IBIMA Publishing
Communications of International Proceedings
https://ibimapublishing.com/p-articles/46AI/2025/4644225/
Vol. 2025 (34), Article ID 4644225

# Improving Digital Marketing Through Analyzing Arabic Customer Reviews by Natural Language Processing*

Mennat Allah HASSAN[1], Silvia GARCÍA-MÉNDEZ[2] and Francisco DE ARRIBA-PÉREZ[2]

[1]Arab Academy for Science, Technology & Maritime Transport, Cairo, Egypt

[2]Information Technologies Group, atlanTTic, University of Vigo,

Correspondence should be addressed to: Mennat Allah HASSAN, mennahani149@aast.edu

* Presented at the 46th IBIMA International Conference, 26-27 November 2025, Ronda, Spain

## Abstract

Digital marketing has become a practical part of modern business strategies, allowing companies to reach wider audiences and engage with customers innovatively. Understanding customer sentiment towards products, services, and experiences is essential for effective marketing campaigns. In this line, Sentiment Analysis (SA), a task of Natural Language Processing, offers valuable insights by automatically analyzing and categorizing opinions expressed in textual data. Consequently, this paper studies different aspect-based approaches to SA in digital marketing and provides a literature review on the most recent studies in Middle East, that focuses on reviews analysis. Note that Arabic Chat Alphabet (ACA) is the informal language used in online chats, social media platforms, and instant messaging applications in the Middle east. This paper highlights the substantial gap in the in the ABSA literature concerning aca across different dialects—particularly the Egyptian dialect, also referred to as Egyptizi., showing that the demand for developing new models capable of mining and evaluating the ACA continues to grow to help analyze the reviews written in ACA to assist personalization and digital marketing.

**Keywords:** NLP, Personalization, ABSA, Franco Arabic.

## Introduction

Data is the primary and most critical asset in today's rapidly evolving information economy. Online reviews written by Customers offer a huge quantity of data. These reviews can offer valuable insights to businesses and aid in their future decision-making processes (Hassan et al. 2016; Lee et al. 2022) as they reflect the customers' opinions, needs, and preferences. Personalization can benefit a lot from the analysis of these reviews. Today, the advancements offered by technology can help businesses meet the user's desire to own a product or service with a personal signature. This process is known as digital personalization (Chandra et al. 2022), an essential aspect of digital marketing. More in detail, digital personalized marketing is a method that utilizes online consumer data to modify the user experience to address customers by name and present shoppers with tailored recommendations, products, and services, aka targeted marketing (Nikolajeva and Teilans 2021).

**Cite this Article as:** Mennat Allah HASSAN, Silvia GARCÍA-MÉNDEZ and Francisco DE ARRIBA-PÉREZ, Vol. 2025 (34) "Improving Digital Marketing Through Analyzing Arabic Customer Reviews by Natural Language Processing " Communications of International Proceedings, Vol. 2025 (34), Article ID 4644225, https://doi.org/10.5171/2025.4644225

Artificial Intelligence (AI) solutions have been increasingly applied in both digital marketing and targeted marketing, leveraging Machine Learning (ML) models and Natural Language Processing (NLP) techniques to enhance consumer engagement and optimize marketing strategies. Among the latter, Sentiment Analysis (SA) and Aspect- Based Sentiment Analysis (ABSA) are highly popular for online review analysis (Saura 2021), consequently, this paper contributes to the debate about the importance of using ABSA in analyzing customers' reviews to achieve digital personalization, especially in the Middle East, where internet users are increasing daily, specifically in Egypt (Alhamad and Kurdy 2020).

This paper is organized as follows. Section 2 presents the digital marketing concept, its history, and the difference between traditional and digital marketing. Section 3 presents the personalization process, the difference between personalization and customization, and the role of customers' feedback in digital personalization. Section 4 presents the online reviews and the role of ml and NLP in this regard. At the same time, Section 5 focuses on SA, ABSA, paying particular attention to the Arabic language, Arab Chat Alphabet (ACA), which is also known as Egyptian Arabizi. Finally, Section 7 concludes the paper and proposes future work.

## Digital Marketing

The digital innovation has transformed the way businesses utilize technology, enabling more interactive, targeted, and measurable methods of engaging with consumers. This approach is widely recognized as digital marketing. This term was first used in the 1990s, during which time it was mainly used to refer to product advertising to customers (Molina Parra et al. 2016).

In 2008, the business intelligence idea was introduced to digital marketing, including a multi-channel marketing strategy (Borah 2008). Table 1 inspired by Borah, (2008) presents the difference between traditional marketing and digital marketing regarding the audience, marketing approach, documentation, cost, marketing analysis, availability, and communication. Between 2000 and 2010, with new social and mobile tools, the digital marketing concept expanded much more. It was transformed from an advertising-oriented technique to creating an experience that links business with customers (Molina Parra et al. 2016).

**Table 1: Difference Between Traditional Marketing and Digital Marketing**

.

| Category | Traditional Marketing | Digital Marketing |
|---|---|---|
| Audience | Local or limited consumers can be targeted. | It can reach target customers all over the globe. |
| Marketing approach | Physical peer-to-peer relationships are formed when delivering goods. | No physical relationship is required between marketers and consumers in digital form. |
| Documentation | Promotion of products is done with brochures, pamphlets, or hard copies of products and services. | It is paperless and done through online videos, websites, and social media pages. |
| Marketing cost | It involves a physical marketing mode, which costs the company much more. | It is much cheaper as it works online with the help of social media. |
| Marketing analysis | Depends on surveys and manual findings; result analysis is complex. | Facts and figures of digital data make analysis quick and efficient. |
| Availability | Not possible to retain 24/7 availability. | Fully available at any time. |
| Communication | One-way communication. | Two-way communication, highly beneficial for organizational growth. |

The importance of establishing digital relationships with consumers through digital marketing continues to grow rapidly. Digital and technological innovations, including intelligent products and ai have become fundamental

drivers of consumer transformation and engagement in the modern marketplace. They are changing corporations' marketing strategies to cover people's demands for a product or service (Molina Parra et al. 2016). Nowadays, businesses use several digital marketing techniques like search engines, social media optimization, and email marketing, among other strategies, to empower their presence on digital platforms and reach their target customers (Agrawal 2021).

Summing up, digital marketing as a concept has evolved from a narrowly defined approach focused on promoting products and services through online channels to a comprehensive framework that integrates digital technologies for customer acquisition, engagement, long-term relationship management, understanding user preferences, brand promotion, increasing customer loyalty and increasing sales (Kannan and Li 2017). Digital marketing includes various activities, techniques, and processes enabled by digital technologies to create, communicate, and deliver value for customers and various business stakeholders. Ultimately, digital marketing is identified by different terms such as internet marketing, online marketing, and web marketing (Borah 2008).

## Personalization

The Oxford definition of personalization is "Designing or producing something that meets someone's requirement". Hence, personalization refers to "offering the right product and service to the right customer at the right time and place" (Sunikka and Bragge 2012). Various terms like individualization, segmentation, one-to-one marketing, and customization are often used interchangeably with personalization to describe tailored offerings centered around the customer (Chandra et al. 2022). However, some nuances must be considered.

### *Difference between Personalization and Customization*

Two primary forms of one-to-one marketing are defined: personalization and customization. Personalization refers to firm-driven decision-making processes that utilize previously collected customer data to design and implement the most effective marketing strategies tailored to individual customers. Customization includes customer behaviors aimed at tailoring specific elements according to their preferences and requirements (Arora et al. 2008). Fig. 1, inspired by Arora et al., (2008), shows the difference between personalization and customization regarding initiation.

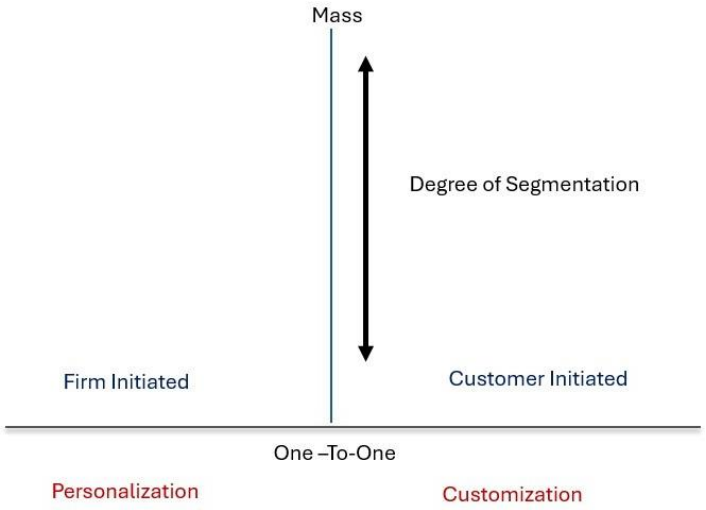


**Fig. 1 Difference between personalization and customization**

Customization is frequently in conflict with Personalization, and the two terms are often used alternately in both academic and practical contexts. However, they can be differentiated conceptually in terms of control ownership. Personalization involves the firm's initiated concept of adjusting the marketing mix to meet individual preferences based on customer data; in contrast, customization involves customers adjusting the marketing strategy to fulfill their needs and enhance satisfaction (Chandra et al. 2022).

## *Personalized Digital Marketing*

Personalized or one-to-one digital marketing is a strategy by which e-businesses apply data analysis to deliver individual marketing messages and develop customized products and services to existing and prospective customers (Behera et al. 2020). In theoretical terms, providing personalized content recommendations to individual customers is a key component of online relationship marketing, as it helps build stronger connections and long-term relationships with customers. Digital personalized marketing can also be defined as "utilizing online consumer data to modify the user experience to address customers by name, present shoppers with tailored recommendations, and more". Mainly, this is targeted marketing at its most raw (Nikolajeva and Teilans 2021). To achieve successful personalized marketing, a business must focus on its relationship management and the customer's data (Chandra et al. 2022).

## *Categories of personalized digital marketing*

Personalization in digital systems can be broadly classified into six primary categories: (i) link-based, (ii) content-based, (iii) search-based, (iv) context-aware, (v) authorization-based, and (vi) humanized personalization. This study primarily concentrates on content personalization, which refers to the dynamic adaptation of information related to a specific product or service to align with distinct user profiles. In practical applications such as e-commerce platforms this form of personalization manifests through differentiated offers or price variations for the same product, tailored according to users' purchasing behavior, preferences, and expressed opinions within written reviews (Mariossbelk 2016).

More in depth, users generate huge amounts of data at each stage of their customer journey when purchasing. This data can be gathered from digital channels, including website visits, emails, and interactions with a brand's mobile application (Desai 2019). Moreover, users may leave reviews after purchasing on e-commerce shopping websites that can be implanted in personalization. These reviews also help inform the purchasing decisions of other users on e-commerce platforms. Therefore, customer reviews play a vital role in personalized marketing, offering companies valuable insights into consumers' behaviors, interests, and preferences. Companies can better understand their customers' desires and requirements by collecting and analyzing their feedback. They can then utilize this information to develop more targeted and efficient marketing campaigns, improving sales and profitability.

## *The Role of Customers' Feedback in Personalized Digital Marketing*

Customer feedback is essential in personalized marketing as it gives companies precious insights into customers' behaviors, interests, and preferences, facilitating a deeper understanding of their desires and needs. Notably, a primary advantage of customer feedback is its ability to generate buyer profiles and fictional representations of dis- tinct customer segments using accurate data and insights. Companies can create more targeted messages and personal offers by understanding the motivations, preferences, and behaviors.

Examples of businesses that uses customer feedback into their personalization strategies include (i) Netflix[i], which uses these data to recommend movies and tv shows based on their viewing history and preferences, (ii ) Amazon[ii], which uses customer feedback and purchase history to personalize product recommendations and shopping experiences, (iii ) Starbucks[iii] which exploits customer feedback to improve its mobile ordering and rewards program, (iv ) Spotify[iv] which takes advantage of customer feedback and analyzes listening data to personalize music recommendations and playlists, and (v ) Airbnb [v] which uses customer feedback to personalize the booking and hosting experience.

## Online Reviews

"Online reviews refer to the empirical information of consumers who convey their opinions or feelings about a product or service they have experienced through text and images that can be studied regarding usage reviews, purchase reviews, and comments posted online" (Lee et al. 2022). Consumers seek the opinions of other consumers, particularly those with more experience, regarding products or services they intend to buy. Additionally, online reviews provide insight into the positive or negative perceptions surrounding their purchases (Lee et al. 2022). Over the past few decades, there has been a significant expansion in the variety of online

platforms that allow Inter- net users to share their opinions about products, services, or news. These platforms encompass social media sites, website review sections, discussion forums, blogs, and others ((Alyami et al. 2022).

In the Middle East, the number of internet users has significantly increased in recent years, accompanied by a growing interest in online shopping and digital purchasing behaviors. As a result, the total number of Internet users in Arab countries has reached 157 million people, as reported by the Arabic Network for Human Rights Information (Alhamad and Kurdy 2020). Online shoppers are widely spread across various countries in the Middle East, with 10.6 million in Saudi Arabia, 6.8 million in the UAE 2.4 million in Kuwait, and 15.2 million in Egypt, among other Arab countries, each at different rates.

The mobile phone is now the top-selling product in the Arab world, as indicated by the director of Amazon. This allows ubiquitous access to these new online platforms. Nowadays, there are several types of online shopping. One major type is using social networking sites such as Facebook[vi], X[vii], Threads[viii], Instagram[ix], and others to market and sell a particular company's product through its pages on these social platforms. The second leading type refers to the company's official website. In both cases, customers can write and express their views on products and services and share their experiences with others (Alzu'Bi et al. 2019).

Individuals and organizations in data mining technologies are firmly getting more interested in studying hidden information about products and services, evaluating and analyzing these online customers' reviews (Alyami et al. 2022). Consequently, one of the most rising related research areas in computer sciences is SA, which aims to extract and analyze the polarity of user opinions (Do et al. 2019).

## *The use of ML in analyzing online reviews*

ML is a topic of rising interest in the digital world, industry discussions, and the advertising world today (Bayoude et al. 2018). Numerous companies employ ml algorithms to dig latent insights within consumer data, aiming to enhance business processes, deliver enhanced customer experiences, and achieve operational efficiencies such as improved speed, cost-effectiveness and heightened precision through data analytics (Bayoude et al. 2018).

At the core of digital personalization, ml can be a solution for the automated analysis of extensive data sets. These models are trainable by both technical and non- technical researchers or marketers, facilitating the extraction of actionable insights and pattern recognition. A diverse range of algorithms can be applied and trained using online reviews (Saura 2021).

More in detail, ml has evolved into what is known as Deep Learning, a technology that has changed how computers analyze languages and images. However, others are used in the review analysis, such as ensemble models (EM), machine vision (MV), and NLP techniques. Table 2, inspired by Saura (2021), explains these approaches in detail. Notably, this work focuses on the NLP approach in analyzing online reviews, particularly on SA and ABSA. The information obtained from such analysis is critical and beneficial in digital personalization. It lets marketers know customers' opinions, preferences, and emotions towards the products and services offered (Saura 2021).

**Table 2 : Machine Learning Methods Used in Digital Marketing**

| Type | Description |
|---|---|
| EM | Performs predictions using a source model resulting from the majority voting of individual models. |
| Deep Learning Neural Networks | Consist of multiple network layers that recognize patterns and can be trained. These models learn from complex datasets and apply the acquired knowledge to other datasets by adopting learned criteria and rules. |
| MV | Allows for visual identification and recognition of objects, people, and products. |
| NLP | Analyses textual content such as online reviews to identify insights and extract meaning. |

### *The use of NLP in analyzing online reviews*

"NLP is a tract of AI and linguistics devoted to making computers understand the statements or words written in human languages. It was discovered to ease the user's work and to satisfy the wish to communicate with computers in natural language" (Khurana et al. 2023).

NLP can be categorized into two primary phases: Natural Language Understanding (NLU) and Natural Language Generation (NLG) (Alhamad and Kurdy 2020). NLG is defined as the segment of NLP where machines automatically produce text in natural human language (Ramos-Soto et al. 2016). At the same time, NLU empowers machines to comprehend human-written natural language and analyze it by extracting concepts, entities, emotions, keywords, and so forth. It finds utility in customer care applications to discern the issues expressed verbally or in written form Khurana et al. (2023). Understanding and analyzing human-written reviews is challenging, and understanding the sentiment adds another layer of complexity (Alhamad and Kurdy 2020).

## SA & ABSA

SA is one of the main tasks within NLP. It involves examining and analyzing users' sentiments and emotions in a sentence. In the context of online reviews, it analyzes texts to discover the polarity conveyed by reviewers' passages (Alhamad and Kurdy 2020).

SA has been generally categorized at three levels. Document level (Tripathy et al. 2017), sentence level (Liu, 2010.), and aspect level (Schouten and Frasincar 2016), respectively, to classify a whole document, a sentence, and the polarity (positive, negative, or neutral) of a particular aspect. Thus, SA focuses on the sentiment of a text rather than its structure (Nazir et al. 2022). Note that the reviews' polarity is of interest for decision-makers to understand and analyze customer satisfaction with their products/services (Alhamad and Kurdy 2020). In contrast, ABSA, the core of our work, focuses on extracting the aspects of the text's significant entities and identifying the sentiment the text expressed for each of them (Pontiki et al. n.d.). ABSA includes three main processing stages, i.e., aspect extraction (AE), sentiment evolution (SE), and aspect sentiment analysis (ASA). Phase one starts with extracting aspects, and the second phase classifies sentiment polarity for a predefined aspect, target, or entity. The third phase involves discovering people's sentiments towards the predefined aspects (Nazir et al. 2022). These phases can be done using varied techniques and models; some will be discussed in the following sections.

### *SA and ABSA in different languages*

SA can vary across different languages. Most SA research has been conducted using data sets in (Wang et al. 2021). To our knowledge, the first work presented for ABSA was in 2004 (Hu and Liu 2004), also for the English language. This paper was among the first to extract product features mentioned by customers, identify opinion-bearing sentences within each review, and determine the corresponding sentiment polarity. After that, many research papers have been published, experimenting with ABSA in languages other than English, such as Indian (Yadav et al. 2021) and Persian (Jafarian et al. 2021). However, the number of ABSA performed using these languages is very low compared to the ABSA research performed on data sets in English Language.

### *SA and ABSA in Arabic language, ACA and Egyptian Arabizi*

Arabic belongs to the Semitic language family, with more than 300 million people speaking it in the Middle Eastern and North African countries (Alyami et al. 2022). Moreover, Arabic is one of the six official languages used by the United Nations (Oueslati et al. 2020). According to the Internet World Stats (Bensoltane and Zaki 2023), in the past two decades, the number of internet users utilizing Arabic has surged by 9348.0 %, making it the fourth most utilized language. The Arabic language is divided into three varieties:

- Classical Arabic: It represents the earliest known form of the Arabic language., served as the language of the Quran, the Holy Book of Islam. During the classical age of the Islamic empire, it was employed for various purposes, including poetry, grammar, medicine, and other sciences (Bensoltane and Zaki 2023). However, due to its complexity, its use is focused today on religious and highly formal contexts.

- Modern Standard Arabic (al-Fusha): is the modern descendant and more accessible version of classical Arabic. It is generally used for formal communication, such as newspapers, school materials, emails, news broadcasts, and government publications (Bensoltane & Zaki, 2023).

- Dialectical Arabic (Aammiyya): is used in informal contexts and daily communications between people. The main groups of Arabic dialects are categorized according to the country, which is Levantine (e.g., Syria), Gulf (e.g., Saudi), Maghrebi (e.g., Morocco and Libya), and Egyptian (e.g., Egypt and Sudan) (Alhamad & Kurdy, 2020;Alyami et al., 2022)

However, in recent years, a new form of Arabic writing has emerged. It is the Arab Chat language (ACL) with the ACA, which consists of Arabic words written using English letters and numbers (Bensoltane and Zaki 2023). In addition, ACA (aka Arabizi, Arabish, Franco-Arab, or Franco) is a recent writing system for Arabic in which English letters are being used in writing instead of Arabic ones. In other words, it is an encoding technique representing every Arabic phenomenon with English letters matching the same pronunciation (Mostafa 2012). For example, the Arabic word "good morning" is "saba7 el 5eir" in Franco-Arab.

Arabizi, is a blend of Arabic and English and is commonly used by Arabic native speakers particularly in texting and typing comments in social media platforms since it allows users to write Arabic words using English letters. Research indicates that Arabizi accounts for approximately 12% of Latin script tweets in Lebanon and 25 % of English/Latin script tweets in Egypt (Tobaili et al. 2019). Consequently, ACA is a natural language that includes short vowels missing in traditional Arabic orthography" (Duwairi et al. 2016). A comparative study between Modern Standard Arabic (MSA) and ACA showed that 86 % of Arabic computer users confirmed that they type faster while using aca (Mostafa 2012). Every Arabic character is mapped into its corresponding Latin letter, as shown in Table 3 inspired by (Duwairi et al. 2016).

Users map the ACA with Latin alpha numerals by their dialect, which may lead to variations in ACA from one country to another. For example, the guttural ق is pronounced as a guttural g in Gulf Arabic but as a ء glottal stop in the Egyptian Arabic dialects. Therefore, it is mapped with the number 2 in the Egyptian Arabizi e.g., the word قلبي which means heart is written as "galby" in the Gulf Arabizi and written as 2alby in the Egyptian Arabizi (Tobaili et al. 2019).

Arabic and Franco Arabic are considered a very sophisticated language to be analyzed and handled by computer systems than other languages, such as English, for many reasons, such as the previously mentioned Arabic variations in different countries, the complexity of its morphology, and the inconsistent and complex orthography (Alyami et al. 2022).

In conflict with the growth of Arab internet users nowadays, the number of written Arabic reviews and web content is relatively low compared to English. Also, few data sets are available for SA, and the number of Arabic ABSA studies and research covered by the scientific community is limited (Bensoltane and Zaki 2023)

**Table 3: Arabic and ACA Mapping**

| Char. in Arabic | Char. in Arabizi | Char. in Arabic | Char. in Arabizi |
|---|---|---|---|
| أ | A | ط | T |
| ب | B | ظ | Z OR 6' |
| ت | T | ع | 3 |
| ث | TH | غ | GH OR 3' |
| ج | G OR J | ف | F |
| ح | 7 | ق | K OR 2 |
| خ | kh OR 5 | ك | K |
| د | D | ل | L |
| ذ | Z | م | M |
| ر | R | ن | N |
| ز | Z | ه | H |

| | | | |
|---|---|---|---|
| س | S | و | W OR O |
| ش | SH | ي | Y |
| ص | S OR 9 | ة | 2 |
| ض | D OR 9’ | ء | 2 |

## Literature Review

To our knowledge, the ABSA in Arabic languages was started by M. Al-Smadi et al., (2015) in Jordan. This paper introduced the first benchmark Arabic dataset, (haad), which consists of book reviews written by Arabic readers. The dataset was preprocessed into xml files and annotated using brat, a web-based annotation tool, after which a baseline

(Areed et al. 2020) generated a data set of government smart app reviews in Arabic, domain aspects, and opinion words. The reviews were manually annotated, and sentiment scores were assigned to all identified opinion words to construct the required sentiment lexicons. The proposed approach integrated lexicon-based techniques with rule-based models to improve sentiment classification accuracy.

(Alhamad and Kurdy 2020) presented a feature-based SA solution that was specifically designed for the Arabic language. It conducts text analysis by decomposing the input into distinct aspects, attributes, and components related to a specific product or service. The reviews were processed and rewritten in xml files; this was followed by sentiment allocation which was done with the Gatex application.

More recently (Voskergian and Saheb 2022) a preprocessed Arabic dataset accompanied by descriptive statistics and a baseline evaluation for ABSA. The data set contains reviews from an Arabic mobile application collected from the three top music applications in the Middle East: Anghami[xi], Spotify, and SoundCloud[xii]. The collected dataset consists of approximately 100,000 application reviews, including 45,769 from Anghami, 31,531 from Spotify, and 23,893 from SoundCloud. Data collection was con- ducted using the Python-based library Google-Play-Scraper. A subset of 2,000 reviews was annotated manually, identifying relevant aspects and assigning polarity values to support subsequent research and analysis.

(Al-Smadi et al. 2022) presented a new ABSA model for application to Arabic reviews. The dataset containing the restaurant reviews was utilized and categorized according to four predefined aspects: price, cleanliness, food quality, and service. A hybrid approach combining ml techniques with domain-specific dictionaries and sentiment word lists was adopted. More than 3,000 reviews were collected from the Jeeran[xiii] restaurant reviews website, a well- known site for reviewing and rating restaurants in Jordan. The annotation of the reviews was done crowdsourcing approach, supported by a dedicated website developed using theasp.net[xiv] framework. A my-sql[xv] database was employed to display the reviews and record the annotations. The annotated data underwent preprocessing, which involved removing punctuation and redundant letter repetitions. The NLTK[xvi] tokenizer was then applied to segment each sentence into individual words. Subsequently, dictionaries and sentiment word lists were extracted from the dataset. Four aspect-specific models were developed using the Support Vec- tor Machine (SVM) model, and another four models were built using Na¨ive Bayes (NB) classifiers, one model for each aspect. In all experiments, the performance of SVM was better than that of the NB classifier.

Ultimately, (Alqarni and Rahman 2023) presented an ABSA model for the tweets written in Arabic language related to the covid-19 pandemic in Saudi Arabia. The collection of the datasets was across two distinct time periods from Riyadh, Dammam, and Jeddah. The annotation process was made manually with three sentiment categories: positive, negative, and neutral. The pre- processing phase involved the manual removal of diacritics, elongated characters, and unnecessary letter repetitions. Convolutional neural networks (CNN) and bi-directional long short memory (BIlSTM). Deep learning algorithms were employed to classify the sentiment of Arabic tweets and determine their corresponding polarity.

Recent efforts in creating lexical resources for Arabic focused mainly on MSA and very few on the ACA (Duwairi et al. 2016). To the best of our knowledge, the first paper focusing on ABSA in ACA is by (Duwairi et al. 2016). They presented a framework for analyzing tweets written in aca in the Jordanian Dialect. Other alternatives in the literature include:

Other studies in literature include the one by Guellil et al., (2018). A novel approach was presented by this study for the automatic construction of sentiment annotation an Algerian data set for the, which is different that Magherbi Arabic.

(Fsih et al. 2022.)proposed and examined a Tunisian dialect corpus. By analyzing this corpus, the efficiency of ml was compared and transferred learning models to perform fine-grained SA.

Ultimately, Mansy et al. (2023) presented a new word embedding model, Ara-Franco, trained using a large corpus of user-generated tweets written in Franco-Arab (aca/Arabizi) with different dialects.

However, to our knowledge, only Gahbiche et al. (2020) and Gabr et al. (2023) applied ABSA in Arabizi with the Egyptian dialect. Gahbiche et al. (2020) proposed lad, a bench- mark data set, and annotated 48 common words of the Egyptian Arabizi to create a subset called the salad. (Gabr et al. 2023) translated Arabic statements into Franco Arabic due to the limitations of the Franco Arabic data sets. They implemented special Python functions to normalize Arabic letters, remove diacritics, and Arabic punctuations.

The above-mentioned literature is summarized in Table 4

**Table 4: Summary of the Literature Review**

| Study | Language and Dialect | Dataset | Methodology | Main Contribution |
|---|---|---|---|---|
| (Al-Smadi et al. 2015) | MSA | Book reviews (HAAD) | XML annotation using BRAT; baseline ABSA evaluation | Introduced the first benchmark ABSA dataset for Arabic. |
| (Areed et al. 2020) | Arabic | Government smart app reviews | Lexicon- and rule-based approach; sentiment lex-icon construction | Combined lexicon-based and rule-based models for improved sentiment classification. |
| (Alhamad and Kurdy 2020) | Arabic | Product/service reviews | Feature-based SA; sentiment allocation using GATE | Developed an aspect-level sentiment model for Arabic using text decomposition. |
| (Voskergian and Saheb 2022) | Arabic | App reviews (Anghami, Spotify, SoundCloud) | Python Google Play Scraper; manual annotation (2,000 reviews) | Built a large-scale Arabic ABSA dataset (100k reviews) with baseline evaluation. |
| (Al-Smadi et al., 2022) | Arabic | Restaurant Reviews (Jeeran.com) | SVM and Na¨ıve Bayes with lexicon & domain dictionaries | Multi-aspect ABSA (price, service, food, cleanliness) using a hybrid ML-lexicon approach. |
| (Alqarni and Rahman 2023) | Arabic | COVID-19 tweets (Saudi Arabia) | CNN and BiLSTM deep learning models | Applied deep learning to Arabic tweet sentiment classification. |
| (Duwairi et al. 2016) | Jordanian ACA | Tweets | Lexicon-based ABSA framework | First ABSA study in Arabic Chat Alphabet (ACA). |
| (Guellil et al. 2018) | Algerian Dialect | Tweets | Automatic sentiment corpus construction | Introduced the first ABSA corpus for the Algerian dialect. |
| (Fsih et al. n.d.) | Tunisian Dialect | Social media corpus | ML and transfer learning comparison | Compared traditional ML and transfer learning for Tunisian ABSA. |
| (Mansy et al. 2023) | Franco-Arab (Arabizi) | Twitter data | Ara-Franco word embeddings | Developed the first embedding model for Franco-Arab dialects. |

| | | | | |
|---|---|---|---|---|
| (Gahbiche et al. 2020) | Egyptian Arabizi | L-AD & SALAD datasets | Manual annotation of 48 Egyptian Arabizi words | Proposed the first ABSA dataset for Egyptian Arabizi. |
| (Gabr et al. 2023) | Egyptian Arabizi | Translated Arabic dataset | Custom Python normalization; diacritic removal | Applied ABSA to Egyptian Arabizi using data normalization and transla-tion. |

## Conclusion

In light of this review, a noticeable research gap exists in ABSA applied to aca. The demand for developing new models capable of mining and evaluating the ACA continues to grow, particularly for the Egyptian Dialect also known as "Egyptizi". To the best of our knowledge, no prior studies have specifically analyzed customer reviews written in the Egyptian Arabizi.

As part of future work, we aim to develop an ABSA model tailored for analyzing reviews written in Egyptian Arabizi. This model is expected to provide significant value to Egyptian organizations by enabling a deeper understanding of customer opinions, satisfaction levels, and preferences regarding the products and services they offer.

---

[i] Available at https://www.netflix.com, November 2025.

ii Available at https://www.amazon.com, November 2025.

iii Available at https://www.starbucks.com, November 2025.

iv Available at https://www.spotify.com, November 2025.

v Available at https://www.airbnb.com, November 2025.

[vi] Available at https://www.facebook.com, November 2025

[vii] Available at https://x.com, November 2025.

[viii] Available at https://www.threads.com, November 2025.

[ix] Available at https://www.instagram.com, November 2025

[x] Available at https://gate.ac.uk/applications/sentiment.html, November 2025.

[xi] Available at https://www.anghami.com, November 2025

[xii] Available at https://soundcloud.com, November 2025.

[xiii] Available at https://jeeran.com, November 2025.

[xiv]Available at https://dotnet.microsoft.com/apps/aspnet, November 2025.

[xv] Available at https://www.mysql.com, November 2025.

[xvi] Available at https://www.nltk.org, November 2025.

## References


- Agrawal, N.K. (2021) An Analytical Study of Digital Marketing and Use of Emerging Tools. 2021 5th International Conference on Information Systems and Computer Networks (ISCON) IEEE1–4.
- Alhamad, G. and Kurdy, M.-B. (2020) Feature-Based Sentiment Analysis for Arabic Language. International Journal of Advanced Computer Science and Applications 11 (11), .
- Alqarni, A. and Rahman, A. (2023) Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach. Big Data and Cognitive Computing 7 (1), 16.

- Al-Smadi, F., Al-Shboul, B., Al-Darras, D. and Al-Qudah, D. (2022) Aspect-Based Sentiment Analysis of Arabic Restaurants Customers' Reviews Using a Hybrid Approach. Proceedings of the 14th International Conference on Management of Digital EcoSystems New York, NY, USA: ACM123–128.
- Al-Smadi, M., Qawasmeh, O., Talafha, B. and Quwaider, M. (2015) Human Annotated Arabic Dataset of Book Reviews for Aspect Based Sentiment Analysis. 2015 3rd International Conference on Future Internet of Things and Cloud IEEE726–730.
- Alyami, S., Alhothali, A. and Jamal, A. (2022) Systematic literature review of arabic aspect-based sentiment analysis. Journal of King Saud University - Computer and Information Sciences 34 (9), 6524–6551.
- Alzu'Bi, S., Alsmadiv, A., Alqatawneh, S., Al-Ayyoub, M., Hawashin, B. and Jararweh, Y. (2019) A Brief Analysis of Amazon Online Reviews. 2019 6th International Conference on Social Networks Analysis, Management and Security, SNAMS 2019 .
- Areed, S., Alqaryouti, O., Siyam, B. and Shaalan, K. (2020) Aspect-Based Sentiment Analysis for Arabic Government Reviews. 143–162.
- Arora, N., Dreze, X., Ghose, A., Hess, J.D., Iyengar, R., Jing, B., Joshi, Y., Kumar, V., Lurie, N., Neslin, S., Sajeesh, S., Su, M., Syam, N., Thomas, J. and Zhang, Z.J. (2008) Putting one-to-one marketing to work: Personalization, customization, and choice. Marketing Letters 19 (3–4), 305–321.
- Bayoude, K., Ouassit, Y., Ardchir, S. and Azouazi, M. (2018) How machine learning potentials are transforming the practice of digital marketing: State of the art. Periodicals of Engineering and Natural Sciences (PEN) 6 (2), 373.
- Behera, R.K., Gunasekaran, A., Gupta, S., Kamboj, S. and Bala, P.K. (2020) Personalized digital marketing recommender engine. Journal of Retailing and Consumer Services 53, 101799.
- Bensoltane, R. and Zaki, T. (2023) Aspect-based sentiment analysis: an overview in the use of Arabic language. Artificial Intelligence Review 56 (3), Springer Nature2325–2363.
- Borah, U. (2008) Digital Marketing: An Analysis. International Journal of Advances in Engineering and Management (IJAEM 2 (1), 582–586.
- Chandra, S., Verma, S., Lim, W.M., Kumar, S. and Donthu, N. (2022) Personalization in personalized marketing: Trends and ways forward. Psychology & Marketing 39 (8), 1529–1562.
- Desai, Dr.Mrs.V. (2019) Digital Marketing: A Review. International Journal of Trend in Scientific Research and Development Special Issue (Special Issue-FIIIIPM2019), 196–200.
- Do, H.H., Prasad, P., Maag, A. and Alsadoon, A. (2019) Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review. Expert Systems with Applications 118, 272–299.
- Duwairi, R.M., Alfaqeh, M., Wardat, M. and Alrabadi, A. (2016) Sentiment analysis for Arabizi text. 2016 7th International Conference on Information and Communication Systems (ICICS) IEEE127–132.
- Fsih, E., Boujelbane, R. and Belguith, L.H. (n.d.) A deep sentiment analysis of Tunisian dialect comments on multi-domain posts in different social media platforms. .
- Gabr, M.A.-E.I.I., Badr, A.Z. and Mahdi, H.M.K. (2023) FROM LATIN TO ARABIC: INNOVATIVE APPROACHES FOR RETRIEVING ORIGINAL TEXTS AND VECTORIZING FRANCO-ARABIC. Indian Journal of Computer Science and Engineering 14 (2), ENGG Journals Publications289–300.
- Gahbiche, S., Gadek, G., Baert, G. and Pauchet, A. (2020) Arabizi Language Models for Sentiment Analysis. COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference Association for Computational Linguistics (ACL)592–603.
- Guellil, I., Adeel, A., Azouaou, F. and Hussain, A. (2018) SentiALG: Automated Corpus Annotation for Algerian Sentiment Analysis. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) .

- Hassan, M.A., Shehab, M.E. and Hamed, E.M.R. (2016) A comparative study of classification algorithms in e-health environment. 2016 Sixth International Conference on Digital Information Processing and Communications (ICDIPC) IEEE42–47.
- Hu, M. and Liu, B. (2004) Mining and summarizing customer reviews. Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining New York, NY, USA: ACM168–177.
- Jafarian, H., Taghavi, A.H., Javaheri, A. and Rawassizadeh, R. (2021) Exploiting BERT to Improve Aspect-Based Sentiment Analysis Performance on Persian Language. 2021 7th International Conference on Web Research, ICWR 2021 Institute of Electrical and Electronics Engineers Inc.5–8.
- Kannan, P.K. and Li, H. "Alice" (2017) Digital marketing: A framework, review and research agenda. International Journal of Research in Marketing 34 (1), 22–45.
- Khurana, D., Koli, A., Khatter, K. and Singh, S. (2023) Natural language processing: state of the art, current trends and challenges. Multimedia Tools and Applications 82 (3), 3713–3744.
- Lee, S., Choi, W. and So, J. (2022) A Study on the Factors Affecting Customer Satisfaction in Delivery Applications: Focusing on Sentiment Analysis of Review Data. Proceedings - 2022 IEEE/ACIS 7th International Conference on Big Data, Cloud Computing, and Data Science, BCD 2022 Institute of Electrical and Electronics Engineers Inc.34–38.
- Liu, B. (n.d.) Sentiment Analysis and Subjectivity. .
- Mansy, A., Rady, S. and Gharib, T. (2023) AraFranco: An Arabic Word Embedding Model for Representing Franco-Arabic User-Generated Text. Lecture Notes on Data Engineering and Communications Technologies Springer Science and Business Media Deutschland GmbH. 587–601.
- Mariossbelk, P. (2016) Human-Computer Interaction Series Human-Centred Web Adaptation and Personalization From Theory to Practice. .
- Molina Parra, P.A., Botero Botero, S. and Montoya Monsalve, J.N. (2016) Family business: concepts and models for analysis. Revista científica Pensamiento y Gestión NA (41), 116–149.
- Mostafa, L. (2012) A SURVEY OF AUTOTAMTED TOOLS FOR TRANSLATING ARAB CHAT ALPHABET INTO ARABIC LANAGUAGE. Academic & Scholarly Research Journal .
- Nazir, A., Rao, Y., Wu, L. and Sun, L. (2022) Issues and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey. IEEE Transactions on Affective Computing 13 (2), 845–863.
- Nikolajeva, A. and Teilans, A. (2021) Machine learning technology overview in terms of digital marketing and personalization. Proceedings - European Council for Modelling and Simulation, ECMS .
- Oueslati, O., Cambria, E., HajHmida, M. Ben and Ounelli, H. (2020) A review of sentiment analysis research in Arabic language. Future Generation Computer Systems 112, 408–430.
- Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., María Jiménez-Zafra, S. and Eryiğit, G. (n.d.) SemEval-2016 Task 5: Aspect Based Sentiment Analysis. .
- Ramos-Soto, A., Bugarín, A. and Barro, S. (2016) On the role of linguistic descriptions of data in the building of natural language generation systems. Fuzzy Sets and Systems 285, 31–51.
- Saura, J.R. (2021) Using Data Sciences in Digital Marketing: Framework, methods, and performance metrics. Journal of Innovation & Knowledge 6 (2), 92–102.
- Schouten, K. and Frasincar, F. (2016) Survey on Aspect-Level Sentiment Analysis. IEEE Transactions on Knowledge and Data Engineering 28 (3), IEEE Computer Society813–830.
- Sunikka, A. and Bragge, J. (2012) Applying text-mining to personalization and customization research literature – Who, what and where? Expert Systems with Applications 39 (11), 10049–10058.

- Tobaili, T., Fernandez, M., Alani, H., Sharafeddine, S., Hajj, H. and Glavaš, G. (2019) Senzi: A sentiment analysis lexicon for the latinised Arabic (Arabizi). International Conference Recent Advances in Natural Language Processing, RANLP Incoma Ltd1203–1211.
- Tripathy, A., Anand, A. and Rath, S.K. (2017) Document-level sentiment classification using hybrid machine learning approach. Knowledge and Information Systems 53 (3), Springer London805–831.
- Voskergian, D. and Saheb, M.H. (2022) AMAR_ABSA: Arabic Mobile App Reviews Dataset Targeting Aspect-based Sentiment Analysis Tasks. 2022 Innovations in Intelligent Systems and Applications Conference (ASYU) IEEE1–7.
- Wang, J., Xu, B. and Zu, Y. (2021) Deep learning for Aspect-based Sentiment Analysis. 2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE) IEEE267–271.
- Yadav, V., Verma, P. and Katiyar, V. (2021) E-Commerce Product Reviews Using Aspect Based Hindi Sentiment Analysis. 2021 International Conference on Computer Communication and Informatics (ICCCI) IEEE1–8.