Journal of Information Systems Engineering and Management
2025, 10(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Battle of Sentiment Lexicons: Wordnet, Sentiwordnet, Textblob and Vader in Web Forum Analysis

Ahmad Taufik Nursal[1], Mohd Faizal Omar[2*], Mohd Nasrun Mohd Nawi[3], Mohamad Sukeri Khalid[4], Mohd Hanizun Hanafi[5], Rafikullah Deraman[6]

[1]*Faculty of Industrial Management, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuhraya Tun Razak, 26300, Gambang, Pahang, Malaysia. ahmadtaufik@umpsa.edu.my*

[2]*School of Quantitative Science, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia. faizal_omar@uum.edu.my*

[3]*Disaster Management Institute (DMI), School of Technology Management and Logistic, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia. nasrun@uum.edu.my*

[4]*Institute of Local Government Studies, School of Government, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia. sukeri@usm.my*

[5]*Faculty of Architecture and Ekistics, Universiti Malaysia Kelantan, 16300, Bachok, Kelantan, Malaysia. hanizun.h@umk.edu.my*

[6]*Project and Facilities Management Research Group, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia. rafikullah@uthm.edu.my*

ARTICLE INFO

Received: 28 Sept 2024
Revised: 30 Nov 2024
Accepted: 15 Dec 2024

ABSTRACT

Introduction: Sentiment analysis plays a crucial role in assessing public perception across various sectors, including the high-rise property industry. Understanding public sentiment provides essential insights for stakeholders, enabling data-driven and informed decision-making. This study examines the applicability and effectiveness of lexicon-based sentiment analysis tools in measuring public sentiment toward high-rise properties in Malaysia.

Objectives: The study aims to investigate the effectiveness of four lexicon-based sentiment analysis tools, namely WordNet, TextBlob, SentiWordNet, and Valence Aware Dictionary and Sentiment Reasoner (VADER). It seeks to evaluate their classification performance and accuracy in identifying positive, negative, and neutral sentiments expressed in public reviews.

Methods: The research employed a real-world case study to analyze public sentiment toward high-rise properties. Reviews were classified using the four lexicon dictionaries, and their performance was assessed by comparing their ability to identify positive, negative, and neutral sentiments.

Results: The analysis revealed that all four lexicon-based tools identified a higher proportion of positive reviews than negative or neutral ones. WordNet recorded the largest number of positive sentiments, closely followed by SentiWordNet and VADER. When identifying negative sentiments, VADER emerged as the most effective, followed by WordNet and SentiWordNet. For neutral sentiments, VADER detected the highest number of reviews, while SentiWordNet and WordNet identified fewer instances. Overall, VADER demonstrated superior performance compared to the other tools. However, the study also highlighted limitations in VADER's performance, such as restricted vocabulary, difficulty in detecting sarcasm, ambiguity, and challenges with misspelled or short terms, which occasionally led to misclassifications.

Conclusions: This study provides valuable insights for researchers, practitioners, and policymakers involved in analyzing public sentiment toward high-rise properties in Malaysia. By understanding the strengths and limitations of lexicon-based sentiment analysis tools, stakeholders can enhance their decision-making processes through more precise sentiment classification

Keywords: Text Analysis, Sentiment Analysis, Housing, Lexicon, NLP, Decision Making.

INTRODUCTION

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The advancement of Information Technology (IT) has led to transformative changes in how individuals express their opinions and experiences towards various subjects, including services, products, events, companies, and individuals [1]–[3]. Web 2.0 platforms, such as microblogging, online forums, blogs, and review websites, have emerged as effective channels for opinion sharing, information dissemination, and idea propagation among online users. The information shared through these platforms is considered authentic feedback, distinct from marketing-oriented information. The pervasive influence of these platforms has significantly increased across almost all sectors of modern society. Consequently, there has been an explosive growth in user-generated content. Daily, millions of posts covering diverse topics, including product reviews, are generated on online forums, Facebook, Twitter, blogs, YouTube, Flickr, Google reviews, and other social networking platforms. As of early 2024, there are approximately 28.68 million social media users in Malaysia, representing 83.1% of the total population. This reflects a significant growth, with an increase of 4.8 million users between 2023 and 2024, marking a 20% rise. The most popular social media platforms include WhatsApp (90.7% of users), followed by Facebook (84.9%), Instagram (77%), and TikTok (68.8%). TikTok has been experiencing rapid growth, with a 48.6% increase in ad reach during the past year. This information serves as a valuable and rich resource not only for online users but also for companies, governments, and other interested parties seeking to extract and analyze public opinions [4]–[6].

The availability of this information allows for the investigation of people's behavior and their interactions on specific topics. However, extracting and analyzing this vast amount of user-generated information poses significant challenges due to the substantial volume of opinionated text found on each platform. Manual analysis by average human readers would be impractical. Therefore, automated methods are required to analyze and summarize this information. Sentiment analysis, often referred to as opinion mining, is a research area dedicated to examining individuals' opinions, sentiments, evaluations, appraisals, attitudes, and emotions toward various subjects, such as products, services, organizations, individuals, issues, or events. It is also characterized as a task within Natural Language Processing (NLP) and information extraction, which seeks to identify whether the sentiment conveyed by the author is positive or negative [7]–[9]. Sentiment Analysis (SA) has attracted significant attention from academic researchers and industry practitioners due to its potential value in the decision-making process. Numerous applications of SA have been proposed to extract insights into public opinions, serving as decision-making tools across various domains, including business, manufacturing, government, healthcare, real estate, and policy-making [10]–[13].

One of the most broadly proposed approaches for sentiment analysis in past studies is lexicon-based approaches. Numerous lexicon dictionaries have been developed and successfully utilised to cater to a wide range of sentiments in various fields. SA has been studied in a multilingual context, encompassing languages such as English-Spanish [18], [19], English-Hindi [20] and Chinese-English [21]. Several studies have compared the effectiveness of different lexicons in conducting sentiment analysis on text [14-17]. However, most of this research focuses on general lexicons, which often overlook the widespread use of internet slang commonly used by online users. Furthermore, platforms associated with Web 2.0, such as microblogs, impose word limitations, further complicating the use of traditional lexicons. Moreover, it is challenging to determine which dictionary is superior across various scenarios, or whether it can provide the same level of coverage as certain lexical approaches. Each lexicon has been developed using distinct methodologies, possesses unique characteristics, and produces varying outcomes when applied to different types of text. This paper seeks to evaluate and compare the performance of WordNet, SentiWordNet (SWN), TextBlob, and the Valence Aware Dictionary and Sentiment Reasoner (VADER) in conducting sentiment analysis on online reviews of high-rise properties, gathered from online forums and Google Reviews. The outcomes of this study aim to provide a comprehensive understanding of which lexicon is most suitable for sentiment analysis, particularly in contexts where the text incorporates informal language, such as slang.

SENTIMENT ANALYSIS WITHIN THE HOUSING INDUSTRY CONTEXT.

The growth of Web 2.0 platforms, such as review websites, social media, and online forums, has spurred increasing interest in sentiment analysis (SA) across various industries, including tourism, politics, and the stock market [22]-[25]. While sentiment analysis has been widely applied in these sectors, its use in the housing market is still emerging. Research has highlighted the importance of emotions and beliefs, in addition to economic factors, in shaping housing investment and purchasing decisions [26], [27]. Online platforms and user-generated content provide valuable insights into consumer sentiment, offering a more cost-effective and scalable alternative to traditional surveys [16]–[28]. Several studies have explored sentiment analysis in the housing sector, utilizing both professional datasets (e.g., newspapers) and user-generated content (e.g., online reviews). Research has shown that sentiment extracted from news articles and online reviews can significantly influence housing market dynamics, affecting factors such as price fluctuations and investment trends [5] [29], [30]. Recent studies have also focused on user opinions in housing-related decision-making models, incorporating factors like neighborhood perceptions, educational environments, and the impact of noise on property prices. These findings underscore the growing importance of sentiment analysis in understanding housing market trends and aiding stakeholders in decision-making [7], [21], [22].

LEXICON APPROACHED

The lexicon-based approach relies on a dictionary or lexicon that includes opinionated words with specific polarity. A sentiment lexicon consists of words categorized by their polarity as positive, negative, or neutral. The development of these lexicons generally follows one of two methods: the dictionary-based approach or the corpus-based approach. The dictionary-based approach generally utilizes existing dictionaries, such as VADER, to extract sentiment words. In contrast, the corpus-based approach involves identifying opinion words within a large corpus [32]. This method is used in two main scenarios: to discover new sentiment words from a domain-specific corpus using a pre-existing list of known opinion words, and to create a new sentiment lexicon [32]. However, the primary drawback of this approach is its need for a comprehensive corpus containing all English words [33], which makes it less effective compared to the dictionary-based approach. Conversely, the dictionary-based approach employs a lexical database with opinion words to extract sentiment. Some of the most popular sentiment lexicons in this category include SentiWordNet [34], [35], and VADER [11].

In WordNet, nouns, verbs, adjectives, and adverbs are grouped into sets of cognitive synonyms called ‘synsets’, each expressing a distinct concept. A synset is a set of words with the same part-of-speech that can be interchanged in a certain context. The WordNet 3.0 currently consist over 117,000 synsets, comprising over 81,000 noun synsets, 13,600 verb synsets, 19,000 adjective synsets and 3,600 adverb synsets [36]. Synsets can be related to each other by semantic relations, such as hyponymy. SentiWordNet (SWN) is a lexicon dictionary based on the Wordnet lexicon. The SentiWordNet has been widely adopted due to its extensive coverage of opinion words which contain almost every emotional word in the English language. SentiWordNet has been developed in three versions; version 1.0, version 2.0, and the latest 3.0 [24]. This dictionary blends three scores with a WordNet dictionary synset to determine positive, negative, or neutral text polarity. The synsets listed in SentiWordNet 3.0 consist of three values, namely positive (P), negative (N) and objective (O).

The sum of these values is equal to 1. For instance, the synset for estimable have the value of P = 0.65, N = 0, O = 0.35. The determination of text polarity is based on PosScore (positive score) and NegScore (negative score). This lexicon consists of a set of most commonly used 147,306 synsets. However, SentiWordNet 3.0 is very sensitive to noise in datasets, especially in social media datasets. In addition, it mainly consists of general lexicons and does not consider internet slang and emoticons. Valance Aware Lexicon and Sentiment Reasoner (VADER) is a lexicon, and rule-based sentiment dictionary specifically developed for sentiment in the social media context [11]. Common abbreviations in social media such as acronyms (LOL), slang, and emoji such as smile indicate a positive emotion. This dictionary was developed by examining and selecting features from other lexicon dictionaries such as Linguistic Inquiry and Word Count (LIWC), Affective Norms for English Words (ANEW), and General Inquirer (GI). As a result, VADER consists of 7500 features. In past studies, Vader has produced high accuracy results when dealing with social media text, movie reviews, and product reviews. In addition, VADER has shown a few advantages as dictionary-based sentiment; it works perfectly on the social media type of text, does not require any training data, supports emoji for sentiment classification, and has high-speed processing time [15]. The lexicon-based methods are also very competitive because they require few efforts in the human-label document and not a required training dataset. Table 1 illustrates the application of the lexicon approach in numerous industry domains.

Table 1: The Application of The Lexicon Approach

Author	Objective	Domain	Data Set	Lexicon Method
Al-Shabi, (2020)	Comparison of Lexicon approach	Financial	Twitter	• VADER • SentiWordNet, • SentiStrength • Liu and Hu opinion lexicon • AFINN
Medagoda et	Sentiment Lexicon Construction	General	Online	• SentiWordNet 3.0

al., (2015)	Using SentiWordNet 3.0		Newspaper	
Bonta et al., (2019)	Measure public sentiment on movie review	Entertainment	Review website	• NLTK • Text blob and • VADER
Ikoro et al., (2018)	To improve the accuracy of the Lexicon approach	Energy	Twitter	• Sentiment R • Hu & Liu opinion lexicon
Vu & Le, (2017)	Proposed a lexicon-based method using sentiment dictionaries with a heuristic method for data pre-processing.	General	Amazon, IMDb, and Yelp	• SentiWordNet 3.0 • Hu & Liu opinion lexicon
Chalothorn & Ellman, (2013)	Comparison between SentiWordNet and SentiStrength	Terrorist	Online forums	• SentiWordNet • SentiStrength
Neidhardt et al., (2017)	To assess user engagement and conduct sentiment analysis within an online travel forum.	Tourism	Online forums	• SentiWordNet
Park (2020)	Evaluation of customer satisfaction	Cosmetic	Reviews website	• SentiWordNet 3.0
Khattak et al. (2020)	To measure customer satisfaction based on novel fuzzy sentiment analysis	Drug, electronic and mobile phone	Twitter and Facebook	• Sentiword Net 3.0
Ruscheinsky et al (2018)	Determine the broader relationship between news media sentiment, US with real estate market	Real estate	Online News	• Harvard General Inquirer (GI)

METHODS

The dataset was extracted from an online property forum in Malaysia and the Google review site. It comprises textual reviews for a specific housing project, focusing on entries from 2019 to 2022. These reviews pertain to three high-rise residential condominium projects in Malaysia. The polarity of the collected data was analyzed utilizing three distinct lexicon-based sentiment analysis methodologies: WordNet, SentiStrength, TextBlob and VADER. The construction of the sentiment analysis model commenced with tokenizing sentences into individual words and eliminating high-frequency stopwords. Subsequently, the words were represented in a bag-of-words (BOW) format, and part-of-speech (POS) tagging was employed to ascertain the syntactic role of each word within the sentences. The lexicons WordNet, SentiWordNet, TextBlob and VADER were then applied to assign positive and negative sentiment scores to each synset associated with the words.

Data Pre-processing

In sentiment analysis, data cleaning and preprocessing are often critical steps to ensure accuracy, although tools like VADER can sometimes bypass these processes due to their ability to handle factors such as capitalization, punctuation, and colloquial expressions. In contrast, tools such as TextBlob, SentiWordNet, and WordNet necessitate rigorous preprocessing to achieve reliable outcomes. The pre-processing phase is crucial for cleaning the data and eliminating unnecessary symbols to ensure accurate sentiment analysis. Initial pre-processing of raw data is essential to prevent inaccuracies and misleading results. In this study, several pre-processing activities were conducted, including tokenization, conversion to lowercase, spell-checking, stemming, filtering of stopwords, and the removal of symbols and punctuation. Additionally, words were tokenized and tagged. For TextBlob, many preprocessing tasks can be streamlined through the use of the NLTK (Natural Language Toolkit) library. A common initial step involves converting all text to lowercase, thereby maintaining consistency across the dataset. For example, "House" is standardized to "house" to simplify analysis. However, while this step is generally useful, it may not always be appropriate, particularly in cases where proper capitalization is crucial for tasks such as named entity recognition or enhancing translation accuracy. The removal of stop words—common words that carry minimal meaning—is subsequently performed using the nltk.corpus library. This is followed by tokenization, which divides the text into smaller units, or tokens, by excluding punctuation and repetitive words that could distort the analysis. NLTK's pre-trained language algorithms, such as RegexpTokenizer(r'\w+'), are employed to efficiently split sentences into words, excluding punctuation marks. The next step in preprocessing involves Part-of-Speech (POS) tagging and lemmatization. POS tagging identifies the grammatical role of each word (e.g., noun, verb, adjective) within the text. After this process, lemmatization is performed to convert words to their base or root forms, for example, changing "attempting" to "attempt".

This step ensures that different forms of a word are treated consistently throughout the analysis. Lemmatization, which uses WordNet for mapping, is more precise than simple stemming, as it takes the word's context into account and effectively manages abbreviations and spelling variations. A notable challenge in natural language processing (NLP) is dealing with abbreviations and acronyms, which can complicate analysis and consume considerable resources. Abbreviations increase storage requirements, slow down search processes, and may not always contribute to the analysis. Therefore, normalization becomes crucial to convert abbreviations and clean noisy data, especially when analyzing informal text from platforms like Google review and online forum. Normalization addresses issues of inconsistent vocabulary and abbreviations, thus improving the quality and accuracy of sentiment analysis. The final stage of data preprocessing is noise removal, which involves filtering out extraneous elements such as figures, domain-specific symbols, and code snippets that are irrelevant to the sentiment analysis. This step ensures that the dataset is free from unnecessary information and ready for accurate processing. These comprehensive pre-processing steps were meticulously applied to enhance data quality. Ultimately, a total of 5,340 reviews in the shape of raw data and only a total of 2,631 reviews consisting of sentiment were prepared for subsequent analysis. Figure 1 outlines the stages involved in the data preprocessing workflow.

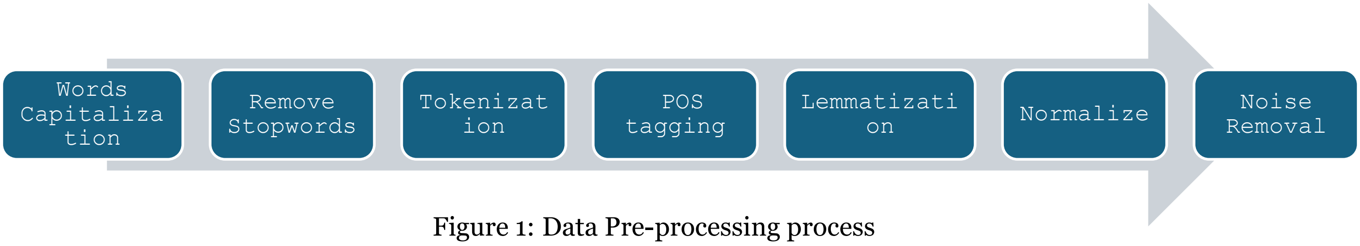

Figure 1: Data Pre-processing process

Sentiment Classification

The next step is to identify the polarity of each review. The lexicon-based model is developed using RapidMiner software, as shown in figure 4.3. The three most used lexicon dictionaries, Wordnet, SentiWordNet, and Valence Aware Dictionary for sentiment Reasoning (VADER), have been conducted. The sentiment score for each review is computed and classified according to specific conditions where;

if(sentiment<0, "Negative",if(sentiment>0 && sentiment !=0, "positive", "Neutral"))

Data Annotation

To ensure the high accuracy of sentiment classification by lexicon dictionary in the previous process, the whole dataset was annotated by humans in this study. Two annotators were proficient in English and well-knowledgeable in the property domain. In this study, proper instruction as a guideline was given to the annotators as in Appendix A. The annotated dataset was used as a baseline for the validation of each dictionary lexicon classification result. The annotators were tasked with assigning a polarity to each review in the dataset, using two categories: positive and negative. The evaluation of the system's performance in this study was based on four key metrics: precision, recall, F-score, and accuracy. Additionally, the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values were computed to further assess the system's effectiveness.

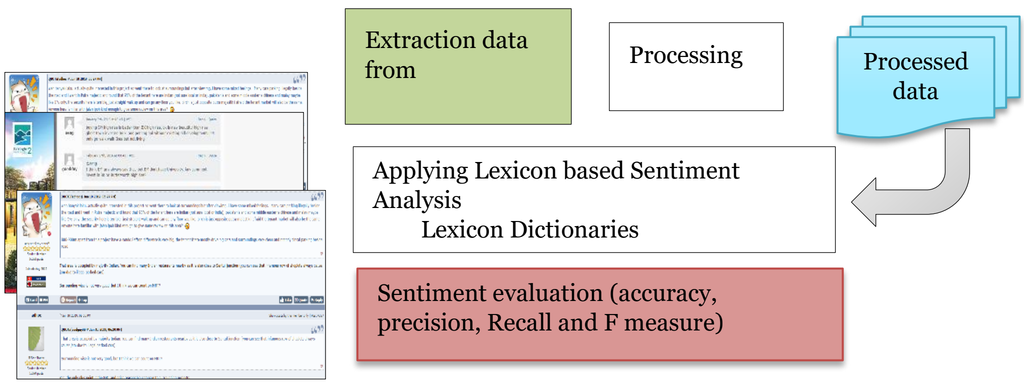

Figure 2: Research framework process

RESULTS

The three most used lexicon dictionaries, Wordnet, SentiWordNet, TextBlob and Valence Aware Dictionary for sentiment Reasoning (VADER). Table 2 presents the distribution of positive, negative, and neutral reviews for the overall projects. It is evident that all three lexicon dictionaries predicted a higher number of positive reviews compared to other polarity sentiments. The sentiment analysis results show that VADER classified 1,528 reviews as positive, 791 as negative, and 312 as neutral. SentiWordNet identified 1,787 reviews as positive, 735 as negative, and 109 as neutral. WordNet classified 1,800 reviews as positive, 759 as negative, and only 72 as neutral. Lastly, TextBlob categorized 1,567 reviews as positive, 710 as negative, and 354 as neutral. These numbers indicate that each dictionary has a different approach to classifying sentiments, with some identifying more positive reviews, while others detect more negative or neutral sentiments. To assess the accuracy of sentiment classification by the lexicon dictionaries, the entire dataset was annotated by human annotators in this study. The annotators were proficient in English and possessed domain knowledge in the property field.

Detailed guidelines and instructions were provided to ensure consistent and accurate annotation. The annotated dataset served as a baseline for validating the classification results of each lexicon dictionary. The annotators assigned each review in the dataset with either positive or negative polarity, as these were the two polarity types considered in this study. The classification results obtained from the lexicon dictionaries were compared with the manually annotated testing data to evaluate their accuracy. Evaluation measures, including accuracy, precision, recall, and F-measure, were calculated and used for comparison. These measures provide a comprehensive assessment of the performance of each lexicon dictionary in sentiment classification. The classification results for each lexicon dictionary are presented in subsequent sections. The utilization of human annotators for manual annotation ensured high accuracy in sentiment classification during the previous process.

This approach enhances the reliability of the baseline dataset used for validating the classification results of the lexicon dictionaries. The expertise and proficiency of the annotators, along with the provided guidelines, contributed to the consistency and quality of the annotations. In this study, the focus was on positive and negative polarities, as they are commonly used in sentiment analysis. Future research could explore additional polarity types or more refined sentiment categories to capture a more nuanced understanding of public sentiment towards high-rise properties. Overall, the combination of manual annotation by human experts and the evaluation measures employed in this study ensures a robust assessment of the lexicon dictionaries' sentiment classification accuracy. These findings contribute to the advancement of sentiment analysis methods and their applicability in understanding public sentiment towards high-rise properties in Malaysia.

Table 2: Sentiment polarity classification results for the reviews

Lexicon dictionary	Positive	Negative	Neutral
VADER	1528	791	312
SentiWordNet	1787	735	109
Wordnet	1800	759	72
TextBlob	1567	710	354

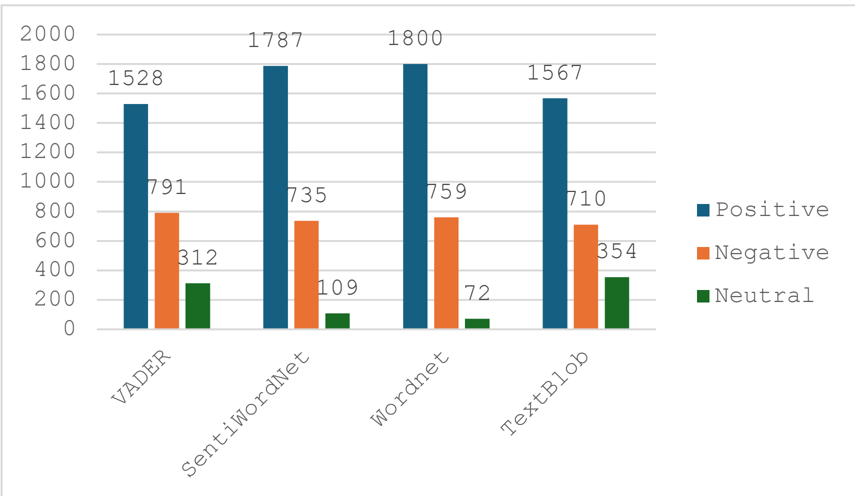

Figure 3: Classification results for the sentiment polarity of the reviews.

Table 3: Annotation's results

Approaches	Positive	Negative
Annotator 1	1362	1269
Annotator 2	1369	1262

This research employs four evaluation metrics: precision, recall, F-score, and accuracy. Additionally, the values for true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) have been computed and utilized to assess the system's. The outcomes from the manual annotation process are displayed in Table 4.

Table 4: The performance measure for each lexicon dictionary

Dictionary	Evaluation matrix	First annotator score	Second annotator score	Average score
VADER	Accuracy	0.778	0.785	0.782
	Precision	0.737	0.749	0.743
	Recall	0.909	0.908	0.909
	F1 score	0.814	0.821	0.817
Wordnet	Accuracy	0.689	0.690	0.690
	Precision	0.648	0.651	0.65
	Recall	0.878	0.877	0.877
	F1 score	0.745	0.747	0.746
SentiWordNet	Accuracy	0.694	0.697	0.695
	Precision	0.653	0.657	0.654
	Recall	0.88	0.886	0.886
	F1 score	0.751	0.754	0.753
TextBlob	Accuracy	0.714	0.734	0.724
	Precision	0.708	0.712	0.71
	Recall	0.890	0.901	0.8955
	F1 score	0.788	0.795	0.7915

As shown in Table 4, the classification results generated by VADER achieved an overall accuracy of 78% for the review sentences, along with a notably high recall rate of 90%. These findings indicate that the VADER lexicon dictionary outperformed the other two dictionaries in terms of sentiment classification. TextBlob as second, SentiWordNet achieved the third-highest accuracy, precision, recall, and F1-score, with only a slight difference in scores with Wordnet. Although the results demonstrate a reasonable level of accuracy for VADER in this study, it is important to examine the instances where misclassifications occurred. Table 6 provides an overview of the reviews that VADER incorrectly classified, offering insight into the specific cases of misclassification by the VADER lexicon.

Table 5: Example of misclassification by lexicon dictionary

Reviews	Incorrect Polarity	Manually Label Polarity
"good analysis sound pollution factor slowly die because pollution"	Positive	Negative
"developer add sliding sun louvre balcony recent project petalz use project npe highway increase privacy reduce serious sound pollution highway"	Negative	Positive
"Parking lot design kind weird design"	Neutral	Negative

Previous studies have highlighted several factors that contribute to misclassification by lexicon-based sentiment dictionaries. These factors include limitations in the lexicon's vocabulary coverage, the presence of sarcasm in reviews, ambiguity in word meanings, and the use of misspelled or abbreviated terms. Such challenges have been identified as key sources of inaccuracies in sentiment classification using lexicon dictionaries.

DISCUSSION

This study aimed to assess the applicability and effectiveness of different lexicon dictionaries in measuring public sentiment towards high-rise properties in Malaysia. The selected lexicon dictionaries, Wordnet, SentiWordNet, and

VADER, were compared based on their classification performance and accuracy. The analysis of overall project reviews revealed that all three lexicon dictionaries predicted a higher number of positive reviews compared to negative and neutral reviews. Wordnet exhibited the highest number of classified positive reviews, followed closely by SentiWordNet and VADER. In terms of negative reviews, VADER classified the most sentences, with TextBlob,Wordnet and SentiWordNet following behind. For neutral reviews, VADER had the highest number of positive polarity reviews, while SentiWordNet and Wordnet had fewer. To evaluate the classification accuracy of each lexicon dictionary, the results were compared with manually annotated testing data. The evaluation measures of accuracy, precision, recall, and F-measure were employed. VADER achieved the highest accuracy of 78% for overall review sentences, with a substantial recall value of 90%.

SentiWordNet also demonstrated respectable performance, albeit with slightly lower accuracy, precision, recall, and F1 score. VADER, incorporates two lexicons specifically designed for social media, which account for the intensity and nuances of colloquial expressions and internet slang. This allows VADER to analyze text more accurately compared to other, which explains the observed results. However, it is important to consider the cases where VADER misclassified reviews. The limitations of lexicon dictionaries, such as word limitations, sarcasm, ambiguity, and misspelled or short terms, were identified as contributing factors to misclassification. These factors highlight the need for further refinement and expansion of lexicons to enhance accuracy in sentiment classification. The findings of this study contribute to the understanding of sentiment analysis in the context of high-rise properties in Malaysia. The predominance of positive sentiment towards these properties aligns with the assumption that they are often associated with desirable features and amenities. Nevertheless, the misclassification cases identified by VADER underscore the challenges in accurately capturing sentiment using lexicon dictionaries alone. Future research should address the limitations of lexicon dictionaries by expanding and updating them to include a wider range of domain-specific words and expressions. Additionally, the incorporation of machine learning and natural language processing techniques can improve sentiment analysis models' performance in capturing contextual nuances and sarcasm. In conclusion, this study emphasizes the importance of carefully selecting and evaluating lexicon dictionaries for sentiment analysis in specific domains. The results provide valuable insights for researchers, practitioners, and policymakers involved in analyzing public sentiment towards high-rise properties in Malaysia, enabling them to make informed decisions based on accurate sentiment classification.

ACKNOWLEDGMENT

This research was supported in part by the Ministry of Higher Education (MOHE) of Malaysia through the Fundamental Research Grant Scheme (FRGS/RDU230148/1/2023/ICT06/UMP/02/3). We also extend our sincere thanks to UMPSA for the financial and other assistance that helped us reach the finish line.

REFERENCES

[1] I. Bessedik, N. Taghezout, and A. Saidi, “Web 2.0 integration in a Web-Based Decision Support System: Effect study of social networking on decision-making,” Second Int. Conf. Innov. Comput. Technol. (INTECH 2012), pp. 365–370, Sep. 2012.

[2] C. -l. Hsu, “Sociology of Hyperlink Networks of Web 1.0, Web 2.0, and Twitter: A Case Study of South Korea,” Soc. Sci. Comput. Rev., vol. 29, no. 3, pp. 354–368, Sep. 2010.

[3] A. Isfandyari-Moghaddam and M. Hosseini-Shoar, “Factors affecting Web 2.0 adoption: a case study,” Progr. Electron. Libr. Inf. Syst., vol. 48, no. 1, pp. 2–15, 2014.

[4] S. A. A. A. Alrababah, K. H. Gan, and T.-P. Tan, “Product Aspect Ranking using Sentiment Analysis: A Survey,” in International Research Journal of Engineering and Technology, 2016, pp. 124–128.

[5] N. Akma, A. Bakar, and M. Allegra, “The Simulation Models for Human Pedestrian Movement of a Departure Process in an Airport Institute for Educational Technology,” in 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2018, pp. 1–6..

[6] J. R. Ruscheinsky, M. Lang, and W. Schäfers, “Real estate media sentiment through textual analysis,” J. Prop. Invest. Financ., vol. 36, no. 5, pp. 410–428, 2018, doi: 10.1108/JPIF-07-2017-0050.

[7] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis,” Found. Trends Inf. Retr., vol. 2, no. 1–2, pp. 1–135, 2008.

[8] B. Liu, “Sentiment analysis and opinion mining,” Synth. Lect. Hum. Lang. Technol., vol. 5, no. 1, pp. 1–167, 2012.

[9] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, “New Avenues in Opinion Mining and Sentiment Analysis,” IEEE Intell. Syst., vol. 28, no. 2, pp. 15–21, Mar. 2013, doi: 10.1109/MIS.2013.30.
[10] N. Godbole and S. Skiena, “Large-Scale Sentiment Analysis for News and Blogs,” 2007.
[11] C. J. Hutto and E. Gilbert, “VADER : A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text,” in Proceeding of the Eighth International AAAI Conference on Weblogs and Social Media, 2014, pp. 216–225.
[12] F. H. Khan, U. Qamar, and S. Bashir, “SentiMI : Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection,” Appl. Soft Comput. J., vol. 39, pp. 140–153, 2016, doi: 10.1016/j.asoc.2015.11.016.
[13] H. Nguyen, R. Al, and K. Academy, “Comparative Study of Sentiment Analysis with Product Reviews Using Machine Learning and Lexicon-Based Approaches,” SMU Data Sci. Rev., vol. 1, no. 4, 2018.
[14] H. T. T. F. Tadashi Kumano Hideki Kashioka, “Construction and analysis of Japanese-English broadcast news corpus with named entity tags,” ACL2003 Work. Multiling. Mix. Named Entity Recognit. July 12 2003 Sapporo Japan 8pp PDF 56KB, 2003.
[15] A. Danielewicz-Betz, H. Kaneda, M. Mozgovoy, and M. Purgina, “Creating English and Japanese Twitter Corpora for Emotion Analysis,” People, vol. 1634, p. 5869, 2015, doi: 10.7763/IJKE.2015.V1.20.
[16] H. Y. Lee and H. Renganathan, “Chinese Sentiment Analysis Using Maximum Entropy,” pp. 89–93, 2011.
[17] A. Alsaffar and N. Omar, “Study on feature selection and machine learning algorithms for Malay sentiment classification,” Proc. 6th Int. Conf. Inf. Technol. Multimed., no. June 2016, pp. 270–275, 2014, doi: 10.1109/ICIMU.2014.7066643.
[18] T. Solorio and Y. Liu, “Part-of-Speech Tagging for English-Spanish Code-Switched Text,” in Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2008, pp. 1051–1060.
[19] S. Vicente, R. Agerri, and G. Rigau, “Simple , Robust and (almost) Unsupervised Generation of Polarity Lexicons for Multiple Languages,” in EACL, 2014, pp. 88–97.
[20] U. Barman, A. Das, J. Wagner, and J. Foster, “Code Mixing: A Challenge for Language Identification in the Language of Social Media,” in EMNLP 2014, 2014, p. 13.
[21] Z. Wang, S. Y. M. Lee, S. Li, and G. Zhou, “Emotion Detection in Code-switching Texts via Bilingual and Sentimental Information,” in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), 2015, pp. 763–768.
[22] M. R. González-Rodríguez, R. Martínez-Torres, and S. Toral, “Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness,” Int. J. Contemp. Hosp. Manag., vol. 28, no. 11, pp. 2609–2627, 2016, doi: 10.1108/IJCHM-02-2015-0057.
[23] K. Schouten and F. Frasincar, “Survey on Aspect-Level Sentiment Analysis,” IEEE Trans. Knowl. Data Eng., vol. 28, no. 3, pp. 813–830, 2016, doi: 10.1109/TKDE.2015.2485209.
[24] R. Varghese and M. Jayasree, “a Survey on Sentiment Analysis and Opinion Mining,” Ijret.Org, vol. 7, no. 1, pp. 312–317, 2013, doi: 10.4304/jetwi.5.4.367-371.
[25] M. S. Mubarok, A. Adiwijaya, and M. D. Aldhi, “Aspect-based sentiment analysis to review products using Naïve Bayes,” AIP Conf. Proc., vol. 1867, no. August, 2017, doi: 10.1063/1.4994463.
[26] T. S. Al-nahdi, E. Nyakwende, A. m. Banamah, and A. A. Jappie, “Factors Affecting Purchasing Behavior in Real Estate in Saudi Arabia,” vol. 6, no. 2, pp. 113–125, 2015.
[27] E. Siahaan, K. A. Fachrudin, and M. L. L. S., “Evaluation of Location Aspect in Buyer ’ s Decision to Purchase Elite Residence in Medan,” Adv. Econ. Bus. Manag. Res., vol. 46, no. Ebic 2017, pp. 623–629, 2018.
[28] E. Brock, “Study shows lack of ethics programs,” Am. City Cty., vol. 123, no. 3, pp. 8–10, 2008.
[29] Y. Fu et al., “Sparse Real Estate Ranking with Online User Reviews and Offline Moving Behaviors,” Proc. - IEEE Int. Conf. Data Mining, ICDM, vol. 2015-Janua, no. January, pp. 120–129, 2015, doi: 10.1109/ICDM.2014.18.
[30] F. Tan, C. Cheng, and Z. Wei, “Modeling real estate for school district identification,” in Proceedings - IEEE International Conference on Data Mining, ICDM, 2017, no. 2, pp. 1227–1232. doi: 10.1109/ICDM.2016.43.
[31] Y. Fan and W. X. Wan, “Public transport , noise complaints , and housing : Evidence from sentiment analysis in Singapore,” J. Reg. Sci., vol. 61, no. November 2020, pp. 570–596, 2021, doi: 10.1111/jors.12524.
[32] R. Feldman, “Techniques and applications for sentiment analysis,” Commun. ACM, vol. 56, no. 4, p. 82, 2013, doi: 10.1145/2436256.2436274.

[33] Q. Zhang, H. Chen, and X. Huang, “Chinese-English Mixed Text Normalization,” in Proceedings of the 7th ACM international conference on Web search and data mining, 2014, pp. 433–442. doi: 10.1145/2556195.2556228.

[34] K. Kapadia and S. Agarwal, “Integration of GIS and Fuzzy MCDM approach for Real Estate Investment analysis,” J. Mech. Civ. Eng., vol. 13, no. 3, pp. 56–66, 2016, doi: 10.9790/1684-1303035666.

[35] K. Denecke, “Using SentiWordNet for multilingual sentiment analysis,” in Proceedings - International Conference on Data Engineering, 2008, pp. 507–512. doi: 10.1109/ICDEW.2008.4498370.

[36] C. Fellbaum, WordNet, no. 2000. New-York: Springler, 2010. doi: 10.1007/978-90-481-8847-5.

[37] S. Baccianella, A. Esuli, and F. Sebastiani, “SentiWordNet 3 . 0 : An Enhanced Lexical Resource for Sentiment Analysis and S ENTI W ORD N ET 3 . 0 : An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining,” in LREC 2010 Proceedings (including the Workshops Proceedings), 2014, no. November.

[38] V. Bonta, N. Kumaresh, and N. Janardhan, “A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis,” no. March, 2019.

[39] N. Medagoda, S. Shanmuganathan, and J. Whalley, “Sentiment Lexicon Construction Using SentiWordNet 3.0,” in 11th International Conference on Natural Computation (ICNC), 2015, pp. 802–807.

[40] J. Park, “Framework for sentiment-driven evaluation of customer satisfaction with cosmetics brands,” IEEE Access, vol. 8, pp. 98526–98538, 2020, doi: 10.1109/ACCESS.2020.2997522.

[41] A. Khattak et al., “Fine-grained sentiment analysis for measuring customer satisfaction using an extended set of fuzzy linguistic hedges,” Int. J. Comput. Intell. Syst., vol. 13, no. 1, pp. 744–756, 2020, doi: 10.2991/ijcis.d.200513.001.

[42] D. Kang and Y. Park, “Measuring customer satisfaction of service based on an analysis of the user generated contents: Sentiment analysis and aggregating function based MCDM approach,” 2012 IEEE Int. Conf. Manag. Innov. Technol., pp. 244–249, 2012, doi: 10.1109/ICMIT.2012.6225812.