on publisher's website:

<https://pamdas.rcm2.pt/event/1/attachments/8/134/PAMDAS-2025-proceedings.pdf>

Publisher's copyright statement:

This is the peer reviewed version of the following article: Barbosa, L. & Coelho, J. V. (2025). Sentiment analysis of patient complaints in healthcare systems using VADER: Can it contribute to a better service?. In Mateus Mendes, José Torres Farinha, Ana Rita Malta, Hugo Raposo (Ed.), Proceedings of PAMDAS 2025 (pp. 443?453). RCM2+ — Research Centre for Asset Management and Systems Engineering.. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Chapter 37

Sentiment analysis of patient complaints in healthcare systems using VADER: Can it contribute to a better service?

Contribution #67

Cite as:

Liliana Barbosa, João Vasco Coelho. Sentiment analysis of patient complaints in healthcare systems using VADER: Can it contribute to a better service?. In *Proceedings of PAMDAS 2025 - International conference on Physical Asset Management and Data Science*, 17-18 Jul. 2025, Coimbra, Portugal. ISBN 978-989-8331-19-9.

```
@inproceedings{pamdass2025-67,  
  title = {Sentiment analysis of patient complaints in  
    healthcare systems using VADER: Can it contribute  
    to a better service?},  
  author = {Liliana Barbosa and João Vasco Coelho},  
  booktitle = {Proceedings of PAMDAS 2025 - International  
    Conference on Physical Asset Management and Data  
    Science},  
  address = {Coimbra, Portugal},  
  month = {July 17/18},  
  year = {2025}  
}
```

Sentiment analysis of patient complaints in healthcare systems using VADER: Can it contribute to a better service?

Liliana da Costa Barbosa ^{1,*}, João Vasco Coelho ²

¹ Polytechnic University of Coimbra, Coimbra - Portugal

² CIES/ISCTE - IUL, Lisbon – Portugal

*Corresponding author:
liliana.barbosa@live.com.pt

Abstract

This study focuses on exploring the effectiveness of using VADER (*Valence Aware Dictionary and sEntiment Reasoner*), a rule-based sentiment analysis tool, in swiftly analyzing patient complaints within healthcare settings aiming to develop overall digital capabilities. Currently, the use of sentiment analysis in the healthcare space is considered to be under-explored. Free-text complaints from patients often contain valuable insights into their experiences. However, analyzing large volumes of textual feedback can be time-consuming and prone to subjective interpretation. Sentiment analysis tools such as VADER can provide meaningful solutions by automatically extracting and quantifying the emotional tone of textual data.

The dataset used in the study comprised the total number and written free-text of patient complaints collected during 2024, in a key surgery service of a 10,000 employee public hospital located in Portugal. In terms of methods, Orange text mining capabilities were intertwined with VADER sentiment analysis. Three key study findings demonstrate the benefits of using VADER to enhance healthcare patient experience management: (a) its ability to balance efficiency and accuracy in sentiment analysis without compromising precision, (b) its role in fostering a culture of *patient-centered* care decision-making, and (c) its support for process optimization and digitalization efforts, namely in terms of case triage and prioritization.

Keywords: Natural language processing (NLP); Sentiment analysis; Text classification; VADER sentiment analysis; BERT analysis; Healthcare; Patient experience; Patient complaints.

1 Introduction

Patient and family complaints are a valuable source of information to provide patient-oriented healthcare. Complaints made by patients and families typically express a request or dissatisfaction regarding health services, and are considered a valuable data source for hospitals to improve the quality of health services [1][2][3].

As healthcare progressively changes from *disease-centered* to *patient-centered*, it is expected that free-text data will be increasingly important as means to understand patient experiences, that can be somehow lost by relying in the more frequently used surveys based in closed-ended questions [2][3]. To support this intent and develop patient-oriented healthcare services, systems and processes need to be improved to collect, organize, analyze and quickly address patient and family complaints, and establish an accurate *patient's voice*, which is proposed to be one of the core building blocks for system building in contemporary complaint management in the healthcare sector [1][2][3].

To extract value from rich free-text data regarding patient experience, establishing trust by devising quicker case resolution loops, is considered to be of key relevance [4][5][6][7][8][9]. Regarding this intent, it

is appropriate to say that the scale and diversity of free-text feedback make it difficult for human agents to swiftly extract and digest actionable information. Sentiment analysis, a pivotal technique in natural language processing (NLP) that involves identifying and extracting subjective information from textual data, can help automate this process by classifying textual feedback into emotional categories such as positive, neutral, or negative [2]. Beyond basic polarity detection, advanced sentiment models can also detect intensity and emotional cues, offering a nuanced understanding of customer grievances [5][6][7][8][9].

The *Valence Aware Dictionary and sEntiment Reasoner* (VADER), a rule-based sentiment analysis tool, is considered to be a reliable solution in this domain [2]. It was initially designed to perform on social media text, but is increasingly being tested and valued as effective in other domains [10][11][12][13]. VADER combines a robust lexicon with heuristic rules to capture contextual nuances in sentiment, making it easy to use and highly accurate [13][14][15][16].

In a context (healthcare) where the use of sentiment analysis, namely in terms of patient experience management support, is considered to be still significantly under-explored [13], this paper aims to explore the extent to which rule-based sentiment analysis tools such as VADER can contribute to more efficient, swift and fair patient complaints resolution by quickly establishing an accurate *patient's voice* [1][2][3], in complex and critical sectors such as healthcare and large-sized organizations aiming to develop their technological capabilities such as public hospitals.

The study's dataset consists of the total number and content of patient complaints (n=63) collected throughout 2024, in one of the key surgery services of a 10,000 employee public hospital located in Portugal. In terms of research methods, the analytical apparatus was based in intertwining the usage of Orange text mining capabilities with its VADER sentiment analysis widget, and in comparing outputs, for accuracy control, with the ones derived a grouped sentiment analysis using BERT (*Bidirectional Encoder Representations from Transformers*), a transformer-based model. The paper is organized as follows: *Section 2* presents a review of the current state of the art, *section 3* presents the dataset and methods that were used, *sections 4* and *5* the research results and respective discussion, including observed challenges and limitations. A conclusion section presenting a vision for future work ends the paper.

2 State of the art

Over the last decade, sentiment analysis, conversational AI and opinion mining have been emerging as critical areas of research, with significant applications across industries such as e-commerce, healthcare, social media, and customer service [2][13][14][15][16]. During this timeframe, sentiment analysis methods, in concrete, have developed significantly, driven by advances in natural language processing (NLP) and machine learning techniques. Current primary methodologies in the sentiment analysis field include [10][11][12][13][14][15][16]:

1. *Rule-based, lexicon-based* approaches that rely on predefined linguistic rules to identify and extract sentiment from text; these approaches typically involve using predefined dictionaries of sentiment-bearing words to determine the sentiment orientation of text; each word in the lexicon is assigned a sentiment score, and the overall sentiment of a text is calculated based on the presence and scores of these words. While these methods are simple and easy to implement, they often struggle with context, sarcasm, and nuanced expressions. Despite the limitations, these methods are still widely used due to their interpretability and low computational requirements;
2. *Machine learning* models trained on labeled datasets, learning patterns in the data to predict sentiments in new, unseen text; these techniques offer greater flexibility and accuracy compared to rule-based approaches;
3. *Deep learning techniques*, such as Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Transformer-based models (e.g., BERT) have demonstrated superior performance in sentiment analysis, by processing complex representations of text that enhance sentiment classification accuracy.

Closely intertwined with the aforementioned developments, sentiment analysis methods have become increasingly explored and applied in industries and domains where understanding customer and user sentiments can drive more effective managerial decision-making and improve overall experiences [4] [5][6][7][8][9]. For instance, businesses have been leveraging sentiment analysis to gain insights into customer feedback, enhancing product offerings and service quality [13]. In healthcare, sentiment analysis has been explored to aid in monitoring patient opinions and emotions, contributing to better patient care and satisfaction [9][16].

As healthcare shifts from a *disease-centered* to a *patient-centered approach*, free-text data is expected to play an increasingly vital role in capturing and promptly handling patient experiences [2][3]. To derive meaningful insights from rich free-text data on patient experiences, build trust and establish faster resolution loops, it is essential to be able to systematize complaints and complaint data related to health services [4] [5][6][7][8][9]. To advance this goal and foster *patient-oriented* healthcare services, systems and processes must be enhanced to efficiently organize and address patient and family complaints. Establishing an accurate *patient voice* is proposed as a fundamental building block in modern complaint management within the healthcare sector. As illustrated by previous studies [3][4][9][13][16], it is frequent to find ambiguity and nuances in free text patient complaints, making it difficult to identify appropriate improvement actions. Complaints caused by hospital staff, as opposed to processes or environment (e.g., doctors, nurses, and other staff) are often more frequent and more reported as complaints triggered by a range of events or attributes (e.g., communication, attitude, nursing skill), with direct comparisons of complaint classification between hospitals and countries being difficult to do [3][4][9].

Despite the recent observed advancements in the healthcare space, sentiment analysis, particularly in patient experience management through free-text data mining, remains significantly under-explored [13]. However, it continues to gain traction as healthcare organizations strive to enhance and optimization the services provided to patients and their families, in particular through process digitalization and automation. The present study aims to contribute to this progress.

3 Methodology

Materials.

The dataset that was used in the study consisted of the total number and content of patient complaints ($n=63$) collected during 2024, in one of the key healthcare services of a 10,000 employee public hospital located in Portugal. Given its size, nature (general surgery) and exposure to be subject considered in the research (patient experience management in healthcare organizations with developing digital capabilities), this healthcare service was positioned as an *exemplary*, pregnant case to address, in empirical terms, the research proposed intents. The reliance of the service on free-text survey data to handle patient complaints was a key factor supporting its suitability to address the research objectives in empirical terms.

Methods.

Data was translated from Portuguese to English, labelled and processed to establish a base *corpus* and perform subsequent sentiment analysis. *Figure 1* presents the dataset labelling per feedback type. Analysis was made by intertwining the usage of Orange text mining capabilities with its VADER sentiment analysis widget. In Orange, VADER is used to analyze English-language text, providing positive, negative, neutral and compound sentiment scores. These scores help users understand the overall sentiment of the established *corpus*. For secondary comparative purposes (in terms of compound sentiment analysis scores accuracy), a grouped sentiment analysis using BERT (used for nuanced and complex sentence structures analysis) [17] was made. The text mining model that was used is presented in *Figure 2*.

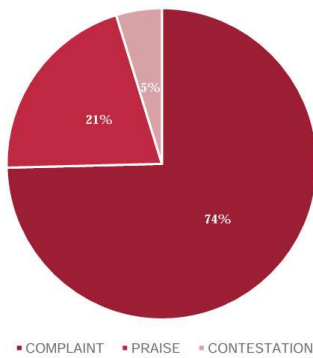


Figure 1. Dataset labelling (per feedback type).

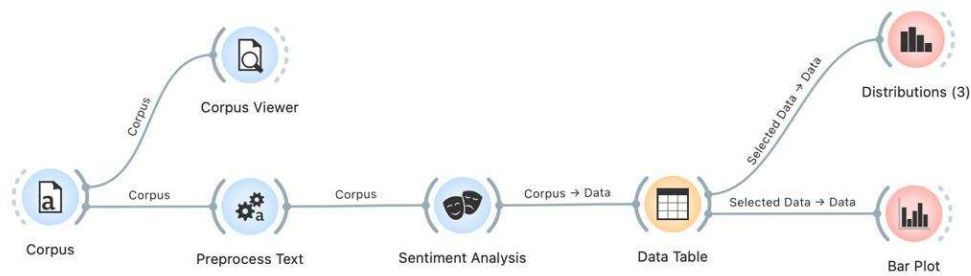


Figure 2. Text mining model (using Orange).

Composite scores were obtained by summing the valence scores of each word in the lexicon, including adjustments based on predefined rules, and normalizing the score to a range between -1 (extremely negative) and +1 (extremely positive). In the classification process, the following criteria was used to determine the sentiment (positive, negative, or neutral) of each comment [14]: *positive sentiment*: compound score ≥ 0.5 ; *neutral sentiment*: (compound score > -0.5) and (compound score < 0.5); *negative sentiment*: compound score ≤ -0.5 .

In terms of analytical results presentation, given the research goal of exploring the potential usefulness or appropriateness of using a sentiment analysis technique and tool as means to digest patient complaints in healthcare settings and support quick resolution actions and/or decisions by key stakeholders, it was decided to split analytical outputs per emotional tone - *negative* sentiments vs. *neutral* or *positive* -, regardless the nature of the feedback being presented by the patients (e.g., a complaint, a compliment or a dispute).

4 Results

Overall, three main research results should be highlighted: 1) the existence of a significant number of neutral and positive scores in complaints free-text suggesting that, although labelled as complaints, the tone isn't necessarily (strongly) emotional; 2) where an overall positive emotional tone was observed (compound score), text labelling was mixed (complaint, praise and/or disputes); 3) to be able to identify and differentiate these nuances in a timely manner can be particularly useful for users and decision-makers to quickly trigger (or not) follow-up case resolution actions.

In a more qualitative perspective, complementing VADER's sentiment analysis output, it was possible to summarize the most frequent content feedback and link it with negative toned feedback (e.g., misdiagnosis and incorrect surgery scheduling, post-surgery pain and neglect, poor communication/updates), neutral feedback (e.g., constructive suggestions of service improvements) and positive feedback (e.g., gratitude for the professionalism and care).

Figure 3 presents a sample of the output that was obtained with VADER that had a negative tone, Figure 4 a output sample where a neutral or a positive tone was observed. In Figure 3, that aggregates negative tone cases, most texts were labelled as complaint, with one case of dispute. Compound scores range from -0.9984 (extremely negative) to -0.0323 (slightly negative), indicating VADER's ability to differentiate the intensity of the sentiment expressed in a complaint. Most complaints have negative compound scores (check Figure 5), reflecting predominantly negative sentiments. The coexistence of neutral and positive scores in complaint texts suggests that, although they are classified as complaints, the tone isn't necessarily (strongly) emotional.

	Text	Label	positive	negative	neutral	compound
15	**COMPLAI...	Complaint	0.038	0.125	0.838	-0.9984
23	**Subject:** ...	Complaint	0.058	0.134	0.808	-0.996
33	**Subject:** ...	Dispute	0.076	0.167	0.757	-0.9948
7	Good Mornin...	Complaint	0.075	0.129	0.796	-0.9921
49	Formal Com...	Complaint	0.056	0.162	0.783	-0.9912
16	To Whom It ...	Complaint	0.05	0.189	0.761	-0.9904
59	**Outeiro Ma...	Complaint	0.064	0.117	0.819	-0.9863
10	**NCC. 998...	Complaint	0.049	0.16	0.791	-0.9861
4	Suspected M...	Complaint	0.063	0.102	0.835	-0.9857
12	**To Whom I...	Complaint	0.098	0.129	0.773	-0.9842
17	Urgent Com...	Complaint	0.061	0.11	0.83	-0.9784
31	**Subject:** ...	Complaint	0.071	0.082	0.847	-0.97
40	Dear Sir/Mad...	Complaint	0.122	0.137	0.741	-0.9692
45	Hello, good a...	Complaint	0.033	0.067	0.901	-0.9682
32	**Subject:** ...	Complaint	0.079	0.119	0.802	-0.9584
2	I underwent ...	Complaint	0.042	0.158	0.8	-0.9555
36	Good aftern...	Complaint	0.062	0.093	0.845	-0.9555
1	I, Maria, a pa...	Complaint	0.055	0.116	0.829	-0.9544
39	**Case Num...	Complaint	0.016	0.049	0.935	-0.9523
26	**Subject:** ...	Complaint	0.049	0.087	0.864	-0.9458
43	Dear Sir/Mad...	Complaint	0.051	0.071	0.879	-0.8373
57	Dr. José Gul...	Complaint	0.084	0.097	0.819	-0.8305
37	I've been wal...	Complaint	0	0.169	0.831	-0.7579
51	Dear Sirs, I a...	Complaint	0.045	0.06	0.895	-0.7556
3	My complain...	Complaint	0	0.067	0.933	-0.7227
8	Good aftern...	Complaint	0.113	0.127	0.759	-0.712
25	The complai...	Complaint	0	0.113	0.887	-0.6705
38	I went to the ...	Complaint	0.034	0.092	0.874	-0.6483
44	I've been wal...	Complaint	0	0.065	0.935	-0.3506
22	**Subject:** ...	Complaint	0.075	0.076	0.849	-0.0451
9	Dear HUC, ...	Complaint	0.069	0.086	0.846	-0.0323

Figure 3. VADER output data table with negative emotional tone (compound score).

In Figure 4, despite overall positively toned answers, text labelling is mixed - complaint, praise or disputes were observed. Compound scores range from 0 (neutral) to 0.9994 (extremely positive), indicating the intensity of the sentiment. The negative scores are generally low, with most values close to 0, indicating minimal negative sentiment. Neutral scores are high (ranging from 0.577 to 1), suggesting that many texts have a neutral tone. Some of these cases likely do not require the immediate direct intervention of a healthcare service member.

	Text	Label	positive	negative	neutral	compound
63	**Good morn...	Praise	0.264	0.046	0.69	0.9994
19	**Subject:** ...	Praise	0.173	0.027	0.8	0.9981
54	Good evenin...	Praise	0.289	0	0.711	0.9961
35	Dear Sirs,I a...	Praise	0.282	0.012	0.705	0.9955
56	**Dear Sirs,*...	Complaint	0.115	0.074	0.811	0.9938
6	Subject: Con...	Praise	0.237	0	0.763	0.9924
30	Luis Joaquin...	Praise	0.311	0	0.689	0.9869
53	Dear Sir,Hea...	Praise	0.423	0	0.577	0.9863
34	**Subject:** ...	Dispute	0.106	0.04	0.854	0.9858
55	Dear Profess...	Praise	0.28	0	0.72	0.9773
42	I would like t...	Praise	0.416	0	0.584	0.976
58	**To the Citiz...	Complaint	0.092	0.014	0.893	0.9698
52	Dear Dr. Alex...	Praise	0.217	0.022	0.761	0.9595
61	To the entire ...	Praise	0.231	0.04	0.728	0.9504
62	Dear Profess...	Praise	0.394	0	0.606	0.9423
13	Good aftern...	Complaint	0.125	0	0.875	0.936
5	Good mornin...	Complaint	0.229	0.037	0.734	0.9235
20	**Subject:** ...	Complaint	0.106	0.013	0.881	0.9197
21	**Subject:** ...	Complaint	0.161	0	0.839	0.9125
28	Request for ...	Complaint	0.141	0	0.859	0.8916
48	Exmo. Sirs,I ...	Dispute	0.068	0.054	0.878	0.8799
60	I would like t...	Praise	0.234	0	0.766	0.8658
27	Complaint R...	Complaint	0.138	0.085	0.777	0.7707
18	Urgent Com...	Complaint	0.126	0.1	0.775	0.7523
46	Dear Sir/Mad...	Complaint	0.061	0.044	0.895	0.743
11	Whom it May...	Complaint	0.103	0.081	0.816	0.7261
50	Good mornin...	Complaint	0.095	0.064	0.841	0.6871
14	**To the Dire...	Complaint	0.121	0.099	0.781	0.5759
24	**Subject:** ...	Complaint	0.046	0.028	0.926	0.5565
41	To the Estee...	Complaint	0.052	0.04	0.908	0.4776
47	Good mornin...	Complaint	0.064	0.051	0.885	0.2732
29	They called ...	Complaint	0	0	1	0

Figure 4. VADER output data table with a neutral or positive emotional tone (compound score).

Figure 5 presents the compound scores obtained with the use of VADER: as mentioned, the existence of a significant number of neutral or positive scores for texts labelled as complaints, suggests that, despite the classification, the tone isn't necessarily (strongly) emotional.

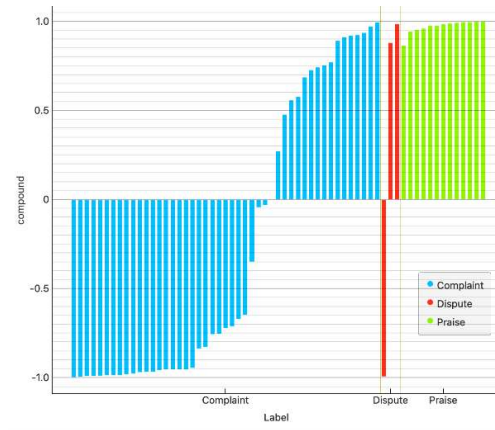


Figure 5. VADER output data table – compound scores.

For analytical control and output comparison purposes, BERT was used to double-check the accuracy of the results achieved with VADER. BERT grouped results are presented in Figure 6.

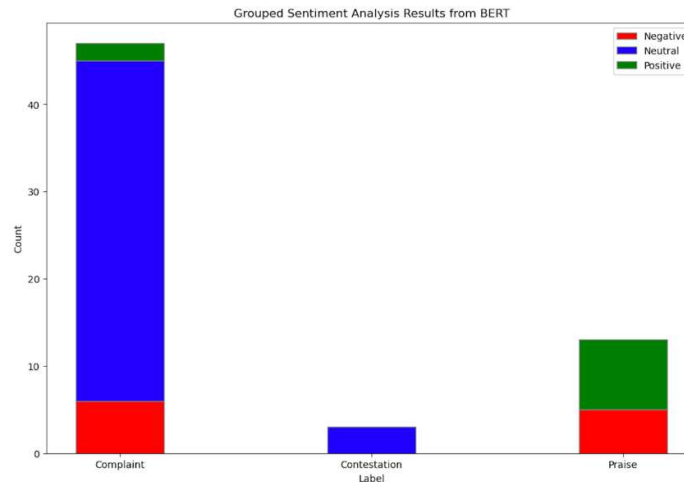


Figure 6. BERT output data table – grouped sentiment analysis.

Overall, as illustrated in the literature [2], the overall orientation of VADER and BERT results can be considered to be similar, namely the existence of a significant number of neutral and positive scores in complaints free-text. VADER analysis, given its rule-based nature, is more straightforward in terms of legibility and call-to-action, but the chance of missing subtle sentiments or contextual-dependent details exists (e.g., a sarcastic comment can be classified as positive; overestimation of neutral or positive sentiments in ambiguous cases can be observed). BERT results are likely to be more accurate, especially for nuanced or sarcastic text. A sarcastic comment classified as neutral by VADER is likely to be classified as negative by BERT, thus providing approach to analytical accuracy [18].

5 Discussion

As illustrated by the present research, the use of sentiment analysis techniques and tools such as VADER to support patient experience management in healthcare settings with developing digital capabilities seems to present significant potential to contribute to transform free-text patient feedback into actionable, prompt insights. As illustrated, when analyzing informal free-text and requiring quick results, VADER is able to provide overall meaningful output, namely if broader, helicopter view or snapshot perspectives are required.

In terms of research results, the research provides an illustration of how VADER sentiment analysis can support the extraction of meaningful insights from rich free-text data on patient experiences, offering a view of customer grievances. As illustrated by past studies [4] [5][6][7][8][9], VADER human-centric design enhances sentiment understanding, making it a potentially valuable tool in healthcare-related analytics, going beyond basic polarity detection, by detecting the intensity of emotional cues. This makes it particularly useful for applications and domains where interpretability and speed are crucial, as is the case of patience experience management and the growing need to craft an accurate *patient's voice* in healthcare settings [1][2][3].

Beyond reinforcing previous studies findings [2][3][13][15][16], it is worthwhile highlighting, in particular, three concrete aspects provided by the present study that suggest the appropriateness of deepening and expanding the usage of VADER into the patience experience management space:

1. *Efficiency – effectiveness (accuracy) balance*: VADER's rule-based approach offers a quick and computationally efficient method for sentiment analysis without sacrificing accuracy. This is particularly beneficial in healthcare settings and patient experience handling where timely, easy-to-understand insights are of paramount importance;
2. *Support to patient-centric healthcare services*: by analyzing patient complaints, healthcare providers

can gain a deeper understanding of patient experiences and identify areas needing improvement. This aligns with the broader trend of *patient-centered* care, which emphasizes the importance of considering patient feedback in healthcare service delivery;

3. *Process optimization support (e.g., case prioritization)*: the sentiment analysis made in the study revealed that most complaints had a negative tone, pinpointing concrete areas where patients frequently encounter problems; this can help healthcare providers prioritize these issues and allocate resources more effectively.

While VADER-based sentiment analysis may offer valuable, actionable insights to patient feedback handling, three challenges and limitations need to be considered:

1. *Dealing with nuances and ambiguity*: Sentiment analysis models tend to struggle with ambiguous expressions where the sentiment is not clear-cut. This can lead to misclassification and reduced accuracy. In the present study, the existence of a significant number of neutral and positive scores in complaints free-text signals possible advantages in using a deeper integration of different models in future work, hybridizing traditional techniques with deep learning models. Developing model capabilities to understand nuanced opinions and subtle sentiments remains a significant challenge;
2. *Unstructured data and domain adaptation*: Models trained on specific datasets often perform poorly when applied to different domains. To this regard, in the present study, a significant challenge was dealing with unstructured data, often containing noise, slang, abbreviations, and inconsistent use of language. Effective preprocessing methods, such as tokenization, stemming, lemmatization, and noise removal, are essential to enhance the quality of the data and improve analysis outcomes;
3. *Contextual dependency (for interpretation)*: Understanding the context in which a sentiment is expressed is crucial for accurate analysis. In the present study, following existing literature [17][18], the use of BERT, with comparative purposes, highlighted the relevance of considering the surrounding context to correctly interpret sentiments, especially in complex sentences.

Continuous advancements in sentiment analysis require maintaining a laser focus on developing innovative approaches to address existing gaps and challenges. As mentioned, future work will include a deeper integration of different models, hybridizing traditional techniques with deep learning models, a way to balance the limitations of each approach. For example, rule or lexicon-based methods can be used to provide initial sentiment scores, subsequently refined using deep learning models. This hybrid approach will help improve accuracy by incorporating both rule-based knowledge and pattern recognition capabilities of deep learning models. Additionally, integrating syntactic and semantic features extracted from traditional methods with embeddings learned by deep learning models can enhance the overall performance of sentiment analysis systems, namely to better understand and detect textual subtle and context-dependent nuances.

6 Conclusion

The present study illustrates the effectiveness and potential usefulness of using VADER sentiment analysis to analyze patient complaints in healthcare settings. Three key study findings suggest the advantages of using VADER to support healthcare patient experience management: a) the efficiency-effectiveness (accuracy) balance offered by VADER, by delivering quick and computationally efficient sentiment analysis without sacrificing accuracy; b) support given to patient-centered care; and c) support given to process optimization intents (via, for example, sentiment data-infused case prioritization). Future work will focus on integrating advanced machine learning models to enhance the accuracy and depth of sentiment analysis, as demonstrated by the comparison with the BERT sentiment analysis model. Additionally, expanding the dataset to include cross-sectional feedback from multiple departments and hospitals will provide a broader and more comprehensive perspective on patient sentiments across healthcare systems, helping to identify recurring issues and prioritize areas for improvement.

Acknowledgments

The authors would like to thank the healthcare organization included in the study for providing access to the data used in this research.

References

- [1] A. Borg and M. Boldt, "Using VADER sentiment and SVM for predicting customer response sentiment," *Expert Syst Appl*, vol. 162, 2020, doi: 10.1016/j.eswa.2020.113746.
- [2] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, 2014. doi: 10.1609/icwsm.v8i1.14550.
- [3] R. Harrison, M. Walton, J. Healy, J. Smith-Merry, and C. Hobbs, "Patient complaints about hospital services: Applying a complaint taxonomy to analyse and respond to complaints," *International Journal for Quality in Health Care*, vol. 28, no. 2, 2016, doi: 10.1093/intqhc/mzw003.
- [4] C. D. Mann, J. A. Howes, A. Buchanan, and D. J. Bowrey, "One-year audit of complaints made against a University Hospital Surgical Department," *ANZ J Surg*, vol. 82, no. 10, 2012, doi: 10.1111/j.1445-2197.2012.06240.x.
- [5] L. Donaldson, "An organisation with memory," *Clin Med (Lond)*, vol. 2, no. 5, 2002, doi: 10.7861/clinmedicine.2-5-452.
- [6] S. Lloyd-Bostock and L. Mulcahy, "The Social Psychology of Making and Responding to Hospital Complaints: An Account Model of Complaint Processes," *Law Policy*, vol. 16, no. 2, 1994, doi: 10.1111/j.1467-9930.1994.tb00120.x.
- [7] L. Mulcahy and J. Q. Tritter, "Pathways, pyramids and icebergs? Mapping the links between dissatisfaction and complaints," *Sociol Health Illn*, vol. 20, no. 6, pp. 825–847, 1998, doi: 10.1111/1467-9566.00131.
- [8] S. Carmel, "Patient complaint strategies in a general hospital," *Hosp Health Serv Adm*, vol. 35, no. 2, 1990.
- [9] T. W. Reader, A. Gillespie, and J. Roberts, "Patient complaints in healthcare systems: A systematic review and coding taxonomy," 2014. doi: 10.1136/bmjqs-2013-002437.
- [10] R. Socher et al., "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank," in *Conference on Empirical Methods in Natural Language Processing*, 2013. [Online]. Available: <https://api.semanticscholar.org/CorpusID:990233>
- [11] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment Analysis of Twitter Data," *Proceedings of the Workshop on Languages in Social Media*, Jan. 2011.
- [12] H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan, *A System for Real-time Twitter Sentiment Analysis of 2012 US Presidential Election Cycle*. 2012.
- [13] Y. Mao, Q. Liu, and Y. Zhang, "Sentiment analysis methods, applications, and challenges: A systematic literature review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no.

4, p. 102048, 2024, doi: <https://doi.org/10.1016/j.jksuci.2024.102048>.

- [14] M. Hoti and J. Ajdari, "Sentiment Analysis Using the Vader Model for Assessing Company Services Based on Posts on Social Media," *SEEU Review*, vol. 18, pp. 19–33, Dec. 2023, doi: 10.2478/seeur-2023-0043.
- [15] D. Youvan, *Understanding Sentiment Analysis with VADER: A Comprehensive Overview and Application*. 2024. doi: 10.13140/RG.2.2.33567.98726.
- [16] G. Al Qarana, L. Rianto, and L. Purnomo, "Analyzing Perceptions of Maternal and Pediatric Care in Jakarta: An Integrated VADER and GloVe Analysis of Google Reviews in Mother and Child Hospitals." [Online]. Available: www.joiv.org/index.php/joiv
- [17] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 2019.
- [18] A. de L. Languré and M. Zareei, "Improving Text Emotion Detection Through Comprehensive Dataset Quality Analysis," *IEEE Access*, vol. 12, pp. 166512–166536, 2024, doi: 10.1109/ACCESS.2024.3491856.