

محمد القحطاني

Nasser-Mohammed survey 03.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:398990756

Submission Date

Jul 16, 2026, 12:05 AM GMT+3

Download Date

Jul 16, 2026, 11:36 AM GMT+3

File Name

Nasser-Mohammed survey 03.docx

File Size

108.5 KB

17 Pages

5,102 Words

30,557 Characters

24% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

-  **67 Not Cited or Quoted 22%**
Matches with neither in-text citation nor quotation marks
-  **11 Missing Quotations 2%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 15%  Internet sources
- 16%  Publications
- 21%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 67 Not Cited or Quoted 22%**
Matches with neither in-text citation nor quotation marks
- 11 Missing Quotations 2%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 15% Internet sources
- 16% Publications
- 21% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Submitted works	Midocean University on 2026-05-02	7%
2	Internet	www.mdpi.com	1%
3	Internet	mdpi-res.com	1%
4	Internet	aclanthology.org	<1%
5	Internet	www.lrec-conf.org	<1%
6	Internet	peerj.com	<1%
7	Internet	theacademic.in	<1%
8	Submitted works	King Abdulaziz University on 2025-01-12	<1%
9	Internet	www.jetir.org	<1%
10	Submitted works	IUBH - Internationale Hochschule Bad Honnef-Bonn on 2026-06-07	<1%

11	Internet	aajsr.com	<1%
12	Internet	www.nvidia.com	<1%
13	Submitted works	British University In Dubai on 2023-05-08	<1%
14	Submitted works	Monash University on 2026-05-22	<1%
15	Submitted works	Mapua University on 2026-07-08	<1%
16	Submitted works	University of Anbar on 2023-08-18	<1%
17	Internet	www.ltrc.lsu.edu	<1%
18	Publication	Salma Khaled, Walaa Medhat, Ensaf Hussein Mohamed. "UniTextFusion: A low-res...	<1%
19	Submitted works	Liverpool John Moores University on 2022-03-03	<1%
20	Internet	tns.ewapub.com	<1%
21	Internet	ttu-ir.tdl.org	<1%
22	Publication	Lamia Ouchene, Sadik Bessou. "Benchmarking sentiment analysis of algerian ara...	<1%
23	Internet	admin.thescipub.com	<1%
24	Internet	ijisrt.com	<1%

25	Publication	Caroline Sabty, Islam Mesabah, Özlem Çetinoğlu, Slim Abdennadher. "Language I...	<1%
26	Submitted works	Midocean University on 2025-09-05	<1%
27	Publication	Souha Al Katat, Chamseddine Zaki, Hussein Hazimeh, Ibrahim Bitar, Rafael Angar...	<1%
28	Submitted works	Midocean University on 2026-05-02	<1%
29	Publication	Pradeep Kumar Roy, Asis Kumar Tripathy, Abhinav Kumar, Yulei Wu. "Social Netw...	<1%
30	Internet	dokumen.pub	<1%
31	Publication	Han Yi, Peilin Lv, Feng Wang, Chao Chen. "Simultaneous multi-component quantif...	<1%
32	Publication	Nora Alturayeif, Hamzah Luqman. "Fine-Grained Sentiment Analysis of Arabic CO...	<1%
33	Internet	public-pages-files-2025.frontiersin.org	<1%
34	Submitted works	Aston University on 2023-05-27	<1%
35	Submitted works	IBAZ on 2025-12-05	<1%
36	Submitted works	Liverpool John Moores University on 2021-03-12	<1%
37	Publication	M. Maheswari, Ammar Hameed Shnain, Ayesha Siddiqua, H C Nagaraj, N. Shilpa. "...	<1%
38	Submitted works	Macao Polytechnic Institute on 2026-04-15	<1%

39	Publication	Ramzi Haraty, Mira Chehade. "Transfer learning and sentiment analysis of lebane...	<1%
40	Submitted works	Thesis 2026 on 2026-07-03	<1%
41	Submitted works	University of Hertfordshire on 2025-09-03	<1%
42	Internet	amtaweb.org	<1%
43	Internet	arxiv.org	<1%
44	Internet	fjs.fudutsinma.edu.ng	<1%
45	Internet	irep.ntu.ac.uk	<1%
46	Internet	metall-mater-eng.com	<1%
47	Publication	"Artificial Intelligence and Data Analytics for Innovative Applications in Engineeri...	<1%
48	Submitted works	Asia Pacific University College of Technology and Innovation (UCTI) on 2025-09-28	<1%
49	Submitted works	Asia Pacific University College of Technology and Innovation (UCTI) on 2026-04-26	<1%
50	Submitted works	Association of Educators on 2023-11-08	<1%
51	Publication	Ewa Ziemba, Jarosław Wątróbski. "Adoption of Emerging Information and Comm...	<1%
52	Publication	Ferihane Kboubi, Anja Habacha Chaibi. "Fine-Grained Arabic Dialect Identification...	<1%

53	Submitted works	ICSDI on 2025-12-25	<1%
54	Publication	Mahander Kumar, Lal Khan, Hsien-Tsung Chang. "Evolving techniques in sentime...	<1%
55	Publication	Tarek Rashed, Ramadan Alfared, Abduelbaset Goweder, Husien Alhammi, Abuba...	<1%
56	Publication	"Advances in Brain Inspired Cognitive Systems", Springer Science and Business M...	<1%
57	Publication	Arwa Alqarni, Atta Rahman. "Arabic Tweets-Based Sentiment Analysis to Investig...	<1%
58	Submitted works	City University of Macau on 2026-01-30	<1%
59	Publication	Hanane Elfaik, El Habib Nfaoui. "Leveraging feature-level fusion representations ...	<1%

28	 مجلة ميدأوشن للبحوث والدراسات	Midocean Journal for Research and Studies https:// e-ISSN:	
----	---	---	--

Arabic Sentiment Analysis: A Survey of Methods, Challenges, and
Future Directions

26

Nasser Alshalwan – Mohammed Alqahtani

Master of Science in Artificial Intelligence – College of Informatics – Midocean University

—

Email: mdulaim@outlook.sa

Dr. Wesam Ahmed, College of Informatics – Midocean University –

Abstract:

Arabic sentiment analysis is an important task in natural language processing because it helps identify opinions, emotions, and attitudes expressed in Arabic text. With the advent of social media and online communication such as reviews, this task has become of greater importance. However, Arabic sentiment analysis remains challenging because Arabic includes Modern Standard Arabic and many regional dialects, which differ in vocabulary, spelling, grammar, and writing style. In addition, social media text often contains informal expressions, emojis, repeated letters, sarcasm, code-switching, and noisy writing. This survey reviews Arabic sentiment analysis methods, challenges, and future directions. It discusses traditional machine learning methods, deep learning models, transformer-based models, and hybrid approaches used in Arabic sentiment classification. The reviewed studies show that transformer-based models such as AraBERT, MARBERT, and CAMELBERT have improved Arabic sentiment analysis because they can represent words according to context. The survey also shows that hybrid approaches are promising because they can combine the strengths of different models and text representations. However, challenges such as dialect variation, domain differences, dataset limitations, noisy user-generated text, computational cost, and evaluation inconsistency still remain. Future research should focus on more diverse datasets, robust models, efficient hybrid methods, and clearer evaluation benchmarks.

Keywords: Arabic Sentiment Analysis; Machine Learning; Deep Learning; Transformer Models; Hybrid Approaches

الملخص:

يُعدُّ تحليل المشاعر العربية من المهام المهمة في معالجة اللغة الطبيعية، إذ يسهم في التعرف على الآراء والمشاعر والاتجاهات التي تعبر عنها النصوص العربية. وقد ازدادت أهمية هذا المجال مع التوسع الكبير في استخدام وسائل التواصل الاجتماعي، والمراجعات الإلكترونية، وأنماط التواصل الرقمي المختلفة. ومع ذلك، لا يزال تحليل المشاعر العربية يواجه تحديات متعددة؛

نظرًا لتنوع العربية بين الفصحى والحديثة ولهجات إقليمية عديدة تختلف في المفردات، والتهجئة، والتراكيب، وأساليب التعبير. كما تزيد نصوص وسائل التواصل الاجتماعي من صعوبة المهمة بسبب ما تتضمنه من تعبيرات غير رسمية، ورموز تعبيرية، وتكرار للحروف، وسخرية، وتناوب لغوي، وكتابة غير معيارية. يستعرض هذا البحث المسحي أبرز طرائق تحليل المشاعر العربية وتحدياتها واتجاهاتها المستقبلية، من خلال تناول أساليب التعلم الآلي التقليدية، ونماذج التعلم العميق، والنماذج القائمة على المحولات، والأساليب الهجينة المستخدمة في تصنيف المشاعر العربية. وتبين الدراسات التي جرى استعراضها أن النماذج القائمة على المحولات، مثل AraBERT و MARBERT و CAMELBERT، أسهمت في تحسين أداء تحليل المشاعر العربية بفضل قدرتها على تمثيل الكلمات وفق سياقها. كما تشير النتائج إلى أن الأساليب الهجينة تُعدُّ اتجاهًا واعدًا؛ لأنها تجمع بين مزايا نماذج وتمثيلات نصية مختلفة. ومع ذلك، ما تزال تحديات تنوع اللهجات، واختلاف المجالات، ومحدودية البيانات، وضوضاء النصوص التي ينتجها المستخدمون، وارتفاع التكلفة الحسابية، وتباين أساليب التقييم قائمة. لذلك، توصي الدراسة بالتركيز على بناء بيانات أكثر تنوعًا، وتطوير نماذج أكثر قدرة على التعميم، وتصميم أساليب هجينة فعالة، واعتماد معايير تقييم أكثر وضوحًا.

الكلمات المفتاحية: تحليل المشاعر العربية؛ التعلم الآلي؛ التعلم العميق؛ نماذج المحولات؛ الأساليب الهجينة

1. Introduction

Sentiment analysis is a key task in natural language processing (NLP). It refers to the task of extracting opinions, emotion, attitudes, and stance expressed in written data. It has gained more attention with the rapid increase of online communication, and it is widely used on social media, review websites, discussion forums, and other Internet sources where users express their opinion in textual format every day. Arabic sentiment analysis has attracted recent interest; however, it is a more challenging task than in high-resource languages due to the linguistic nature of Arabic. Arabic language exists in different forms, such as Modern Standard Arabic and many dialectal Arabic forms, which can differ in vocabulary, lexicon, writing style, and grammar. This variation

makes automatic sentiment classification more difficult, especially in noisy and informal user generated data such as tweets and comments that include emojis, spelling errors, repeated characters, emphasis and sarcasm, code-switching and mixing different dialects. The problem is also more difficult across domains where sentiment systems learned from one domain, such as product review domains, may not work well in other domains, such as forums or news comments. For example, the same words or expressions might have different sentiment interpretation or importance across domains or dialects, and the variation of this becomes even more diverse over social media. Main past work shows that cross-domain and cross-dialect Arabic sentiment classification remains a hard problem and more robust methods are needed to account for this linguistic variation (El Mekki et al., 2021). Moreover, a recent survey showed that Arabic sentiment analysis is still less advanced than in high resource languages especially in the area of deep learning and transformer based deep models (Shi & Agrawal, 2025). The development of Arabic sentiment analysis passed through three major stages. Traditional Machine learning classifiers, such as logistic regression and support vector machines were widely applied in combination with term frequency-inverse document frequency (TF-IDF) features. Later, Deep neural network models, such as convolutional neural networks and long short term memory were used to learn more useful text features from data. Recently, transformer based models, such as AraBERT, MARBERT, and CAMELBERT were demonstrated to have powerful contextual understanding ability when representing Arabic texts. Some of these models are especially useful for dialectal and noisy Arabic texts. The main issues with existing work are that most models are tested on smaller datasets, specific dialects or limited domain settings, some models are computationally expensive or are hard to understand, and benchmark quality, data bias, data pre-processing and evaluation setting can influence the results. Therefore a general survey is needed to review existing state of the art methods for Arabic sentiment analysis, highlight strengths and shortcomings of them, and suggest practical future directions. This survey paper categorizes Arabic sentiment analysis methods into three main classes: traditional machine learning approaches, deep learning approaches, and transformer based approaches, and also presents hybrid approaches combining the contextual representation power of transformer models with deep learning or machine learning classifiers. The value of this survey is that it defines the challenges of Arabic sentiment analysis and proposes practical future directions to develop balanced, domain-sensitive, and more efficient methods.

2. Previous Studies

In this section, the common Arabic sentiment analysis approaches are discussed. The approach used is traditional machine learning approach, deep learning approach, Transformer-based approaches, and the most commonly used evaluation metrics. This section is crucial to illustrate the evolution of Arabic sentiment analysis approaches starting from the use of feature-based classifiers up to advanced and sophisticated methods based on contexts.

2.1. Sentiment Analysis using Machine Learning

Another frequently adopted method in sentiment analysis is (ML) where the text is converted into numerical features such as bag-of-words and (TF-IDF), and then classification models are used to assign the sentiment class. In the field of Arabic sentiment analysis, logistic regression, (SVM), naive Bayes, decision tree, and XGBoost were employed as classifiers to classify the texts into positive, negative and neutral sentiments. They have been employed because of their simplicity, interpretability and are practical to variety of text classification works but, due to Arabic's rich morphology, variety in dialects, spelling in an informal manner, and social media's writings (noisy), the ML methods' performances rely heavily on the quality of preprocessing and feature extraction steps (Alyami & Olatunji, 2020; Taher, 2024).

2.1.1. Logistic Regression

Logistic Regression is one of the traditional machine learning classifiers for text classification and sentiment analysis which approximates the probability that a text belongs to a specific sentiment class using numerical features obtained from it. It is often used along with (TF-IDF) in text feature representation. It is relatively fast, easy to explain, and simpler classifier that makes its efficient use for several classification tasks. However, it relies on the quality of features extracted and it cannot effectively understand complex contexts or specific slang meanings that can occur in the text (Aljabri et al., 2021; Taher, 2024).

2.1.2. Support Vector Machine

Support vector machine (SVM) is another classification technique popular for use in sentiment analysis. This technique identifies a classification boundary based on the features of each text in numerical form to separate them into sentiment classes. This machine learning classifier performs well on many text-based categorization issues, primarily as it works well with feature spaces that have high dimensionality such as (TF-IDF) vectors. In terms of Arabic language-based sentiment classification, (SVM) is employed for various text based content like tweets, into various sentiment types. Just like all the machine learning techniques, this model can be significantly affected by issues such as processing, the characteristics of features, and the challenges involved

in interpreting various dialect and informal word usages of the Arabic language (Alyami & Olatunji, 2020).

2.1.3. XGBoost

XGBoost is a machine learning model based on gradient boosting decision trees. It builds several decision trees step by step, where each new tree tries to correct the errors made by previous trees. In sentiment analysis, XGBoost can be used after converting text into numerical features such as TF-IDF. It is useful because it can model non-linear relationships between features and sentiment classes. However, it still depends on the quality of text representation and may require careful parameter tuning. In Arabic sentiment analysis, this can be difficult when the data includes dialects, noisy writing, and informal social media language (Alghamdi, 2024; Taher, 2024).

2.2. Sentiment Analysis using Deep Learning

Deep learning (DL) is a subset of machine learning that uses neural networks in learning patterns from data. (DL) models can automatically learn text features instead of using more manually designed features to carry out a task like sentiment analysis. These models can capture word order, localized expressions and patterns within a sentence, etc. (DL) models have been used in the analysis of Arabic sentiment, as Arabic text presents many challenges such as morphology, dialectal variations, and non-standard word usage. These models, however, have been found to demand more data, training time and more hyperparameter tuning (Elhassan et al., 2023; Minaee et al., 2021).

2.2.1. Convolutional Neural Network (CNN)

Convolutional neural network (CNN), which is one of the deep learning methods, are suitable for classifying the text, for the sentiment classification, (CNN) finds meaningful local patterns in text such as short word phrases or combination of words related to positive/negative tone, as they are capable of extracting features by themselves. The drawback is they may perform poorly where sentiment depends on distant context, relationship of distant words and context of entire sentence (Elhassan et al., 2023; Minaee et al., 2021).

2.2.2. Long Short-Term Memory (LSTM)

Long short-term memory (LSTM) is a recurrent neural network model designed to process sequential data. In sentiment analysis, (LSTM) reads the text as a sequence of words and learns how earlier words may affect later meaning. This makes it useful for capturing sentence structure and word order. In Arabic sentiment analysis, (LSTM) can help model longer expressions and contextual sentiment clues. However, it may require more training time and can be affected by limited or noisy Arabic datasets (Basiri et al., 2021; Alosaimi et al., 2024).

2.2.3. Bidirectional Gated Recurrent Unit (BiGRU)

Bidirectional gated recurrent unit (BiGRU) is a deep learning model based on recurrence and capable of reading text from front to back and from back to front. This mechanism enhances the learning of context prior to and subsequent to each word. In the task of sentiment analysis, (BiGRU) can perform well when a word meaning is reliant on other nearby words and is more straightforward than several recurrent architectures. However, the efficiency is largely reliant on the training data set as well as the configurations made on the model (Alhumoud, 2024; Almaqtari et al., 2024).

2.3. Sentiment Analysis Using Transformer-Based Models

Transformer-based models are recent deep learning models widely used in many natural language processing tasks. These models use an attention mechanism to represent the relationships between words in a sentence, which means that the words can be represented in the context in which they are used, as opposed to representing words as standalone words, this is very beneficial for Arabic sentiment analysis as the word can have different meanings depending on the dialect, context and writing style of the sentence, and we will be able to cater for different styles and contexts, therefore being most useful for Arabic text classification comparing to other traditional machine learning and simple deep learning models (Devlin et al., 2019; Antoun et al., 2020; Abdul-Mageed et al., 2021; Inoue et al., 2021).

2.3.1. AraBERT

AraBERT is an Arabic transformer-based model designed for Arabic language understanding. It was pre-trained on large Arabic text data to support different Arabic natural language processing

tasks. Regarding the sentiment analysis tasks, AraBERT models the context of each Arabic word, so that the Arabic word representations will improve classification where the expressions or the structure of the sentence is changing and thus, AraBERT is also a widely adopted architecture in the Arabic sentiment classification task, due to its powerful Arabic representation, but its accuracy might not improve with slang and noise on the social media, and also it might perform differently according to the dataset used in pre-training and fine-tuning the model (Antoun et al., 2020; Alosaimi et al., 2024).

2.3.2. MARBERT

MARBERT was developed as an Arabic transformer-based language model, with special emphasis on dialectal Arabic and the analysis of social media texts. The model was trained on large volumes of Arabic tweets and online vernacular, making it highly suitable for Arabic sentiment analysis tasks on datasets containing social media text that may contain dialects, informal orthography, and other sources of noise typical of user-generated language. This allows MARBERT to grasp contextual meanings in social media Arabic more effectively than models that primarily focus on standard Arabic. MARBERT could pose challenges, however, if further adaptation (fine-tuning) is needed or computational limitations are a concern when using conventional approaches (Abdul-Mageed et al., 2021; Belbachir, 2023).

2.3.3. CAMeLBERT

CAMeLBERT is another Arabic pre-trained language model developed to support Arabic natural language processing tasks across different language variants. It was introduced to study the effect of Arabic language variety, model size, and task type on model performance. In sentiment analysis, CAMeLBERT can represent Arabic text based on context and can be useful for both formal and dialectal Arabic. But like all transformer-based models, performance is based on dataset, task target domain, and fine-tuning (Inoue et al., 2021; Mansour et al., 2025).

2.4. Performance Metrics

Performance metrics are used to evaluate how well sentiment classification models predict the correct sentiment labels. In classification tasks, the results are commonly described using four outcomes: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

31

24

48

TP and TN refer to correctly classified positive and negative cases, while FP and FN refer to incorrectly classified cases. Common metrics used in sentiment analysis include accuracy, precision, recall, and F1-score. Accuracy measures how many predictions were right across the board, precision measures how many of the predicted positives were right, and recall measures how many of the right positives were found. F1-score, similarly to the accuracy calculates the balance of precision and recall as one index. It is used for instance to determine accuracy when used for comparing accuracy between datasets and algorithms used for text classification or sentiment analysis (Jurafsky & Martin, 2025; Terven et al., 2025).

2

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

$$\text{Precision} = TP / (TP + FP) \quad (2)$$

$$\text{Recall} = TP / (TP + FN) \quad (3)$$

$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

3. Related Work

In this section, a number of selected recent research articles related to Arabic sentiment analysis are presented. They are chosen based on their nature of using machine learning, deep learning, transformer-based models and hybrid models for Arabic sentiment classification. The selected articles in this paper represent some of the important directions of the field including dialectal sentiment analysis, noisy social media text, benchmark datasets, transformer-based models and hybrid classification approaches and are summarized in Table 1 following in chronological order.

Belbachir (2023) used MARBERT with lexical and emoji-based filtering for sentiment analysis of Algerian Arabic tweets. The study showed that targeted filtering can slightly improve the F1-score in dialectal Arabic sentiment classification. It is important because it focused on Algerian Arabic, which is one of the less-covered Arabic dialects in sentiment analysis. However, its results are limited to one dialect and may not generalize to other Arabic dialects.

Alosaimi et al. (2024) proposed a hybrid model that combines AraBERT with long short-term memory for Arabic sentiment analysis. The model was tested on Arabic benchmark datasets and achieved high accuracy. This study is relevant because it shows the value of combining transformer-based representations with sequential deep learning. However, the study was based mainly on benchmark datasets and may not fully represent noisy social media text.

Almaqtari et al. (2024) proposed a hybrid model for sentiment analysis in Arabic consumer reviews. The model used AraBERT, multi-channel CNN and bidirectional gated recurrent units. They evaluated the model on Arabic reviews and obtained good results. This work exemplifies how integrating contextual, local and sequential features can enhance sentiment classification. However, the study focused on review domains rather than tweets or social media text.

Mansour et al. (2025) proposed a transformer-based ensemble model for dialectal Arabic sentiment classification. This research compared the results of TF-IDF baselines, classical machine learning models, deep learning models and transformer based models (namely CAMELBERT, MARBERT, and XLM-RoBERTa) over Arabic tweet datasets. The ensemble was successful and produced good results on ASAD in particular. This paper is noteworthy as it directly produced models for dialectal Arabic tweet sentiment analysis, although the ensemble model is somewhat complex and much less interpretable than simpler models.

Jaber et al. (2025) proposed an ensemble machine learning approach using pre-trained embeddings and voting methods for Arabic sentiment analysis. The study showed that ensemble machine learning can still achieve strong performance when supported by effective embeddings. This indicates that strong performance in Arabic sentiment analysis does not always require a fully transformer-based pipeline. However, the approach remains more traditional compared with recent transformer-based frameworks.

In summary, the state of Arabic sentiment analysis is shifting to hybrid and transformers-based techniques, mostly for dialectal and noisy Arabic. Nevertheless, there are still existing challenges that face the available work such as domain, dialect, dataset quality, model cost, and explainability. The main objectives, methods, results, and limitations of the reviewed studies are summarized in Table 1, Comparative Study of Related Work.

Table 1. Comparative Study of Related Work

Authors	Year	Objective	Dataset / Domain	Model Technique	Key Result	Limitation
Abdul-Mageed et al.	2022	To benchmark dialect identification and dialectal sentiment analysis on Arabic tweets	NADI 2022 Arabic tweets	Shared-task benchmark systems	Winning system achieved 75.16 F1 in sentiment subtask	Benchmark paper, not a single standalone model
Abdel-Salam	2022	To identify dialect and sentiment in nuanced Arabic tweets	NADI 2022 Arabic tweets	Prompt-based, fine-tuned and multitask BERT ensemble	Achieved 1st place with 75.15 F1-PN	Designed mainly for shared-task setting
Bayrak and Issifu	2022	To improve nuanced Arabic dialect and sentiment identification through domain adaptation	NADI 2022 Arabic tweets	Domain-adapted MARBERT and Arabic BERT models	Best sentiment model achieved 74.29 Macro F1-PN	Lower than stronger recent ensemble systems
Qaddoumi	2022	To build an ensemble for Arabic tweet sentiment analysis	NADI 2022 Arabic dialect tweets	Average ensemble of Arabic BERT models	72.49 Macro-PN-F1 on dev and 69.86 F1-PN on ranked test 3rd	Simpler ensemble performed below top systems
Belbachir	2023	To improve sentiment analysis of Algerian Arabic tweets	SemEval-2023 Task 12 Algerian Arabic tweets	MARBERT with lexical and emoji-based filtering	Improved from 0.732 F1 to 0.736 F1	Focused on one dialect only
Alosaimi et al.	2024	To improve Arabic sentiment analysis using a hybrid model	Benchmark Arabic sentiment datasets	AraBERT LSTM	Reported accuracy above 97%	General benchmark setting, not tweet-specific

Baniata and Kang	2024	To address Arabic dialect sentiment analysis using advanced transformers	Arabic dialect sentiment datasets	Switch Transformer with Mixture of Experts	Achieved 84.02% accuracy in one setting	High computational cost
Almaqtari et al.	2024	To improve Arabic sentiment classification in consumer reviews	ACRD, HARD, LABR Arabic reviews	AraBERT, Multi-channel CNN, BiGRU	96.92% on ACRD, 96.68% on HARD, 92.93% on LABR	Review-based, not tweet-based
Alawneh et al.	2024	To study the effect of preprocessing and emoji encoding on Arabic tweet sentiment analysis	ASTC Arabic Twitter corpus	CNN-LSTM with preprocessing and emoji encoding	Best setup achieved 91.85% accuracy	Evaluated on one tweet dataset only
Mansour et al.	2025	To improve dialectal Arabic sentiment classification with transformer ensembles	ASTD, ASAD, TEAD Arabic tweets	Ensemble of CAMeLBERT, MARBERT, XLM-RoBERTa and others	90.4% average accuracy, 87.7% F1, 95% F1 on ASAD	Ensemble is heavier and less interpretable
Aljomah et al.	2025	To classify noisy Arabic sentiment text using transformers	MARSA noisy Arabic text	Fine-tuned CAMeLBERT vs AraBERTv02-tweet and MARBERT	92% accuracy and 91.5% F1-score	One noisy-text dataset only
Jaber et al.	2025	To improve Arabic sentiment analysis using	ArTwitter, Syria_Tweets	Pre-trained embeddings, ensemble voting	92.43% accuracy and 92.00% F1-score on ArTwitter	More classical than recent transformer frameworks

		embeddings and ensemble ML				
Alharbi et al.	2025	To evaluate sentiment analysis for Arabic dialects in hospitality domain	Multi-dialect hospitality reviews	Shared-task systems	Top system achieved 0.81 Micro-F1	Domain is reviews, not tweets
Alghamdi et al.	2026	To analyze Saudi-English code-switching sentiment and introduce a new dataset	SESA-CSD Saudi-English code-switching social media text	Small language models, best with XLM-RoBERTa	95.50% accuracy and 95.53% Macro-F1 with 3% and 11% improvement over baseline	Code-switching focus, not fully Arabic-only tweets
Alsahafi et al.	2026	To analyze sentiment of Arabic tweets during Riyadh Season	33,300 tweets from X (Twitter)	CNN, LSTM, CNN-BiLSTM, BiLSTM attention with AraBERT embeddings	Best model achieved 83.37% accuracy	Domain-specific to Riyadh Season

4. Challenges and limitations

This section discusses the main difficulties and limitations of Arabic sentiment analysis that still exist, covering: the dialect variation, noisy text in social media, domain divergence and the availability of data sets (El Mekki et al., 2021; Shi & Agrawal, 2025).

- Dialect variation-different Arabic dialects possess different vocabulary, spelling, and expressions which are challenging to sentiment classification.
- Noisy social media text - tweets and online posts have features of emoji, misspelling, sarcasm, hashtags, slang, and the informal style.
- Domain differences: a trained model on one domain (reviews) does not usually perform well on a different domain (tweets).

- Limitations of datasets -some datasets are very small and imbalanced or have specific dialects and domains.

5. Future directions

This section includes the main future directions proposed to take our knowledge further for Arabic sentiment analysis, which include creation of various datasets, enhancing robustness, minimizing computational expense and using more transparent evaluation methods (Mansour et al., 2025; Shi & Agrawal, 2025).

- Diverse datasets: Future studies should build larger datasets covering different dialects, domains, and writing styles.
- Improved models: Existing models should be further improved to effectively handle various dialects, noisy text and domain variations.
- Cost-effective hybrid techniques: Future systems should be both accurate and simple in order to be practical.
- Concise assessments: New standard datasets and evaluation metrics should be a top priority.

6. Conclusion

This survey reviewed recent methods, challenges, and future directions in Arabic sentiment analysis. It covered the techniques for Arabic sentiment analysis that are based on conventional machine learning techniques, deep learning approaches, transformer-based models, and hybrid systems. From the discussed research, it can be seen that the Arabic sentiment analysis area has achieved significant progress during recent years, particularly through adopting transformer-based models such as AraBERT, MARBERT, and CAMELBERT. It appears that the use of these models is a promising choice for analyzing sentiments of Arabic texts, especially dialectal and user-generated texts, as it captures well contextual information in text. Nonetheless, there still remain challenges such as dealing with the dialectal variations in Arabic language, noise in user-generated text, different domains of texts, scarce dataset sizes, and non-standardized evaluation metrics. The reviewed studies also show that hybrid systems are a promising direction as they combine advantages of multiple models and representation approaches. Future work is to collect bigger datasets, bigger domains, to build a more powerful model for cross dialect and cross domain analyses, lower computational requirements and use a well-defined evaluation metric. In general Arabic sentiment analysis is an area that is still developing, it should produce practical, efficient and context-based model.

References

Abdel-Salam, R. (2022). Dialect and sentiment identification in nuanced Arabic tweets using an ensemble of prompt-based, fine-tuned, and multitask BERT-based models. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP)* (pp. 452–457). Association for Computational Linguistics.

Abdul-Mageed, M., Elmadany, A., & Nagoudi, E. M. B. (2021). ARBERT & MARBERT: Deep bidirectional transformers for Arabic. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 7088–7105). Association for Computational Linguistics.

Abdul-Mageed, M., Zhang, C., Elmadany, A., Bouamor, H., & Habash, N. (2022). NADI 2022: The third nuanced Arabic dialect identification shared task. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP)* (pp. 85–97). Association for Computational Linguistics.

Alawneh, H., Hasasneh, A., & Maree, M. (2024). On the utilization of emoji encoding and data preprocessing with a combined CNN-LSTM framework for Arabic sentiment analysis. *Modelling*, 5(4), 1469–1489.

Alghamdi, H. M. (2024). Unveiling sentiments: A comprehensive analysis of Arabic Hajj-related tweets from 2017–2022 utilizing advanced AI models. *Big Data and Cognitive Computing*, 8(1), 5.

Alghamdi, S., Rayson, P., & Alotibi, R. (2026). Improving on state-of-the-art models for sentiment analysis on Saudi-English code-switching text. In *Proceedings of the 2nd Workshop on NLP for Languages Using Arabic Script*. Association for Computational Linguistics.

Alharbi, M. I., Chafik, S., Ezzini, S., Mitkov, R., Ranasinghe, T., & Hettiarachchi, H. (2025). AHaSIS: Shared task on sentiment analysis for Arabic dialects. In *Proceedings of the Shared Task on Sentiment Analysis for Arabic Dialects* (pp. 1–6). INCOMA Ltd.

Alhumoud, S. O. (2024). CNN-BiGRU-Focus: A hybrid deep learning classifier for sentiment and hate speech analysis of Ashura-Arabic content for policy makers. *International Journal of Advanced Computer Science and Applications*, 15(11).

Aljabri, M., Chrouf, S. M. B., Alzahrani, N. A., Alghamdi, L., Alfahaid, R., Alqarawi, R., Alhuthayfi, J., & Alduhailan, N. (2021). Sentiment analysis of Arabic tweets regarding distance learning in Saudi Arabia during the COVID-19 pandemic. *Sensors*, 21(16), 5431.

Aljomah, F., Aldhafeeri, L., Alfadel, M., Alshahrani, S., Abbas, Q., & Alhumoud, S. (2025). Enhancing Arabic sentiment analysis with pre-trained CAMELBERT: A case study on noisy texts. *Computers, Materials & Continua*, 84(3), 5317–5335.

Almaqtari, H., Zeng, F., & Mohammed, A. N. (2024). Enhancing Arabic sentiment analysis of consumer reviews: Machine learning and deep learning methods based on NLP. *Algorithms*, 17(11), 495.

Alosaimi, W., Saleh, H., Hamzah, A. A., El-Rashidy, N., Alharb, A., Elaraby, A., & Mostafa, S. (2024). ArabBert-LSTM: Improving Arabic sentiment analysis based on transformer model and long short-term memory. *Frontiers in Artificial Intelligence*, 7, 1408845.

Alsahafi, S., Alshutayri, A., & Alahdal, S. (2026). Sentiment analysis of tweets during the Riyadh Season using deep learning. *Engineering, Technology & Applied Science Research*, 16(1), 32327–32335.

Alyami, S. N., & Olatunji, S. O. (2020). Application of support vector machine for Arabic sentiment classification using Twitter-based dataset. *Journal of Information & Knowledge Management*, 19(1), 2040018.

Antoun, W., Baly, F., & Hajj, H. (2020). AraBERT: Transformer-based model for Arabic language understanding. In *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools* (pp. 9–15). European Language Resources Association.

Baniata, L. H., & Kang, S. (2024). Switch-transformer sentiment analysis model for Arabic dialects that utilizes a mixture of experts mechanism. *Mathematics*, 12(2), 242.

Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future Generation Computer Systems*, 115, 279–294.

Bayrak, G., & Issifu, A. M. (2022). Domain-adapted BERT-based models for nuanced Arabic dialect identification and tweet sentiment analysis. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP)* (pp. 425–430). Association for Computational Linguistics.

Belbachir, F. (2023). Foul at SemEval-2023 Task 12: MARBERT language model and lexical filtering for sentiment analysis of tweets in Algerian Arabic. In *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)* (pp. 363–369). Association for Computational Linguistics.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long and Short Papers)* (pp. 4171–4186). Association for Computational Linguistics.

El Mekki, A., El Mahdaouy, A., Berrada, I., & Khoumsi, A. (2021). Domain adaptation for Arabic cross-domain and cross-dialect sentiment analysis from contextualized word embedding. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 2824–2837). Association for Computational Linguistics.

Elhassan, N., Varone, G., Ahmed, R., Gogate, M., Dashtipour, K., Almoamari, H., El-Affendi, M. A., Al-Tamimi, B. N., Albalwy, F., & Hussain, A. (2023). Arabic sentiment analysis based on word embeddings and deep learning. *Computers*, 12(6), 126.

Inoue, G., Alhafni, B., Baimukan, N., Bouamor, H., & Habash, N. (2021). The interplay of variant, size, and task type in Arabic pre-trained language models. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop* (pp. 92–104). Association for Computational Linguistics.

Jaber, A., Bahati, I., & Martínez, P. (2025). Leveraging pre-trained embeddings in an ensemble machine learning approach for Arabic sentiment analysis. *Frontiers in Artificial Intelligence*, 8, 1653728.

Jurafsky, D., & Martin, J. H. (2025). *Speech and language processing* (3rd ed., draft). Stanford University and University of Colorado Boulder.

Mansour, O., Aboelela, E. M., Talaat, R., & Bustami, M. (2025). Transformer-based ensemble model for dialectal Arabic sentiment classification. *PeerJ Computer Science*, 11, e2644.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep learning-based text classification: A comprehensive review. *ACM Computing Surveys*, 54(3), Article 62.

Qaddoumi, A. (2022). Arabic sentiment analysis by pretrained ensemble. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP)* (pp. 447–451). Association for Computational Linguistics.

Shi, Z., & Agrawal, R. (2025). A comprehensive survey of contemporary Arabic sentiment analysis: Methods, challenges, and future directions. In *Findings of the Association for Computational Linguistics: NAACL 2025* (pp. 3760–3772). Association for Computational Linguistics.

Taher, H. A. (2024). Dual-language sentiment analysis: A comprehensive evaluating SVM, logistic regression, XGBoost, and decision tree using TF-IDF on Arabic and English dataset. *Wasit Journal for Pure Sciences*, 3(4), 59–69.

Terven, J., Cordova-Esparza, D.-M., Romero-González, J.-A., Ramírez-Pedraza, A., & Chávez-Urbiola, E. A. (2025). A comprehensive survey of loss functions and metrics in deep learning. *Artificial Intelligence Review*, 58, 195.