



**Sentiment analysis of reviews written in the Saudi
dialect to classify customer satisfaction levels in
e-commerce.**

Prepared By

Areej Haddaj Alqarni

252001729

**Submitted in fulfillment of the requirements for the degree Master
of Science in Artificial Intelligence**

At

College of Informatics

Midocean University

UAE

Supervisor: Dr. Yasser Abdul Hamid

College of Informatics, Midocean University, Comoros

[2026- Third Trimester]

Abstract

Sentiment analysis has become a fundamental tool for understanding customer satisfaction in e-commerce, especially with the increasing reliance on digital platforms for purchasing decisions. This study aims to develop a sentiment analysis model for customer reviews written in Arabic, with a particular focus on the Saudi dialect, in order to enable companies to evaluate the quality of products and services more accurately.

The research begins with a literature review on Arabic sentiment analysis, covering six scholarly papers that contributed to strengthening the theoretical foundation of the study. Subsequently, a dataset of customer reviews in the Saudi dialect was collected and prepared, applying text preprocessing techniques such as data cleaning, normalization, and feature extraction. A sentiment analysis model was then developed using machine learning algorithms, including various classifiers and Aspect-Based Sentiment Analysis (ABSA), with the goal of categorizing customer satisfaction into positive, negative, or neutral.

This study provides a specialized model for analyzing Arabic text in the Saudi dialect and establishes a clear methodology for processing informal texts. Additionally, it offers an annotated dataset that can be used in future research. The findings also enhance the understanding of customer satisfaction patterns in the Saudi market and support data-driven decision-making for e-commerce companies.

Table of Contents

Abstract	2
List of Abbreviations	4
Introduction	5
1.1 Problem Definition	5
1.2 Thesis Objectives	5
1.3 Thesis Contribution	6
1.4 Thesis Organization	6
Background	7
2.1 The Role of Sentiment Analysis in E-Commerce	7
2.2 Sentiment analysis in the Arabic language	7
2.3 Machine Learning Approaches	7
2.4 Pre-trained Language Models	8
2.5 Datasets in Arabic Sentiment Analysis	8
Related work	10
Literature Review Matrix	12
Research Gaps	13
Conclusion	13
References	14

List of Abbreviations

- NLP: Natural Language Processing
- SA: Sentiment Analysis
- SVM: Support Vector Machine
- NB: Naive Bayes
- KNN: K-Nearest Neighbors
- DT: Decision Tree
- RF: Random Forest
- ABSA: Aspect-Based Sentiment Analysis
- TF-IDF: Term Frequency – Inverse Document Frequency
- BOW: bag of words
- POS: Part-of-Speech Tagging

Chapter 1

Introduction

1.1 Problem Definition

E-commerce has witnessed rapid growth in recent years, and dependence on digital platforms has become an essential component of customers' decision-making processes. Customer reviews and textual evaluations now represent a significant source for understanding user experiences with products and services. With the increasing amount of content generated by customers on e-commerce platforms, companies are facing a major challenge in analyzing these data effectively, especially since relying on traditional methods or manual analysis has become impractical given the massive and continuously growing volume of data.

Accordingly, there has emerged a growing need to develop more advanced approaches based on artificial intelligence techniques to understand customers' opinions and accurately analyze their sentiments, enabling companies to improve their products, enhance the quality of services provided, and support future decision-making. Although several Arabic studies have addressed the field of sentiment analysis, there remains a need for further development in terms of the models used, the accuracy achieved, and the diversity of data sources. Moreover, there is still a clear need to integrate more advanced natural language processing techniques capable of handling dialectal language and analyzing textual content more precisely and effectively.

1.2 Thesis Objectives

The objectives of this thesis are as follows:

1. To collect and prepare a dataset of customer reviews written in the Saudi dialect from e-commerce platforms.
2. To apply Natural Language Processing techniques to preprocess, clean, and analyze informal text written in Saudi dialect.
3. To develop a sentiment analysis model, including Aspect-Based Sentiment Analysis (ABSA), capable of classifying customer satisfaction levels (positive, negative, and neutral).

4. To evaluate and compare the performance of various machine learning and deep learning algorithms in sentiment classification.
5. To identify the most influential linguistic features that contribute to determining sentiment in Saudi dialect customer reviews.

1.3 Thesis Contribution

This thesis provides several scientific contributions, the most prominent of which are:

1. Developing a sentiment analysis model specifically designed for customer reviews written in the Saudi dialect, within a research area that remains relatively limited.
2. Establishing a clear methodology for processing informal Arabic text in the context of e-commerce, including preprocessing, analysis, and feature extraction stages.
3. Providing an annotated dataset that can be utilized in future research related to sentiment analysis in the Arabic language.
4. Enhancing the understanding of customer satisfaction patterns in the Saudi market and supporting data-driven decision-making for e-commerce platforms.

1.4 Thesis Organization

This thesis is organized into several chapters as follows:

1. Chapter 1: Introduction – covers the research topic, objectives, research problem, and the significance of the study.
2. Chapter 2: Background – provides the theoretical foundation of the research.
3. Chapter 3: Literature Review – presents previous studies and compares them.
4. Chapter 4: Methodology – describes the research design, data collection methods, preprocessing steps, and the machine learning techniques used.
5. Chapter 5: Results – presents the experimental results and evaluates the performance of the models.
6. Chapter 6: Conclusion and Future Work – summarizes the main findings, contributions, limitations, and provides recommendations for future research.

Chapter 2

2 Background

2.1 The Role of Sentiment Analysis in E-Commerce

In the last few years, e-commerce has seen unprecedented growth, and it has become an inseparable part of consumers' buying behavior. Along with this growth, organizations have become increasingly aware of the need to understand customers in a digital environment through opinion analysis and reviews coming from every online platform and social networking site. The modern markets cannot solely depend on conventional feedback mechanisms but need intelligent tools that can interpret customer digital impressions for service quality improvements, data-driven decision-making, and increasing customer satisfaction.

Therefore, it is important to have digital tools that help companies understand their customers' opinions about products, especially given the vast number of reviews that cannot be analyzed manually. Sentiment analysis provides insights into customer opinions by examining their reviews and determining whether they are positive, negative, or neutral through the processing of textual content.

2.2 Sentiment analysis in the Arabic language

Sentiment analysis in the Arabic language has become a crucial area of research, especially for understanding customer opinions on social media and online reviews. Researchers face significant challenges, including dialectal variation, morphological complexity, and ambiguity in informal or abbreviated texts. To address these issues, previous studies have emphasized the importance of building corpora and lexicons specific to different dialects, such as Jordanian, Egyptian, and Saudi Arabic, to develop more accurate and effective models for sentiment analysis in Arabic.

In the field of text analysis and sentiment classification, a set of traditional and modern algorithms are used, each with its own characteristics and way of functioning. Among the most widely used algorithms in previous studies are the following:

2.3 Machine Learning Approaches

Many previous studies have relied on machine learning techniques for Arabic sentiment analysis, where these studies mainly focused on specific stages beginning with text preprocessing to improve data quality prior to the classification process. This stage includes noise removal, text normalization, and handling the various linguistic characteristics of the

Arabic language. Subsequently, feature extraction techniques were applied to represent textual data numerically, enabling Machine learning algorithms to process it effectively. Alrefae et al. (2024), the researchers employed six different feature extraction combinations to identify the most effective approach for improving model accuracy. Several machine learning classifiers were evaluated, including SVM, NB, RF, and KNN. The results demonstrated that the SVM model combined with TF-IDF, POS tagging, and bigram features achieved the highest performance among the evaluated models.

Furthermore, the study conducted by Dahbi et al. (2019) showed that converting texts from local dialects into Modern Standard Arabic (MSA) led to a significant improvement in the performance of machine learning algorithms. In particular, the SVM classifier achieved the highest performance after this conversion, highlighting the importance of linguistic normalization in Arabic sentiment analysis. Louati et al. (2023) demonstrated that the SVM algorithm achieved strong performance that was very close to that of the deep learning model CAMELBERT, which confirms the efficiency and effectiveness of SVM in Arabic sentiment analysis tasks.

2.4 Pre-trained Language Models

They are artificial intelligence models initially trained on large volumes of texts to learn about linguistic structures, context, and relationships of words. Later, they can be fine-tuned for specific tasks such as text classification, sentiment analysis, question answering, or information extraction, with no need to train them from scratch, which saves time and resources and increases accuracy.

There are clear gaps in previous studies, as the datasets used to train the models were small and were often collected manually from multiple platforms. This lack of sufficient data has directly affected the models' performance, and the limited dataset may have contributed to the inaccuracy of the models' evaluation results.

2.5 Datasets in Arabic Sentiment Analysis

In previous studies on sentiment analysis, datasets are often collected from social media platforms or online review websites, such as Twitter or e-commerce platforms. These datasets typically consist of user-generated texts in Modern Standard Arabic or various Arabic dialects. The texts are usually manually labeled into categories such as positive, negative, or neutral, and preprocessing techniques such as noise removal, normalization, and stemming are applied to improve data quality and enhance algorithm performance. However, building high-quality Arabic sentiment datasets remains challenging due to the diversity of dialects, the prevalence of informal language, and the limited availability of publicly annotated resources. Moreover, in the study by Alrefae, et al., (2024), researchers

applied Aspect-Based Sentiment Analysis (ABSA) instead of traditional full-text sentiment analysis. This approach requires a carefully annotated dataset at the aspect level, covering attributes such as price, quality, service, and others, in order to accurately identify sentiments related to each aspect. The results demonstrated that ABSA provides a more detailed and in-depth analysis compared to general sentiment analysis, offering actionable and precise insights into customer opinions and enhancing the understanding of user experience.

Chapter 3

Related work

In study [11], the researchers constructed a dataset of Arabic tweets written in the Jordanian dialect. The data were collected using keywords related to general Jordanian topics. The study employed several preprocessing strategies, including data cleaning, removal of URLs, unnecessary characters, repeated words, and hashtags. The tweets were then manually annotated and classified into positive, negative, and neutral sentiments. The dataset was trained using traditional machine learning models such as Support Vector Machine (SVM) and Naïve Bayes. The results demonstrated that using TF-IDF representation combined with stemming and bigram features within the SVM model achieved the best performance in terms of accuracy, F1-score, and recall.

Similarly, in [9] focused on analyzing citizens' sentiments regarding traffic conditions in three Moroccan cities. Opinions were collected from social media posts on both Twitter and Facebook written in the Moroccan dialect. The study applied traditional machine learning algorithms, and the results showed that the Support Vector Machine (SVM) model outperformed other algorithms after translating the texts from the Moroccan dialect to Modern Standard Arabic (MSA), achieving an accuracy of 92%. The findings also indicated a predominance of negative sentiments, particularly in the city of Rabat.

In 2024, study [1] applied Aspect-Based Sentiment Analysis (ABSA) to analyze customer reviews written in the Saudi dialect using Natural Language Processing techniques and machine learning algorithms. The researchers collected two separate datasets, one for clothing products and the other for electronics, which were extracted from e-commerce platforms such as Amazon and Shein. The data were annotated by identifying relevant aspects, such as quality and price in the electronics category, as well as size and color in the clothing category, and assigning sentiment values of 1 for positive, -1 for negative, and 0 for neutral. The evaluation results showed that the Support Vector Machine (SVM) algorithm outperformed other models, achieving an accuracy of 86.5%, an F1-score of 85.82%, and a recall of 86.50% when using a combination of TF-IDF, POS, and bigram features.

In study[13] , the researchers analyzed the sentiment of comments on the YouTube platform to enable content creators to assess viewers' satisfaction with their content. This study addressed the challenge resulting from the removal of the dislike count feature on YouTube by relying on user comments as an alternative indicator of audience perception. The data were collected using a Python-based script connected to the YouTube Application Programming Interface (API), through which 4,760 comments were gathered from various video categories, The sentiments were manually labeled by five annotators into two classes: positive and negative. After removing irrelevant comments, a balanced dataset consisting of 4,212 comments was obtained, equally divided into 2,106 positive and 2,106 negative comments. Following text preprocessing, several machine learning classifiers were evaluated. The results demonstrated that the Naïve Bayes classifier outperformed the other classifiers, achieving an accuracy of 94.62%, a precision of 94.64%, a recall of 94.64%, and an F1-score of 94.62%.

In study [12], the researchers aimed to develop a sentiment analysis system for opinions written in Arabic by higher education students, with the objective of enhancing and improving the educational process. The study proposed the SVM-SAA-SCR algorithm, which relies on text preprocessing followed by classifying opinions into positive, negative, or neutral sentiments using the Support Vector Machine (SVM) algorithm. The researchers used a real-world dataset collected from Prince Sattam bin Abdulaziz University (PSAU), which includes students' opinions on various educational aspects. To evaluate the system's performance, it was compared with the advanced CAMELBERT model. The results revealed very minor differences between the two models, confirming the robustness and effectiveness of the proposed SVM-SAA-SCR algorithm.

Literature Review Matrix

This table provides a summary of the most relevant previous studies related to sentiment analysis in the context of e-commerce, highlighting the methodologies used and the key findings of each study.

Study	Algorithm	Dataset	Key Findings
Alrefae, et al., (2024)	SVM, Random Forest (RF), Naive Bayes (NB), KNN	Amazon, Jarir, Shein, Noon, and Twitter	-For electronics dataset, the RF algorithm combined with TF-IDF, POS tagging, and trigrams achieved the highest effectiveness, with an average accuracy of 86.26% per aspect. -For clothing dataset , the algorithm combined with TF-IDF, POS tagging, and bigrams achieved an average accuracy of 86.51% per aspect.
Al-Thubaity, et al.,(2018)	Lexicon-based approach	Collect data in twitter	SauDiSenti performed better for neutral tweets and showed promising results despite being smaller than AraSenTi.
Alomari, et al.,(2017)	SVM +Naive Bayes	collected from the X platform(Twitter).	The results showed that SVM outperformed Naive Bayes in terms of accuracy and F-score.
Dahbi,et al.,(2019)	Naive Bayes, SVM,KNN, Decision Tree	using a Python script from Twitter and Facebook	SVM algorithm achieved the highest performance after converting the texts from the Moroccan dialect to Modern Standard Arabic (MSA).
Louati, et al., (2023)	SVM	real student opinions from the Department of Computer Science and Information Systems at Prince Sattam University.	The SVM-SAA-SCR model achieved strong performance with an accuracy of 84.7% ,Precision of 86% and Recall 85%.
Musleh, et al.,(2023)	NB, SVM,KNN, DT,RF	YouTube comments	Achieved an accuracy of 94.62%, precision of 94.64%, recall of 94.64%, and F1-score of 94.62%, making it the best-performing classifier among the six evaluated classifiers.

Research Gaps

Previous studies on Arabic sentiment analysis face a significant lack of resources and datasets, as most research relied on limited-sized datasets that do not cover all dialects and domains. In addition, there is a lack of adequate utilization of linguistic lexicons and dictionaries, and they have not been integrated with modern techniques such as Transfer Learning, which limits the models' ability to understand nuanced and figurative expressions and accurately analyze sentiments, delaying the development of reliable and effective tools for Arabic text analysis.

Conclusion

The studies highlight the importance of sentiment analysis in business development and in leveraging information extracted from social media texts to enhance and improve business practices. They demonstrate that Arabic text analysis is feasible and effective, whether in Modern Standard Arabic or in various dialects, with the necessity of performing text preprocessing, feature extraction, and considering dialectal differences to improve accuracy. The results also showed the effectiveness of machine learning algorithms, particularly the SVM algorithm, which achieved the best performance and was considered the most efficient scenario among the evaluated classifiers.

References

- [1] Alrefae, R., Alqahmi, R., Alduraibi, M., Almatrafi, S., & Alayed, A. (2024). Enhancing Customer Experience Through Arabic Aspect-Based Sentiment Analysis of Saudi Reviews. *International Journal of Advanced Computer Science & Applications*, 15(7).
- [2] Hicham, N., Karim, S., & Habbat, N. (2023). Enhancing Arabic Sentiment Analysis in E-Commerce Reviews on Social Media Through a Stacked Ensemble Deep Learning Approach. *Mathematical Modelling of Engineering Problems*, 10(3).
- [3] Al-Thubaity, A., Alqahtani, Q., & Aljandal, A. (2018). Sentiment lexicon for sentiment analysis of Saudi dialect tweets. *Procedia computer science*, 142, 301-307.
- [4] Alotaibi, S., Mehmood, R., & Katib, I. (2019, June). Sentiment analysis of Arabic tweets in smart cities: A review of Saudi dialect. In *2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC)* (pp. 330-335). IEEE.
- [5] Abo, M. E. M., Qazi, A., Saad, A. A. A., Elsahib, H. A., & Abdelaziz, A. (2024, August). Sentiment Analysis in Saudi Arabic Dialect for HAJJ Season Services Using Twitter Data. In *2024 1st International Conference on Logistics (ICL)* (pp. 1-5). IEEE.
- [6] Alahmary, R. M., Al-Dossari, H. Z., & Emam, A. Z. (2019, January). Sentiment analysis of Saudi dialect using deep learning techniques. In *2019 International conference on electronics, information, and communication (ICEIC)* (pp. 1-6). IEEE.
- [7] Kaseb, A., & Farouk, M. (2022, December). SAIDS: A novel approach for sentiment analysis informed of dialect and sarcasm. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP)* (pp. 22-30).
- [8] AlFutamani, A. A., & Al-Baity, H. H. (2021). Emotional Analysis of Arabic Saudi Dialect Tweets Using a Supervised Learning Approach. *Intelligent Automation & Soft Computing*, 29(1).
- [9] Dahbi, M., Saadane, R., & Mbarki, S. (2019, October). Social media sentiment monitoring in smart cities: an application to Moroccan dialects. In *Proceedings of the 4th International Conference on Smart City Applications* (pp. 1-6).
- [10] AlFutamani, A. A., & Al-Baity, H. H. (2021). Emotional Analysis of Arabic Saudi Dialect Tweets Using a Supervised Learning Approach. *Intelligent Automation & Soft Computing*, 29(1).

[11] Alomari, K. M., ElSherif, H. M., & Shaalan, K. (2017, June). Arabic tweets sentimental analysis using machine learning. In *International conference on industrial, engineering and other applications of applied intelligent systems* (pp. 602-610). Cham: Springer International Publishing.

[12] Louati, A., Louati, H., Kariri, E., Alaskar, F., & Alotaibi, A. (2023). Sentiment analysis of Arabic course reviews of a Saudi university using support vector machine. *Applied Sciences*, 13(23), 12539.

[13] Musleh, D. A., Alkhwaja, I., Alkhwaja, A., Alghamdi, M., Abahussain, H., Alfawaz, F., ... & Abdulqader, M. M. (2023). Arabic sentiment analysis of youtube comments: Nlp-based machine learning approaches for content evaluation. *big data and cognitive computing*, 7(3), 127.