مجلة جامعة بني وليد للعلوم الإنسانية والتطبيقية
Bani Waleed University Journal of Humanities and Applied Sciences
تصدر عن– جامعة بني وليد – ليبيا
Website:https://jhas-bwu.com/index.php/bwjhas/index

ISSN 3005-3900 | (348 -331) الصفحات | المجلد الحادي عشر – العدد الثالث – 2026

A Arabic Language: Sentiment Analysis (Opinion Mining) using Sentic Computing Models, Tools and Techniques

Amer Saleh Musbah Gadowa*
Department of Computer Programming, College of Computer Technology Tripoli, Tripoli. Libya
amer_s_gadowa@cctt.edu.ly

اللغة العربية: تحليل المشاعر (استخراج الآراء) باستخدام نماذج حوسبة السنتك ، الأدوات و الأساليب

عامر صالح مصباح قدوعة [1]*
قسم برمجيات الحاسوب، كلية تقنية الحاسوب طرابلس، طرابلس، ليبيا

Received: 05-04-2026	Accepted: 12-05-2026	Published: 20-05-2026

الملخص:

تعتبر اللغة العربية لغة غنية جداً بمفرداتها، ومن هذا المنطلق يعتبر علم تحليل العواطف واستخراج المشاعر من نصوصها أمراً يحمل تحدياً كبيراً في مجال حوسبة السلوك العاطفية. ونظراً لهذه الخصائص اللغوية والدلالات الفريدة تتميز اللغة العربية بتراكيب معقدة من ناحية النحو والصرف وكذلك تعدد معاني المفردات، إضافةً للتنوع والغنى في الأساليب البلاغية و التعبيرات المجازية. علاوةً على ذلك مع تطور اللغة وتعدد لهجاتها أصبحت لكل رقعة جغرافية في الحاضرة العربية لهجة تخصها. مع توافر هذه الخصائص لهذه اللغة جعلت تحليل و تقييم العواطف عملية دقيقة تتطلب استخدام أدوات حديثة لمعالجة اللغة الطبيعية مثل Farasa Tools، بالإضافة الى معاجم المشاعر السلوكية للغة العربية، ونماذج التعليم العميق مثل Word2Vec ، BERT و AraBERT . تستخدم هذه الأدوات لتصنيف النصوص على مستويات عدة (النص الكامل، الجملة والسياق)، أيضاً تحديد المشاعر الإيجابية، السلبية والمحايدة، وعلاوة على التعامل مع الاختصاصات و الرموز التعبيرية الشائعة في وسائل التواصل الاجتماعي المتنوعة.

من هذا المنطلق تطرح هذه الدراسة التساؤل التالي: إلى أي مدى تساهم هذه الخصائص اللغوية والدلالية التي تميز اللغة العربية في التأثير على تحليل واستخراج الآراء و اكتشافها ؟ وما هي أكثر نماذج حاسوبية وأدوات تقنية استخداما في تحاليل مقاصد الأنماط السلوكية؟

تتمثل فرضيات هذه الدراسة في دقة تحليل المشاعر واستخلاص الآراء والتي ترتبط ارتباطا إيجابياً باستعمال نماذج متقدمة للحوسبة الوجدانية ومثالاً على ذلك نماذج التعليم العميق. و بحسب السياق ففي النصوص العربية تحقق الأنظمة الخاصة بتحليل المشاعر و التي تعتمد على معاجم و قواعد لغوية رصينة ذات نتائج

أكثر دقة مقارنة بتلك التي تعتمد على الأساليب الإحصائية. مع الإشارة بأن البيئة التي أُستخدمت في جميع العمليات البرمجية تم تنفيذها بلغة البرمجة بايثون.
أظهرت هذه الدراسة ان دمج التحليل البلاغي والسياقي مع تقنيات التعليم العميق يساهم بشكل جوهري في تطوير و تحسين دقة استخراج المشاعر و الآراء من النصوص العربية، خاصةً النصوص القصيرة مثل المنشورات والتغريدات.
نتائج هذه الدراسة تبرز و تسلط الضوء على إنشاء قواعد بيانات متكاملة تشمل اللغة العربية الفصحى وايضاً مختلف اللهجات، وتطبيق تحليل الآراء و المشاعر على مستويات متعددة. كذلك سيساهم ذلك في توظيف هذه التقنيات في مجالات متنوعة على سبيل المثال لا الحصر مجال التسوق، الرصد الاجتماعي، التعليم الرقمي و تحسين تجربة آراء المستخدم. من هذه الدراسة وعلى رغم التقدم باستخدام أبجديات الأدوات و النماذج التقنية المتوفرة لا تزال هناك تحديات قائمة تتعلق بغزارة و تنوع اللهجات، اللغة المجازية والبيانات الغير مهيلكة، جميع هذه المعوقات تفتح الأفق أمام تطوير نماذج أكثر دقة وفعالية في مجال الحوسبة السلوكية للغة العربية.

الكلمات الدالة: اللغة العربية، استخلاص الآراء، التعلم الألي، تحليل الأنماط السلوكية، التعليم العميق ، الحوسبة المتعلقة باكتشاف المشاعر.

Abstract

Emotion analysis and sentiment extraction from Arabic texts present a significant challenge in the field of affective computing, given the distinctive linguistic and semantic features of the Arabic language. Arabic is characterized by complex morphological and syntactic structures, multiple meanings of words, and a rich use of rhetoric and figurative language, in addition to the existence of diverse local dialects alongside Modern Standard Arabic. These characteristics make emotion assessment a delicate process requiring sophisticated natural language processing tools, such as Farasa and CAMEL Tools, Arabic emotion dictionaries, and deep learning models like Word2Vec, BERT, and AraBERT. These tools are used to classify texts at multiple levels (full text, sentence, and context), identify positive, negative, and neutral emotions, and handle abbreviations and emojis commonly used in social media, the implementation has been carried out using python programming language environment.
The study hypotheses were the accuracy of sentiment analysis and opinion extraction is positively correlated with the use of sophisticated sentiment computing models (such as deep learning). Furthermore, in Arabic texts, sentiment analysis systems that utilize specialized dictionaries and grammars yield more accurate results than those relying solely on statistical methods; hence the question to what extent the linguistic, semantic, and contextual features characterize the Arabic language and influence emotion analysis and opinion extraction?
What are the most common affective computing models, tools, and techniques used in emotion analysis? This
study have shown that integrating semantic, contextual, and rhetorical analysis with deep learning techniques enhances the accuracy of emotion extraction from Arabic texts, especially short texts such as tweets and posts.
The results moreover highlight the highly significant of establishing comprehensive databases that include Modern Standard Arabic and various dialects, and applying emotion analysis at multiple levels. This facilitates the use of these technologies in diverse fields such as marketing, social monitoring, digital education, and improving user experience. Despite the remarkable progress in tools and models, obstacles remain related to dialectal diversity, figurative language, and unstructured data, which open the door to developing more accurate and effective models for emotional computing in Arabic.

Keywords: sentiments Analysis, Opinion Extraction, Arabic language, Affective Computing, Machine Learning, Deep Learning.

Introduction

Sentiment analysis in context is defined as the investigation of individual's thoughts, feelings, evaluations, viewpoints, attitudes, and emotions toward things, embracing goods, services, organizations, people, problems, events, topics, furthermore their characteristics, are all subject to sentiment analysis. Five elements are used to determine opinion: the opinion holder (P), the individual or entity expressing the opinion, and the time (T), when the viewpoint is expressed. Sentence analysis is a highly challenging field of study with numerous and diverse objectives. Among its most prominent activities are subjective classification, emotion classification, vocabulary development, aspect extraction, classification of emotions associated with those aspects, and identification of disturbing messages in opinions [1]. Hence, the main question is given the linguistic, semantic, and contextual characteristics of Arabic writings, how effective are affective computing models, tools, and procedures in analysing emotions and extracting opinions?

Therefore, this study includes sections on background information about sentiment analysis, a description of the difficulties of impression analysis in Arabic, a summary of the most important studies conducted in this field, a description of the linguistic characteristics of Modern Standard Arabic, a discussion of the results, and a conclusion. The study addresses the clear shortcomings of emotion analysis models and methods when applied to the Arabic language compared to other languages, particularly English. The Arabic language is characterized by its multiple levels (Classical Arabic, colloquial Arabic, and regional dialects), the complexity of its morphological and grammatical structure, and the depth of its rhetorical and contextual meanings. For this reason, traditional computer models face difficulty in accurately understanding emotional meanings, and this task becomes more difficult when trying to extract implicit opinions and feelings, that is, those are not explicitly stated through emotional vocabulary, but rather stem from the context, style, or linguistic structures such as negation, irony, and metaphor. Many of the available tools that rely on statistical properties or limited dictionaries fail to produce accurate and reliable results when analysing Arabic texts, and this may lead to a misunderstanding of the emotional positions underlying Arabic discourse or to a distorted perception of public opinion trends.

1 Background of Research

1.1 Arabic language

Social media is spread everywhere in Arabic countries and Arabic sentiment analysis is needed because of a large scale audience. The history of Arabic language is challenging and interesting, the strategic importance of its people, the regions they occupy, and its culture and literary heritage. Arabic is the formal language of 22 nations [2] and it is extremely highly of 420 million people as a mother language, and by 250 million as a second language; it's ranked the 5th language in the world as Arabic-speaking users online about 65 million and that equivalent to 18.8% of all over the world cyberspace community.

Arabic is categorized within three sorts: *Classical Arabic*, *Modern Arabic*, and *Colloquial Arabic*. First, during and before Islam era the Classical Arabic is used and the Holy Quran was subsequently written on it, it's a prosperous of lexicon and sophisticated grammar. Modern Arabic is emanating from the classical form and today it became the formal language of education, media, besides literature. Colloquial Arabic is the language of life, that be used between people in everyday communication and it's slightly different from country to another due to each country has its own dialect [2]. Arabic language has 28 letters and orientation of writing is from right to left; it has no upper or lower case like in English. Furthermore, it has extreme complicated composition; the predominance of expression has a tri-letter root, and what remains either a quad-letter root, penta-letter root or hexa-letter root. The vowels letters in Arabic are (ى "Yaa" ، و "Waw" ، أ "Alef") and the remaining letters are constants [3].

In Arabic countries social networks have changed the way of people thinking, especially in youth generation. This influence has been appeared in the political unrest of the Middle East and North Africa; consequently, sentiment analysis and opinions mining are an indispensible basically for Arabic slang language which is a common that people used via global social networks.

Social networks sites, blogs are provide a mechanism for interacted people to leave a comment or a message which are perform a crucial capacity in the communication in between source of events and destination of online viewers [2]. Participants' opinions and comment is a major principal for the development of the services quality and improvement facilities. The reviews for services/ products are usually based on opinion expressed in much unstructured format. Sentiment classification studies are used to gather reviewers' customers' data for some purposes.

1.2 The Problem of Sentiment Analysis in Arabic Language

Formal Arabic Language has tools exiles which are "ليست، لم، ليس، ما، لن، لا" and those in English are "laysat, lm, laysa, ma, ln, la", mean "No, Not" and negative suffixes and prefixes, which are exiles in English; there is another exile tool "ش" that can be added as a suffix of verbs like "يحكم" which is "Govern" and to get the opposite meaning which is “يحكمش”, where it’s “لا يحكم” informal Arabic and this means in English “Not Govern”. In addition, usually, we use this exile tool in another way by adding character “م ” at the incipient of the verb (replacing the formal exile too “ما”), and then the previous example will be “ميحكمش ” which gives the same meaning [2]. To make it clear for the previous exile tool, this example (“بايدن ميحكمش امريكا بعد 2025”, “Biden mayahkomesh America paad 2025”) which means “Biden is not govern the USA after 2025” since the word “ميحكمش” comprises the prefix “م” and suffix “ش”; that leads to another exile tool a suffix “مش” which means “لا” as formal Arabic and means “No” in English; the previous example after this derivation will be (“بايدن مش يحكم امريكا بعد 2025”, “Biden mesh yahkom America paad 2025”). However, in my country Libya has a slight difference in colloquial language than the other Arabic countries. The opinion mining has the capability for analysing people opinion, evaluation, sentiment, appraisal, behaviour and emotion towards entities such as the product or services. Sentiment analysis is mostly focused on negative or positive sentiments. For instances, the comment “ نوعية الثلاجة ما تنفعش” which is “nowayet altalaja ma tnfash” this Arabic (Libyan) slang language which means in English “The mark of fridge is bad”, in this sentence the tool exile “ما” has a negative opinion. Sometimes, the opinion has two different meaning like in this statement “هذه السيارة صقع” which is “Hadehe alsyara sagga” and in English “This car is cool” this statement gives both positive and negative opinion, the positive one means the car unbelievable, pretty, and attractive modern car, where negative opinion means the car is very expensive and it isn’t affordable to buy it. Users’ comments are sometimes written sentimentally and subjectively in Arabic with English characters and numbers which is called FrankoArabic language in free text as it is unconventional and arbitrary; furthermore, the annotations perpetually contain several syntax errors which make the miming process to be hard [2]. For example of these errors, the user comment likely restate a character in a term like “هدددددددف” rather than “هدف” which represent “Goal”, also the word “مبرووووك ” instead of “مبروك ” and means “Congratulation” in English, all of these need concentration to improve the result of colloquial language mining.

1.3 Feature Selection in Sentiment Classification

Temperament Analysis is the computational study of people’s opinions, attitudes and emotions toward an entity [4], [5]. Sentiment analysis activity is perceived a sentiment classification issue as the first primary in the sentiment categorizing challenge is extracting and selecting content attribute. There is a portion of these features, for instances, *terms presence* and *frequency*: these are individual words n-gram and their frequency counts, and it has two possibilities either provides the words binary weighting or it utilizes term periodical weights to point out the associate significance of features; *parts of speech*: discovering and detecting adjectives; o*pinion words and phrases*: these are words usually used to express opinions, negations. Sometimes the opinions orientation changed by characteristic of negative words. Figure1 illustrates Sentiment Analysis process and steps on product review.

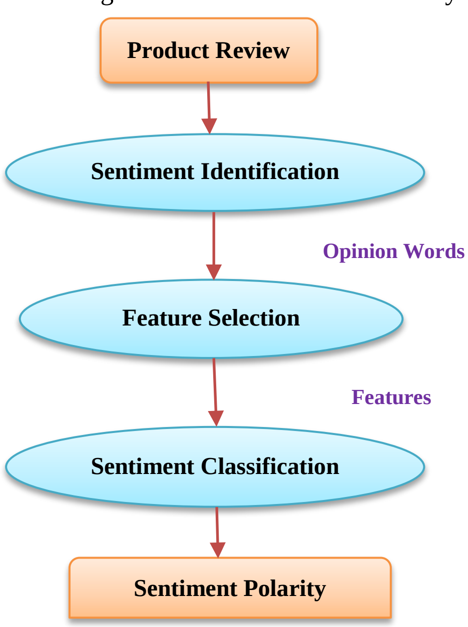

Figure 1. Emotion Analysis procedure on product review [6]

Feature of picking strategies can be categorized into two methods: Lexicon-based methods that need user annotation and statistical methods which are automatic ones and are more habitually applied to obtain a large lexicon they bootstrap this assortment across synonym discovering or networked resources. The feature selection approaches manage the documents in two ways either as collection of words (Bag of Words) or as a string that retains the sequence of vocabulary in the documents; the simplicity of bag of words for its classification process make it more often to use [6].

The next diagram (Figure2) summarizes and shows some of the sentiment classification techniques.

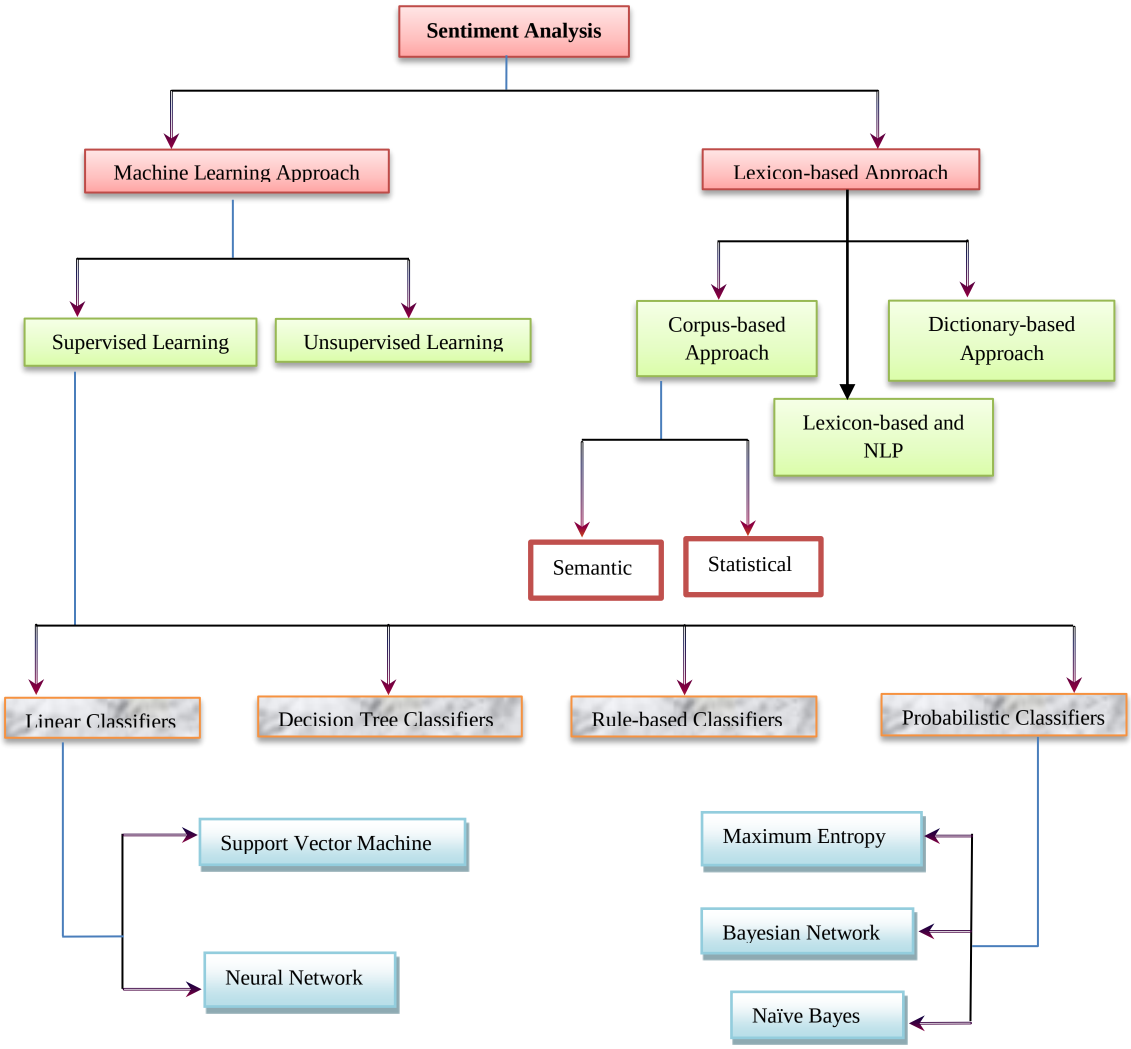

Figure 2. Sentiment Classification Techniques [4]

Study problem

This study addresses the clear shortcomings of sentiment analysis models and methods when applied to the Arabic language compared to other languages, particularly English. The Arabic language is characterized by its multiple levels (*Classical Arabic, colloquial Arabic, and regional dialects*), the complexity of its morphological and grammatical structure, and the depth of its rhetorical and contextual meanings. For this reason, traditional computer models face difficulty in accurately understanding emotional meanings, and this task becomes more difficult when trying to extract implicit opinions and feelings, that is, those that are not explicitly stated through emotional vocabulary, but rather stem from the context, style, or linguistic structures such as negation, irony, and metaphor. Many of the available tools that rely on statistical properties or limited dictionaries fail to produce accurate and reliable results when analysing Arabic texts, and this may lead to a misunderstanding of the emotional positions underlying Arabic discourse or to a distorted perception of public opinion trends.

The main question of the study: Given the linguistic, semantic, and contextual characteristics of Arabic writings, how effective are affective computing models, tools, and procedures in analysing emotions and extracting opinions?

2. Related Studies

Sentiment analysis, also famous as opinion analysis, is a technique for determining and extracting sentimental information throughout written texts using machine learning and natural language processing. Because it allows us to acquire comprehensive sympathetic of public attitudes and opinions toward specific topics, goods, or services, sentiment analysis is a crucial tool today; sentiment analysis has emerged as a major research area due to the multiplication of reviews, evaluations, endorsements, as well as other forms of expression facilitated by the interactive internet. Its applications are evident in diverse fields, including education, health, politics, commerce, and tourism. Many studies have been dedicated to sentiment analysis due to its importance; however, few of these investigations have pointed out sentiment analysis in morphologically abundant languages such as Arabic. Over the past decade, sentiment analysis in Arabic has intrigued significant concentration from specialists and scientists due to the intensifying value of Arab internet users and the rapid expansion of Arabic materials available online [7].

The study of [8] explained sentiment analysis, or opinion mining, aims to gain increasing attention in both industry and academia. Despite significant developments in building sentiment analysis models, this domain persist lively research focus for varied languages around the world, especially Arabic, which is the fifth most widely spoken and the fourth utmost language globally on the Internet? Accompany by fact that researches activities have been heightening in opinion mining in Arabic; numerous researchers have conducted reviews to keep track of developments in this field. While these studies highlight a treasure trove of substantial advances, the rapid stride of developments in machine learning and NLP involves an unceasing require to update the literature. The mission of this paper aims to yield an exhaustive overview of the contemporary developments in Arabic opinion mining. This review shifts beyond simply examining previous work that focused principally on classification models; it offers a holistic perspective on the system, covering advancements in various aspects of opinion mining. These include developments in NLP tools, text sources and collections for sentiment analysis, classification models, and opinion extraction applications. The review also explores future trends in Arabic opinion mining, incorporating the latest developments in the field, including deep learning applications for Arabic opinion mining. The article set off up-to-date information to assist novice and experienced researchers, in addition to industry developers seeking of deploying a comprehensive and effective opinion mining system. Each section concludes with key insights into specific characteristic of the opinion mining system, allowing the reader to focus on the areas of interest [8].

The changeable terrain of political in numerous Arab countries prompted this contemporary research [9], driven by the continuous proliferation of online journalistic content. Hence, the aim is to initiate a media scrutinizing system for analysing political opinions. This mechanism permits political performers; nevertheless, the vast amount of online materials, to remain persistently enlightened about opinions demonstrated online. This enables them to accurately inspect their actual situation, guide the communication strategies, and set up the election campaigns. The evolved system relies on a linguistic technique that utilizes the NooJ language engine to standardize automatic identification principle and employ them to a vigorous set of journalistic articles. The applied rules preliminary granted for the identification besides characterization of various political individual (performers and political institutions). These characterizations are then applied in our media monitoring system to determine the standpoint correlated with the identified entities. The system relies primarily on a set of locally developed grammatical rules for identifying the structures of various political opinion statements. These rules utilize opinion-specific lexicon inputs containing different expressive words (verbs, adjectives, nouns), with each input linked to the analogous semantic designator (polarity and intensity). The upgraded system is capable of identifying and characterizing the sentiment possessor, the sentiment's objective, further to the polarity (nominal or verbal) of the idiomatic feeling. The recent experiments have shown that the adopted extraction method is consistent with an F-measure of 0.83.

This following paper [10] presented the use of data extracted and statistical methods in preparing coherent literature reviews (SLRs). Sentiment analysis (SA) and opinion mining (OM) in the field of social media had been deliberated as a case study to generate a model for a systematic review. The prepared systematic review reviewed the discipline of sentiment analysis and opinion mining and discussed the latest dilemmas inside mining user material in the field of social media. The chief objective of the research deed illustrated in this systematic review was to scrutinize applying the techniques combined of both sentiment analysis and opinion mining for investigating social media; where the current era cyber phenomena (Facebook, Twitter, Google Plus, and LinkedIn) has been dominant and impact our life in such diversity. Accordingly, each person has the capability to obtain these networks to publish their daily news or to articulate their individual opinions or emotional perspective towards different topics. Thus, a massive magnitude of data is broadcasted typically by users in worldwide scale. However, manually processing and analysing this vast volume of data can be impractical. Consequently, scientists have managed numerous studies to reveal new mechanism, tools for analysing and exploring this enormous dataset, particularly data generated through social media. This paper draws on systematic literature reviews and evaluates numerous pertinent investigation papers addressing particular research inquiries. In reaction to the insufficient of recent methodical reappraisal with regarded to sentiment analysis in digital platforms, the pioneers expose this comprehensive and structured review. It defines the field of emotion analysis and opinion mining techniques utilized to data derived from internet communities and explore the demanding of attitude mining via social media framework. The paper commences with details the various methods for performing opinion extraction [12].

Sentiment analysis is a vital and modern field of research in NLP, promoted by means of the ever-increasing magnitude of online impression data. Substantially all tactical strategies in this vicinity are based on English owing to the scarcity of temperament reservoir for other languages, in particular Arabic, and its wide assortment of vernaculars. At the peak of applications of emotion analysis, great sentiment sources play a pivotal purpose. This article reviews a handful of explicitly obtainable sentiment analysis resources for Arabic. It presents the Arabic Sentimental Lexicon, an index of 3,880 positive and negative synonym sets illustrated by their grammatical classification, degrees of polarity, dialectal synonym sets, and conjugation patterns. The article also presents a Multi-Domain Arabic Sentimental Set (MASC) with 8,860 positive and negative reviews from various domains. An in-depth study of five feature set patterns was conducted to leverage productive features and investigate the impact on the accomplishment of Arabic attitude analysis. The intention is of evaluating the quality of the enhanced linguistic materials. The Arabic Sentimental Lexicon is accustomed to generate feature vectors. Five common machine learning algorithms were employed: Naive Bayes, Nearest Neighbours (kNN), Support Vector Machines (SVMs), logistic linear regression, and neural networks as primary classifiers for each set of attributes. An extensive of comparative investigation was executed on standard Arabic datasets, the discussion had been presented, and inferences were drawn. The empirical achievements indicate that the Arabic sentimental lexicon is a valuable resource for analysing Arabic sentiments. It was shown that classifiers pre-processed on attribute vectors extracted from the set employed the Arabic sentimental lexicon are more accurate than classifiers trained based on the initial set [11].

Opinion analysis is a branch of NLP, information fetched, and words mining. It is the procedure of extracting individuals' point of view and awareness from disorganized words contents. This has turn to be an effective, attractive, and furthermore stimulating area on account of emergence of networked social channels and the enormous magnitude of user annotations. There is a variety of research including alternative directions and methodologies within this field; however the dearth of an in-depth initiative that examines it via all aspects is palpable. The paper, presents an entire, multifaceted, similarly methodical review of impression analysis and sentiment analysis to designate the obtainable approaches and distinguish their pros and cons. This aims to gain a superior cognition of the setbacks and available resolution for illuminating the upcoming direction. Accordingly, the study presents a suitable framework for opinion analysis, along with all procedure. It had proven the observe, classify, summarize, and fully contrast the conceived techniques for extracting aspects, classifying opinions, producing summaries, and evaluating them, based on key scholarly works. Towards accomplish a perfect comparison; it had suggested some components in each and every individual that could assist of achieving more appropriate interpretation of the advantages and disadvantages of dissimilar methods [12].

Sentiment analysis of textual documents [11] is an active research area, and online community provides desirable intelligence sources in diverse languages, including all sort of interaction throughout the media. These resources can be studied to look into individual's emotion, behaviours, opinion, and feelings toward different topics and products. Therefore, the paper aimed to analyse customer sentiment on Arabic social media, focusing on the real estate as well as automotive sectors. From this perspective, mechanized analyser systems are conceived of categorizing the sentiment polarity of each customer reconsider or comment on social media as "positive," "negative," or "mixed." This is attained by collecting data from web-based subscriber who interacted with reviews about real estate or cars. These connected convergences are among the vast social media platforms for customers discussing opinions about real estate and cars, specifically in the dialects of the Gulf Cooperation Council (GCC) countries, furthermore in Arabic in commonly. Datasets are available in combined of Gulf dialects and Modern Standard Arabic (MSA) [13].

This research [14] presents an innovative impression investigation methodology precisely individualized for the Arabic language, a field that presents unparalleled impediment on account of its sophisticated morphological structure and heterogeneous colloquialisms. Using a tremendous dataset of over 108,000 overviews from Arab institutions, the principal objective was to establish a vigorous and authentic sentiment quantifying scheme that handles the complexities of the Arabic language and aims to help organizations more effectively understand customer sentiment. Their framework included a comprehensive pre-processing period, essential of setting up the data for rigorous evaluation. Here the cycle involved reforming emojis and smileys into textual depiction, an evolutionary as well as passionate feature of computerized interaction, displacing Arabic diacritics to standardize the text input, and adjusting the words for constancy. The pre-processing was scrupulously devised to signify the precise difficulties prompted by Arabic manuscript and colloquial deviations. For the nucleus mood detection, it has appointed two powerful machine learning paradigms: logistic regression and Naïve Bayes. Every model was carefully selected and reformed for its authentic effectiveness in classifying texts, particularly in manipulating the nuances of opinion mining throughout Arabic text. This endeavour is not limited to exploiting the advantages of these two models, but extends for adapting them to consider the semantic aspects of Arabic. The achievement of the study supplied substantial participation of the scope of impression analysis in Arabic by addressing the distinctive barriers of the language and accommodating traditional machine learning techniques. Consequently the study supplies precious insights and techniques for companies and investigators concerned with the Arab commerce. These clarifications are essential for firms hunting, for awareness and retort to subscriber sentiment in a morphological miscellaneous furthermore culturally affluent territory [14].

Internet interaction platforms are crucial for disseminating data, news, and opinions. Social media people has the capabilities of engaging and had their views on various duties and merchandise [15]. Sentiment analysis is one approach to scrutinizing user opinions for mining effective information. The purpose of the research was evaluation of sentiment analysis model. The framework primarily consists of four cycles: data assembly, pre-processing, sentiment evaluation, categorizing plus evaluation. Data retrieving involves gathering Arabic documents or feedback from digital communication platforms such as Twitter. Antecedent treatment includes text segmentation, punctuation exclusion, standardization, and word stemming. Sentiment evaluation and classification address several key topics, including negation verification, amplification processing, emotion identification, and emotion classification. The appraisal phase assesses the performance of the sentiment analysis model. The viewpoint analysis tool is endorsed by a range of Arabic glossary compilations, specifically the diversity of Arabic punctuation words, positive and negative emotions, affirmative and undesirable modifiers, moreover suffixes for the slight stemmer. The impression scrutiny tool promotes categorize the interactive environment for participant critiques into positive, negative, or neutral (emotional polarization). The approved attitude analysis technique is used for the identity emotions in both Modern Standard Arabic, and permits the probe also the discovering in Informal Arabic (colloquial Arabic), to what media place spread and operate. Tangible requirements such as accuracy, recollect, validity, and flaw ratio are used to estimate the functioning of the sentiment analysis pattern. Numerous observations are conducted using triple main concepts of Arabic words: negation, affection, and amplifiers. The prototype's attitude changes and is influenced by the use of these themes, provided that individually or in combination. The archetype's execution is likewise impacted by the category of Arabic sentence and the Arabic linguistic approach. The sentiment analysis model demonstrates good performance and provides satisfactory validity worth. The validity values for predicted affirmative opinions are 98.2%, 91.8%, and 85.8%, whereas the values for unpleasant reviews are 93.2%, 92.6%, and 70.1%, respectively, for Modern Standard Arabic, Mixed Arabic, and Informal Arabic styles [15].

2.1 Commentary on Previous Studies

A review of previous studies reveals a growing interest in sentiment analysis and emotion retrieving, particularly within the context of the Arabic language, given its linguistic and structural challenges compared to other languages. Some studies, such as Badaro & Bali (2019), have focused on providing a comprehensive and integrated view of Arabic sentiment analysis systems, addressing tools, resources, models, and applications. These studies also emphasize the rapid development of deep learning techniques, reflecting a relative maturity in the field, though the need to update theoretical and technical frameworks remains On the other hand, studies like Najjar & Musfir (2017) have explored the practical applications of sentiment analysis in the political sphere, relying on linguistic and grammatical approaches. While these studies have shown good accuracy, this approach remains limited in its flexibility and generalizability compared to modern machine learning models. Some studies have also focused on the descriptive and methodological aspects, such as those by Salah & Al-Ghuwairi (2019) and Hamtian & Sohrabi (2019), which sought to classify different approaches and techniques in sentiment analysis and opinion mining, highlighting the challenges associated with social media data. This underscores the need for models more adapted to the nature of short, informal texts. Regarding linguistic resources, Al-Musallami & Al-Bard (2018) emphasized the crucial importance of Arabic sentimental dictionaries and demonstrated that employing dedicated linguistic resources contributes to improving the accuracy of models. However, the limitations of these resources in terms of size and dialectal coverage remain an obstacle to achieving more comprehensive results.

Some studies have also addressed sectoral applications of sentiment analysis, such as Alsemaree et al (2024) study in the real estate and automotive sectors, and Fouda & Ahmed's (2024) study on the evaluation of Arab companies. This reflects an expansion in the use of sentiment analysis to support decision-making, although reliance on traditional or semi-traditional models remains prevalent compared to modern, in-depth models. Ruby & Badawi's (2018) study evaluated the performance of sentiment analysis models in Arabic social media comments, focusing on Modern Standard Arabic and colloquial Arabic. The results showed good accuracy, but the influence of linguistic style and sentence structure remains a significant factor requiring further investigation.

In general, it appears that most previous studies have focused on models, resources, or applications without comprehensively integrating sentiment computing models, tools, and techniques within a unified framework. Furthermore, there is a scarcity of studies that address the Arabic language at its various levels (Modern Standard Arabic and dialects) using advanced, modern models within an integrated system, which is what this study aims to address.

What distinguishes previous studies?

Most previous research on sentiment analysis and opinion inference in Arabic has focused on reviewing literature or building linguistic resources, such as sentiment dictionaries and text collections, or on employing traditional classification methods in text processing, such as classical machine learning (e.g., vector support machines, Naïve Bayes, and logistic regression). Many of these studies have also been limited to specific applications in particular fields, such as political affairs or product evaluation via social media, and have not addressed the integration of conceptual and cognitive aspects of sentiment within the analysis. These works have been confined to standard text processing, without considering the advanced capabilities of cognitive processing models that allow for a deeper understanding of the emotional and semantic context of Arabic texts and their diverse dialects.

This research is distinguished by its reliance on an emotional computing model that integrates the emotional, cognitive, and conceptual representation of texts, enabling a deeper and more comprehensive understanding of emotions. Your research goes beyond limited applications or superficial analysis of texts to include the study of different Arabic dialects and the contextual diversity of content, using advanced tools and techniques to process emotions within the contextual framework and implicit meanings. This approach provides a more accurate and complete understanding of emotions and enables the provision of integrated analytical solutions that surpass traditional models, making your study unique in combining applied and theoretical aspects, linguistic and mental processing, and text analysis at multiple levels.

3. Theoretical Framework

3.1 The Theoretical Framework of Emotional Computing and Emotion Analysis

Developing intelligent systems those have the capability of understanding, and interpreting the emotional nuances inherent in textual data in a way that mimics human perception has become crucial in the digital age, characterized by the ever increasing volume of textual data circulating online. Emotional computing is one of the most prominent areas of artificial intelligence, aiming to enable computers to see, understand, and interact with human emotions in ways that go beyond simple textual comprehension. This approach has enabled the transformation of emotional input into measurable and interpretable information, leading to a better understanding of public opinion trends, improved user experience, and enhanced decision-making across various disciplines [16].

Polarity scrutiny, again renowned as opinion mining, is one of the key uses of sentiment computing. It uses written texts, such as news, social media posts, and product reviews, to identify and assess users' psychological orientations and emotional attitudes. This field uses a variety of machine learning and NLP models furthermore the techniques that enable the categorization of emotions into broader categories that include multiple emotions, or into positive, negative, or neutral polarities. As AI tools continue to evolve, the scope of sentiment analysis has expanded to include contextual understanding, the identification of complex emotions, and the extraction of implicit sentiments. This represents a shift from traditional text processing to models that are better able to simulate human understanding [17]. For instances of what mentioned above, automated lexical resources in Arabic language sentence subjectivity, lexical resources for sentiment words is a collection of words those are express the subjectivity in text. There are two types of opinion words, positive words which used to show the felling or opinion actions for example ("أخاذ" "magnificent", "حب" "love", "ممتاز" "distinction"), and negative opinion words that expressing the undesirable situation like ("شنيع" "awful", "كره" "hate", "أناني" "selfish") [18]. The developing sentic computing based-on sentiment mining techniques for extracting sentiments from Arabic natural language text is still a broad area of computational linguistics, text mining, NLP and machine learning, which needs more efforts and stress to be flourish field.

3.2 The Concept and Importance of Emotional Computing and Emotion Analysis

Emotional process data is designated as the analysing and progression of tools and systems capable of identifying, interpreting, processing, and simulating emotional phenomena. While some developments in this field can be pursued to premature philosophical studies of emotion, Picard's seminal contributions led to the emergence of a further contemporary branch of computer science, which has evolved into a multidisciplinary field encompassing computer science, psychology, and cognitive science. The ability to equip robots with emotional intelligence, enabling them to recognize human emotions and adjust their behaviour accordingly, is one of the driving forces behind this field [19].

Artificial intelligence enhances user-system interaction by making these systems emotionally sensitive as well as intelligent. Although the way the human brain works differs fundamentally, research in emotional computing is hampered by basic hardware-related problems. The physical structure of the human brain, often divided into rational and emotional components, allows for the simultaneous processing of information both emotionally and rationally. However, this is not the case in the field of social robotics [20].

This paradox in robot behaviour means that conventional computers are unable to efficiently process information in parallel. This means that modelling the human emotional system, known for its ability to exhibit simultaneous emotional experiences (such as joy and sadness), will be incompatible with these frameworks. Given all this, the enormous volume of data that future sensors will produce, the increasing complexity of control, planning, interaction, and reasoning systems, and the fact that future robots will operate in diverse environments; especially those that will facilitate human interactions-all combine to create this problem. And when these robots are connected via Internet of Things frameworks, the problem will worsen. All of these points to the need for future robotic systems to have enormous computing power [21]. At the same time, the properties of entanglement, superposition, and parallelism are said to give quantum systems amazing capabilities. There is widespread agreement and growing interest in using quantum sensors, quantum algorithms, and quantum controllers for important roles in future automated integrated systems and robots.

3.3 The importance of emotional computing, emotion analysis, and their applications

Emotional computing has gained widespread popularity due to its aim to improve human-computer interaction by creating devices and systems capable of identifying, expressing, and processing human emotions. Giving machines genuine intelligence and authentic human communication is crucial, as evidenced by the increasing demand for diverse emotion computing applications. To enhance emotion-related applications, there is a pressing need to develop further emotion computing capabilities, such as human-computer interfaces, emotion recognition, and analysis. Enterprise systems encompass numerous emotion-related applications, and their importance to businesses in the global economy cannot be overstated. They restructure the portfolio of transaction processing systems and applications within an organization to integrate business processes, systems, information, and data analytics. Various technologies have been employed to study enterprise systems due to their significant potential benefits. These technologies, spanning multiple fields, can dramatically improve the performance of enterprise systems. Production planning, procurement, marketing, and customer support are just a few examples of enterprise systems applications that process data related to user emotions [22].

Most enterprise IT technologies used today may overlook the importance of emotion, and enterprise systems can be supported by emotion computing techniques on a variety of devices and services. In e-commerce, computers can determine when a consumer is experiencing a problem and provide them with more help or information based on available emotional indicators. To target online purchases, measuring customer emotions in the digital world has become crucial. Fong noted that by predicting and modelling consumer behaviour, which is often influenced by the emotional state of the potential customer-marketing tactics, can be optimally tailored [23].

The investigation of [24] has shown that emotional factors have a significant impact on customer satisfaction and service provider change behaviour. This suggests that as the internet service provider market evolves, providers who focus on emotional factors and foster relationships with their customers will have a competitive advantage. When survey respondents are prone to lying and have strong emotions, Christopher demonstrated how using emotional data collection techniques in survey research provide insightful information. This study examines emotion computing techniques in a critical area of enterprise services and customer satisfaction measurement, given the importance of emotion computing in enterprise services.

3.4 Models and Techniques for Emotion Analysis

Computational modelling of human emotional processes has witnessed remarkable growth in recent years. Given its potential in fundamental research on emotion and cognition, as well as its promising and ever-expanding applications, this book serves as an excellent example of the fruitful collaboration that has emerged from emotion research in computer science and psychology. It explores the history of computational modelling of emotion, including its diverse applications and the theoretical traditions that have influenced its development, and how these applications and traditions are reflected in its fundamental structure. The aim is to understand this collaboration and its potential to revolutionize emotion research practices across various disciplines, revealing new and imperative research areas. While this may seem confusing to those unfamiliar with the subject, over the past fifteen years, a wide range of integrated and competing computational models have been developed, and a classification of some of the substantial models and the theoretical traditions from which they emerged is provided. Despite the increasing volume of research, this field is still far from mature. Research is rarely cumulative and often returns to initial stimulating ideas rather than building upon previous computational approaches [25].

Models are rarely compared in terms of their ability to achieve their stated goals, and the goals for which models are built are usually not explicitly defined. This chapter explores research on computational models of emotion in an attempt to uncover their common uses and the basic techniques and assumptions on which they are built. This contrasts with the fact that computational models are complex systems involving a number of design decisions and assumptions that may not be explicitly stated, and are inherited from the psychological and computational traditions from which they originated, a situation exacerbated by the lack of agreed-upon terminology to define these differences. Our aim is to provide common terminology and conceptual distinctions that facilitate the comparison and discussion of competing models, furthermore increase accessibility to this field for external researchers.

Hopefully, this will help establish a vocabulary that supports the field's progress toward more forward-looking research [25].

3.4.1 Techniques and Tools used in Emotion Analysis

The Web 2.0 era has evoked colossal amounts of unstructured data as individuals share their thoughts and experiences on social media platforms, including blogs, forums, rating sites, and social networks. Given the rapid growth of user-generated content, it has become essential for businesses, policymakers, service providers, and researchers to analyse these opinions to understand societal trends and improve decision-making. Sentiment analysis has emerged as an effective tool for extracting opinions and feelings from digital writing [26]. For transforming unstructured data into analysable information, sentiment analysis employs a set of integrated techniques and tools.

First, data is collected from social media platforms using web mining tools and content aggregation software.

Second, the text is prepared and filtered, removing irrelevant data and non-textual content, and addressing variations in writing styles, slang, and specialized terminology.

Third, NLP techniques are applied to extract subjective sentences that express the user's opinion or feelings, while objective sentences containing factual information are discarded.

3.5 The Specificity of the Arabic Language in Emotion Analysis and its Impact on Models and Tools

There are numerous individuals who produce and publish content across various online platforms. Social media platforms allow users to discuss and share their thoughts and feelings, making them a vital source of information. Before these platforms emerged, it was difficult to ascertain people's opinions on specific products, services, or even events. Thanks to social media sites like Twitter, where people freely and frequently express their views and even consult others' opinions when making decisions, this important information has become readily available. Users' opinions on specific services, events, or goods can be obtained by analysing these opinions. NLP techniques, terms extraction, computational linguistics, and machine learning can be used in this analysis to identify, study, and categorize human attitudes, feelings, opinions, sensations, evaluations, and emotions. However, emotion analysis is a key task within natural language processing, aiming to uncover subjective information encompassing feelings and experiences, which are then categorized as neutral, negative, or positive[27].

There are two types of textual data: opinions and truths. Truths are goal claims about items, events, and their hallmark. Opinions, on the other hand, are ordinarily subjective articulation of people's perceptions, evaluations, and outlooks of things, events, and their characteristics. Some research focuses on the polarity of opinion (positive or negative) when developing a model for assessing and classifying emotions. Notwithstanding, for developing a classifier which is capable of distinguishing between subjective and objective opinions, many researchers deliberate objective expression as a neutral category in emotion analysis. Temperament analysis and ordering differ from sensory analysis in that it delves deeper into describing an individual's feelings, allowing for more comprehensive classification and analysis; joy, anger, and sadness are examples of emotional states related to feelings, which are themselves components of the nervous system [28].

Emotion analysis techniques may involve identifying any emotions present in a text by classifying them into the appropriate emotional category. Twitter users post more than 400 million tweets daily, with information being disseminated as text messages in multiple languages. These tweets express users' feelings and emotions in different languages. The abundance of spelling and grammatical errors, colloquial language, social abbreviations, and multimedia content makes analysing emotions in tweets extremely difficult. Due to these challenges, no researcher has attempted to categorize attitudes in Arabic tweets, although multitudes of investigations have been fulfilled for emotions in English tweets. Most emotion analysis in Arabic tweets highlight on classifying emotions either positively or negatively. Research and resources in the speciality of social feeling analysis in Arabic are still scarce. Models that extract and classify emotions from Arabic tweets could be useful for various purposes, such as enhancing e-learning programs, helping psychologists identify terrorist activity, improving customer service, raising product quality, and much more [29].

3.6 Linguistic and Semantic Characteristics of the Arabic Language that Influence Emotion Analysis

Given its diverse applications across various sectors, including social media, marketing, government services, public opinion analysis, political monitoring, and more, sentiment analysis is one of the most crucial branches of NLP in the modern era. Sentiment analysis relies on extracting attitudes and emotions from written texts, identifying the type of emotion (such as joy, sadness, anger, fear, etc.), and determining whether these texts reflect positive, negative, or neutral feelings. While sentiment analysis has witnessed significant progress in Western languages, Arabic faces unique challenges that add to its complexity. These challenges stem from the distinctive linguistic and semantic characteristics of the Arabic language [29].

Grammatical and morphological rules interact with semantics and context, which directly affects the ability of computer systems to understand Arabic texts. The Arabic language is characterized by its wide morphological diversity, the richness of verb and noun conjugations, the complexity of its grammatical structures, the multiplicity of its meanings (the existence of words that carry more than one meaning), and its use of rhetorical styles such as metaphor, simile, and figurative language. Its complexity increases with the presence of dialects and colloquial language. The meaning of words and phrases in the Arabic language may change from one sentence to another or according to the relationships between words in the text. Therefore, context is a decisive factor.

3.7 The linguistic characteristics of the Arabic language and its effect on emotions

The Arabic language is distinguished by its unique textual and phonetic structure, where vocabulary carries explicit or implicit, positive or negative connotations that can directly influence the feelings and psychological states of the reader or listener. Expressive styles in Arabic are diverse, encompassing prose, poetry, praise, and satire, thus granting texts multiple semantic structures that allow for a clear or rhetorical understanding of emotions. Through these structures, the writer's style can be explored, and various aspects of emotion can be identified, such as sarcasm, praise, condemnation, resentment, affection, and enmity. The study of emotions has become particularly important within the framework of social media platforms like Twitter and Facebook, where the data generated is used to understand individuals' attitudes toward specific topics or products. However, Arabic Sentiment Analysis (ALSA) requires meticulous preliminary processing involving phonology, morphology, sentence segmentation, part-of-speech classification, semantic analysis, identification of named entities, as well as subjective and figurative analysis, and sometimes manual recording of opinions using dictionaries and linguistic lexicons [30].

Arabic sentiment analysis is usually accomplished by using a couple of basic approaches: rule-based classifiers or automated classifiers that rely on statistical machine learning algorithms to detect trends and opinions. Studies have shown that sentiment analysis in Arabic faces greater challenges compared to other languages, due to the complexity of its linguistic structure and the multiplicity of rhetorical and figurative expression levels. Therefore, researching the linguistic features of Arabic and their effect on sentiment helps in understanding how emotions and positive or negative attitudes are expressed in texts, and opens up practical horizons in various fields, including media, public opinion monitoring, marketing, and scientific research related to artificial intelligence and natural language processing [31].

3.8 Emotional Computing Tools and Techniques in the Arabic Language: Challenges and Opportunities

Sentiment analysis (SA) uses models from the fields of words mining, computational linguistics, and NLP to perform analyses on individual entry. Initially academic article dedicated to sentiment analysis in English had been released in 2008 by Bang et al. written material that express opinions, such as status upgrade, annotation, and discussions, are the raw material for the immense amounts of modern computerized data made available by the boom of internet communication, web-based forums, and website verification. Consequently, the study of sentiment analysis has promptly expanded to assimilate the repository and evaluate of this excessive of data. Sentiment analysis has been escalated and dominated for the areas of NLP research in the wake of early 21^{st} century. English is still at the peak of concentration for research in this field due to the ubiquitous usability of appropriate datasets recently. Accompany by increasing expression of emotions on merchandising and cyberspaces, where the individuals periodically share their reflection around specific component [32], [33]. All critiques perform as a provision for innovative end-users, producers, besides merchandising squads, yielding those entities together with worthwhile insights penetrate product excellence and promoting viewers compose enlightened verdicts around purchasing, production, or retail. The demanding of including these inquisitive studies is that, they have a pertinent cognition in textual form; this requires appropriate process to activate of retrieving purposeful data. Sentiment analysis tackles this difficulty by executing, categorizing, or grouping overviews dependent on client prerequisites. Hence, central intention of sentiment analysis is to ascertain if a portion of documenting either objective or subjective. Objective content has the absence for conveying any opinion, in the meantime the subjective text does reliant on the emotion expressed; subjective writing is broadly possibly to be categorized as positive, negative, or neutral [34].

To ascertain the spirit or affection manifested in the text, sentiment analysis involves determining the textual data. This can be useful for a diversity of purposes, to illustrate determining the overall sound of a document, identifying the feelings of precise component or themes among the text, and assessing the impression of particular scripts or lines. Arabic was chosen in the role of working language for several reasons. Owing to its structure and complex landscape, hence several of constrained resources are available for Arabic sentiment analysis (ASA). Despite this, the importance of analysing Arab is growing due to the size of its audience and the prominent position of the Arabic language itself, given its prosperous heritage, the strategic importance of its individual and territory, in addition to its cultural tradition, making it both arduous and appealing. The intricacies of the language, embracing its diverse accents and circumscribed provisions, present substantial difficulties in analysing Arab sentiment [35].

The Arabic language expresses unique setbacks for analysing emotions due to its complex linguistic precision. The complexities morphological structure of Arabic, which encompasses many intonations and discrepancies, potentially ambiguous the underlying orientation of emotion. The lacks of diacritical marks as well as prevailing ambiguity surrounding punctuation hinder accurate interpretation. One major ingredient of perceiving is the enormous spectrum of patois along the Arab community. That signifies the way of viewers convey their emotions, may vary greatly between dissimilar regions. Consequently, specific strategies must be adapted to accommodate these differences. Current approaches often struggle to effectively handle the subtle emotional phenomena of Arabic. Pervasive phenomena such as code mixing between Arabic and other languages, the ubiquitous employ of sarcasm, moreover the scarcity of a complete structure able with ease for misinforming techniques. The construction of multi-word expressions and phrases often deviates from the emotions of their constituent words in difficult ways. The emotional polarity of phrases is often domain and dialect dependent, necessitating dynamic, context-aware modelling [35].

4 Study Methodologies

4.1 Methodology

This study employed a descriptive-analytical approach, yearning for portray the nature of Arabic texts published on social media platforms and analyse their sentiments using affective computing techniques and natural language processing models. An applied experimental approach was also utilized to evaluate the performance of various models and tools in extracting and classifying opinions. This was achieved by applying these methods to a set of textual data and determining the accuracy and efficiency of the analysis.

4.2 Study Population

The study population comprised all Arabic tweets and textual posts on social media platforms such as Twitter, Facebook, and Instagram during the period from January 2024 to December 2025. The population was designed to represent the diversity of users in terms of age, gender, geographic location, and the dialect used in the texts.
The Study Sample: A stratified random sample of Arabic tweets was selected. The data was categorized by topic (news, products, services, social events) and dialect (Modern Standard Arabic, local dialects) to ensure representation of all forms of Arabic used. The sample consisted of 10,000 tweets, carefully distributed across different categories to ensure the accuracy and generalizability of the results.

4.3 Data Collection Tools

Data on model performance metrics (Accuracy, F1-Score, Precision, and Recall) were extracted for Word2Vec, BERT (multilingual), AraBERT, CAMeLBERT, and combinations involving Farasa and CamelTools. The datasets referenced in these studies primarily originated from social media platforms, ensuring relevance to the specified scope. The collected performance metrics were then compiled into a structured table for direct comparison. Additionally, libraries packages of Python scripts utilizing (`pandas`, `matplotlib` , `seaborn`, `numpy`, `scikit-learn`, `arabert`, `gensim`, `torch`, `aravec`) were to generate visual representations of the data, including table shows for Accuracy and F1-Score, and a line plot for an aggregated “Global Score” to illustrate overall model efficiency.
Data was collected using the following tools and techniques:

- Social media Application Programming Interfaces (APIs), such as the Twitter API, for collecting text tweets and analyses all contents, hence this called (X API) and gives the capabilities for programmatic access to X’s employed website. Furthermore, this comeback with representing objects (structured JSON).
- Text scraping tools for organizing and storing data suitable for analysis.
- Extensive and readily available Arabic text databases for training models, such as AraVec and Arabic Sentiment Lexicons.
  - AraVec: it is a powerful model that used for distributed embedding of word (Arabic Word Embedding Models), it is free to use (open source) which employed of supplying the Arabic NLP research intention. The Soliman et al. [36] illustrate the various steps for this model. Obtaining of information from standard twitter API, as well as discussed how the model build; where it had given the spectrum of discussion about clustering for both sentiment words and named entities. This model builds using gensim python library. However, it is necessary to install and import this tool to this project by using the command `pip install aravec`
- Arabic natural language processing tools, such as Farasa and CAMEL Tools, for analyzing and cleaning the text (removing symbols, correcting spelling and standardizing).

- CAMel Tools: is an assortment of development paradigm of natural language processing for Arabic that first come out at New York University Abu Dhabi by the CAMel laboratory; the website [37] render satisfactory guide for this techniques, it is stand for Computational Approaches to Modeling Language
- Farasa : it is Arabic language processing segmenter tools that applies support vector machine for classification, the website address [38] has prosperous reservoir of information which leads to cope with this technique and a swift acquiring knowledge. However, another handful of resources [39], [40] those facilitate the process of data understanding.
- Word2Vec: is a technique that performs machine learning mechanism , and it is adapt deep learning as methodology of work to train the models for understanding word semantic and embedding relationships, the sources [41] provide beneficial information to start with; as well as they have an impressive discussion for discovering progressively about this area.
- AraBERT technique: it is a pre-trained model for Arabic natural language processing based on BERT (*Bidirectional Encoder Representations from Transformers*) mechanism from Google; the AraBERT was invented by ***M**achine **IN**telligence **D**evelopment* (MIND) Lab at American University in Beirut-Lebanon [42], [43]. It is compulsory to install this tool inside Python environment by using Terminal command `pip install arabert`.
- Python 3.6.5 programming language: the implementation carried out by using JetBrains PyCharm Community Edition 2018.1.1 x64

Study Procedures

- Data Collection: Arabic text tweets were collected according to predefined categories, topics, and dialects.
- Initial Data Processing: This included cleaning the text, removing duplicates, correcting spelling errors and special characters, and standardizing the phrasing across different dialects as much as possible.
- Sentiment Analysis: Natural Language Processing (NLP) models such as Word2Vec, BERT, and AraBERT were applied to the categorized texts to classify emotions as positive, negative, and neutral.
- Performance Evaluation: The results of the different models were compared using metrics such as accuracy, recall, and F1-Sore.
- Results Analysis: The categorized data were analyzed to determine the overall distribution of emotions within the target population and the impact of Arabic linguistic specificities on model performance.

5. Results and Discussion

5.1 Model Performance Metrics

The following table summarizes the performance metrics of the evaluated sentiment analysis models on various Arabic social platforms datasets:

Model	Dataset Source	Accuracy	F1-Score	Precision	Recall	Global Score
Word2Vec (Skip-Gram)	Twitter/Facebook	0.785	0.772	0.768	0.776	**0.77525**
BERT (Multilingual)	Mixed Social Media	0.700	0.800	0.750	0.780	**0.7575**
AraBERT	Twitter (ArTwitter)	0.920	0.918	0.915	0.921	**0.9185**
CAMeLBERT	Noisy Social Media	0.920	0.915	0.917	0.923	**0.91875**
Farasa + AraBERT	ASTD (Twitter)	0.932	0.928	0.925	0.931	**0.929**
CamelTools + AraBERT	**Dialectal Arabic**	**0.895**	**0.889**	**0.882**	**0.896**	**0.8905**

Table (1) Sentiment Analysis Techniques Performance.

As depicted in the table above, AraBERT and CAMeLBERT demonstrate significantly higher Accuracy and F1-Scores compared to Word2Vec (Skip-Gram) and the generic BERT (Multilingual) model. This superior performance

can be attributed to their pre-training specifically on large Arabic corpora, enabling them to better capture the nuances of the Arabic language.
The combination of Farasa pre-processing with AraBERT yields the highest performance across all metrics, with an impressive Accuracy of 93.2% and an F1-Score of 92.8% on the ASTD (Twitter) dataset. This underscores the critical role of robust Arabic linguistic tools in enhancing the performance of even advanced transformer models by handling complex morphological structures and dialects. Similarly, CamelTools + AraBERT also show strong performance on dialectal Arabic, further emphasizing the benefit of specialized pre-processing for Arabic NLP tasks.

5.2Overall Model Efficiency

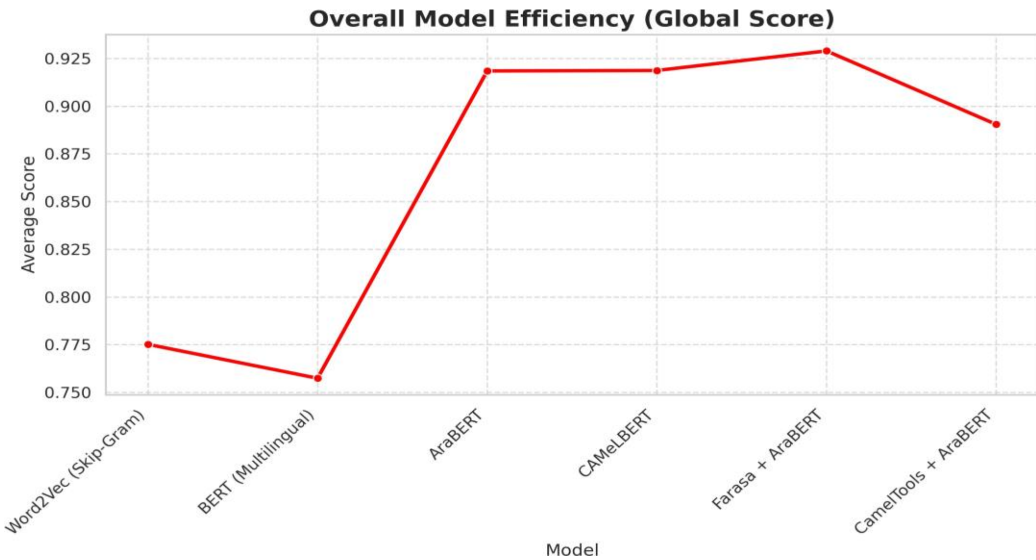

Figure (3) Global Score Techniques Performance.

The line plot illustrating the “Global Score” (average of Accuracy, F1-Score, Precision, and Recall) further reinforces these observations. Models leveraging Arabic-specific preliminary training and pre-processing, such as AraBERT, CAMeLBERT, and their combinations with Farasa or CamelTools, achieve substantially higher overall efficiency. The generic BERT (Multilingual) model, while better than Word2Vec in F1-score, shows lower accuracy, indicating its limitations when dealing with the specific complexities of Arabic without specialized adaptation.

5.3. Temporal and Spatial Boundaries

- Temporal Boundaries: The study includes Arabic tweets and posts published between January 2016 and 2026.
- Spatial Boundaries: The study focuses on Arabic-speaking users in Libya, taking into account dialectal and cultural diversity.

6. Conclusion

This study is of paramount importance because sentiment analysis and opinion extraction are emerging fields within artificial intelligence and affective computing. These disciplines are essential for understanding digital interactions between humans and deciphering the emotional patterns embedded in textual language. By elucidating the scientific foundations and various models of affective computing and sentiment analysis, this study contributes to a deeper theoretical understanding of these topics, thereby advancing scientific research in this field and building a more comprehensive and accurate knowledge framework.
From this point of view, What do affective computing and emotion analysis mean, and what are the theoretical foundations of these terms in the context of the Arabic language?; additionally, how linguistic, semantic, and contextual features characterize the Arabic language and influence emotion analysis and opinion extraction?
Hence, the objectives of this research investigate the most common affective computing models, tools, and techniques used in emotion analysis and how successful are currently available affective analysis models and techniques when applied to Arabic texts?
The implementation of sentiment analysis models demonstrate a remarkable ability to classify Arabic texts with varying degrees of accuracy, deep learning models outperform traditional models in analysing Arabic sentiment. The

duty shows the impact of morphological complexity and the multiplicity of Arabic dialects on the accuracy of analysis results. The importance of pre-processing texts in enhancing the efficiency of sentiment analysis models have been deliberated with emphasis on the objectives of the effectiveness of hybrid models in improving sentiment classification.
During fulfilling this investigation of this project, there had been spectrum of facts and figures such as the role of social media platforms as a primary data source for analysing Arabic sentiment. Likewise the limited availability of classified Arabic language resources hinders model development, and finally the potential for leveraging Arabic sentiment analysis to support decision-making across various fields.

6.1 Study Recommendations

1. Expand the development of Arabic-language resources specifically designed for sentiment analysis, covering Modern Standard Arabic and various dialects.
2. Employ modern deep learning models, particularly transformer-based models, in Arabic sentiment analysis studies.
3. Improve Arabic text preprocessing techniques to align with the language's morphological and syntactic characteristics.
4. Encourage the use of hybrid models that combine linguistic approaches and machine learning to enhance assessment accuracy.
5. Standardize assessment criteria for Arabic sentiment analysis models to ensure valid comparability across different research areas.
6. Focus on assessing emotions in Arabic dialects used on social media.
7. Support the creation of open-source Arabic databases for future research.
8. Expand the application of Arabic sentiment analysis in decision-making, public opinion assessment, and service review.

References

[1] Leo B., (2012). *Emotion Analysis and Opinion Extraction*. Synthetic Lectures in Human Language Technology, vol. 5, No.1, pp. 1-167.

[2] Taysir, H., et al. (2013). *Mining Social Networks' Arabic Slang Comments*. In Proceedings of IADIS European Conference on Data Mining, Czech Republic.

[3] Fouzi, Harrag. (2013). *Estimating the Sentiment of Arabic Social Media Contents*: A Survey.

[4] Khalid, K., and Nazlia, O., (2014). *A Hybrid Method Using Lexicon-Based Approach and Naive Bayes Classifier for Arabic Opinion Question Answering*. Journal of Computer Science, pp. 1961-1968.

[5] Basant, A., et al. (2015). *Concept-Level Sentiment Analysis with Dependency-based Semantic Parsing. Cognitive Computation manuscript*, pp. 1-13.

[6] Walaa, M, Ahmed, H. and, Korashy, H. (2014). *Sentiment analysis algorithms and applications: A survey*. Ain Shams Engineering Journal, pp. 1093-1113.

[7]Abdel-Majeed,M., M. T. Diab, (2011), *Subjectivity and Emotion Analysis in Modern Standard Arabic*, presented in the proceedings of the 49th Annual Meeting of the Computational Linguistics Association: Human Language Technologies: Short Papers , vol, 2, pp. 587-591.

[8]Badaro, J., Bali, R., and Haj, H. (2019). *A survey of opinion extraction in Arabic: A comprehensive systems perspective covering challenges and developments in tools, resources, models, applications, and concepts*. ACM Transactions in Information Processing in Asian and Resource-Limited Languages (TALLIP), vol. 18(3), pp. 1-52.

[9] Najjar, D., and Musfir, (2017). *Extracting opinions and analysing sentiments from Arabic texts published on the Internet: An application to the political field*. International Journal of Speech Technology, vol. 20(3), pp. 575-585.

[10] Salah, Z., and Al-Ghuwairi, A. R. (2019). *A systematic review of opinion extraction and sentiment analysis in social media.* International Journal of Business Information Systems, vol. 31(4), pp. 530-554

[11] Al-Muslimi, T., and Al-Bard, M. (2018). *A Dictionary of Arabic Emotions: Building Publicly Accessible Linguistic Resources for Analyzing Arabic Emotions*. Journal of Information Science, vol. 44(3), pp. 345-362.

[12] Hamtian, F., and Sohrabi, (2019). *A survey of classification techniques for opinion extraction and sentiment analysis*. AI Review, vol. 52(3), pp. 1495-1545.

[13] Alsemaree, O. (2024). *Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions*. School of Electronic Engineering and Computer Science, Queen Mary University of London, London, E1 4NS, UK

[14] Fouda, A. E., and Ahmed, K. S. (2024). *Sentiment Analysis in Arab Corporate Valuations*. In the 6th International Conference on Computing and Informatics (ICCI), pp. 418-428.
[15] Ruby, E., and Badawi, M. (2018). *Evaluating the performance of a sentiment analysis model used for Arabic Facebook comments*. Journal of Theoretical and Applied Information Technology, vol. 96(21), pp. 7098-7112.
[16] Ma, H., and Yarosh, S. (2021). *A review of affective computing research based on the Function and Component Representation Framework.* IEEE Transactions in Affective Computing, vol. 14(2), pp. 1655-1674.
[17]Tao, J., and Tan, T. (2005). *Emotional Computing: A Review*. In the International Conference on Emotional Computing and Intelligent Interaction. Berlin, Heidelberg: Springer Berlin-Heidelberg, pp. 981-995.
[18] Mobarz, H., Rashown, M., and Farag, I., (2014). *Using Automated Lexical Resources in Arabic Sentence Subjectivity.* International Journal of Artificial Intelligence & Applications (IJAIA), vol. 5, No. 6, pp. 1-14.
[19] Muñoz-Saavedra, L. et al. (2020). *Affective state assistant for helping Users with Cognitive Disabilities Using Neural Networks*. Electronics 4843, No.9, pp. 1-22.
[20] Horn, J. (2018). The Robot Brain Servant: *Designing a Partnership Between Humans and Artificial Systems*. Advanced Intelligent Computing Systems, 722, pp. 531-536.
[21] Yan, F., and Iliasso, A. M. (2021). *A conceptual framework for quantum affective computing and its application in integrating multiple robot emotions*. Electronics, vol. 10, No.2, pp. 1-17.
[22] Shaw, L. (2011). *Enterprise Systems: Current Status and Future Trends*. IEEE Transactions in Industrial and Information Technology, vol. 7(4), pp. 630-640.
[23] Fong, B. (2012). *Emotional computing in consumer electronics*, IEEE Transactions in Emotional Computing, vol. 3(2), pp. 129-131.
[24] Ren, F., and Cowan, S. (2012). *Emotion analysis and recognition using language to measure consumer satisfaction: an application of emotional computing*, Information Technology and Management, vol. 13(4), pp. 321-332.
[25] Marcella, S., Grach, J., and Beta, B. (2010). *Computational models of emotion*. A blueprint for emotional computing - a reference book and guide, vol. 11(1), pp. 21-46.
[26] D'Andrea, A., et al. (2015). *Methods, Tools, and Applications for Conducting Sentiment Analysis*. International Journal of Computer Applications (0975 – 8887), vol.125, No.3, pp. 26-33.
[27] Shukla, A. and Shukla, S. (2015). *A survey on classifying and analyzing emotions using data mining*. International Journal of Advanced Research in Computer Science, vol. 6, No. 7, pp. 20-25.
[28] Qiu, X., Sun, T. (2020). *Pre-trained models for natural language processing: a survey*. Chinese Journal of Science and Technology, vol. 3, pp. 1-26.
[29] Al-Qahtani, J. and Al-Uthaim, A. (2022). *Sentiment analysis in Arabic tweets: Language models and available resources*. Frontiers in Artificil Intelligence, vol. 5, pp. 843-838.
[30] Al-Halis, A.M. (2017*), Extracting Arabic opinions using distributed document representations*. In: Proceedings of the Palestinian International Conference on Information and Communication Technology (PICICT), Gaza, Palestine. pp. 28-33.
[31] Nouri Hicham, Sabri Karim1, Nassera Habbat.(2023). *Customer sentiment analysis for Arabic social media using a novel ensemble machine learning approach.* International Journal of Electrical and Computer Engineering (IJECE) vol. 13, No. 4, pp. 4504-4515.
[32] Abdullah, N., et al. (2013). *Arabic Sentiments Analysis: Lexicon-based and Corpus-based*, IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), IEEE, pp. 1-6.
[33] Sekhri, O. (2016). *Problem of terminology of the Arabic language*: *The Arabic language and the translation of emotive expressions into English*. pp. 1-20.
[34]Farah, N., Shleita, E. and Abu-Assi, R. (2010). *Extracting Sentence-Level and Document-Level Sentences from Arabic Texts*, in: IEEE International Conference on Data Mining Workshops, IEEE, pp. 1114-1119.
[35]Al-Alimi, A. A. and Al-Zahrani, A., (2024). Developments and challenges in Arab sentiment analysis: A decade of methodologies, applications, and resource development. Helion, 10.(21)
[36] Soliman, A., Eisa, K. and El-Beltagy, S. R. (2017). *AraVec: A set of Arabic Word Embedding Models for use in Arabic NLP*. In: proceedings of the 3rd International Conference on Arabic Computational Linguistics, Dubai, UAE.
[37] CAMel tools (Online). *CAMeL Tools Documentation*. Retrieved on 19/01/2026 from https://camel-tools.readthedocs.io/en/latest/
[38] Farasa main website (Online). Retrieved on 11/01/2026 from https://farasa.qcri.org/
[39] The founder of Farasa tool. (Online) retrieved on 22/02/2026 from https://mohamed.eldesouki.com/project/farasa
[40] Abdelali, A., Darwish, K., Durrani, N. and Mubarak, H. (2016). *Farasa: A Fast and Furious Segmenter for Arabic.* Qatar Computing Research Institute Hamad Bin Khalifa University. Proceedings of NAACL-HLT, pp. 11–16, San Diego, California.

[41] Coveo website (on line). *Word2Vec Explained: What IT Leaders Need to Know About NLP & AI.* Retrieved on 13/01/2026 from https://www.coveo.com/blog/word2vec-explained/
[42] Fay Aljomah, F., et. al (2025). *Enhancing Arabic Sentiment Analysis with Pre-Trained CAMeLBERT: A Case Study on Noisy Texts.* Tech Science Press, pp. 5317-5335.
[43] Antoun, W., Baly, F., and Hajj, H. (2020). *AraBERT: Transformer-based Model for Arabic Language Understanding*. Proceedings of the Twelfth International Conference on Language Resources and Evaluation (LREC 2020), Marseille, France, pp.1-7.
[44] AUB's **M**achine **IN**telligence **D**evelopment Lab (Online).*The Use of Arabic Language in Artificial Intelligence on the Researcher*. Retrieved on 24/12/2025 from https://sites.aub.edu.lb/mindlab/