iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60250

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Analyse the Enhancement of Sentiment Analysis in Arabic by doing a Comparative Study of Several Machine Learning Techniques

Mohamed Fathy Abd-Elshafy[1], Dr.Tarek Aly[2], Prof. Mervat Gheith[3]
Faculty of Graduate Studies for Statistical Research Cairo University

Abstract: ***Sentiment analysis is a crucial component of natural language processing that seeks to determine the emotional sentiment expressed in a given text. This study investigates sentiment analysis in the Arabic language through a comprehensive approach that integrates traditional machine learning methods with sophisticated deep learning models. We examine the efficacy of conventional algorithms such as Support Vector Machines (SVM) and Naive Bayes, as well as sophisticated neural network architectures such as Convolutional Neural Networks (CNNs), Long Short-Term Memory networks (LSTMs), and the Arabic variant of Bidirectional Encoder Representations from Transformers (BERT). The primary novelty of our approach is in the ensemble method, which combines many approaches to enhance the precision of sentiment categorization in Arabic text. To address the particular challenges presented by Arabic sentiment analysis, such as the intricate structure of the language and the diverse regional variations, we utilize a tailored preprocessing pipeline to effectively handle the nuances of Arabic text. Our comprehensive analysis of various datasets demonstrates that the ensemble technique outperforms individual model benchmarks and offers novel insights into the interplay between different machine learning paradigms in Arabic NLP. The results emphasize the ability of hybrid approaches to improve Arabic sentiment analysis, providing a solid basis for future research and practical applications in understanding the sentiments of Arabic consumers. This study is a significant addition to the expanding domain of Arabic Natural Language Processing (NLP). This resource offers a comprehensive and advanced methodology for utilizing machine learning and deep learning methods to comprehend and analyse the intricate aspects of sentiment in the Arabic language.***

Keywords: Arabic Sentiment Analysis, Traditional Machine Learning, Deep Learning, CNN, LSTM, BERT, Ensemble Methods, Natural Language Processing

I. INTRODUCTION

Sentiment analysis is a prominent field within Natural Language Processing (NLP) that utilizes computational linguistics, text analysis, and machine learning to detect, extract, and measure emotional states and subjective information from text. The use of this technology is especially difficult in Arabic because of the language's complex morphology and wide range of dialects.

A. Natural Language Processing (NLP) Components

Comprehending sentiment in text requires the utilization of multiple Natural Language Processing (NLP) components:

1) Syntax refers to the organization of words and phrases in a language to generate grammatically correct sentences.
2) Semantics: The significance transmitted by written language, which is essential for understanding sentiment.
3) Pragmatics refers to the contextual factors that play a crucial role in determining the intended meaning of a document. It is particularly important for doing nuanced sentiment analysis.

B. Application of Machine Learning in Sentiment Analysis

C. Machine learning algorithms simplify the process of extracting sentiment from text by adjusting to the intricate and subtle nuances of language.

D. Algorithms for Machine Learning

Every algorithm presents a distinct method for categorizing sentiment:

1) *Support Vector Machine (SVM)* is a powerful method for text categorization that utilizes hyperplanes in a high-dimensional space to distinguish between distinct sentiment groups [1].

©IJRASET: All Rights are Reserved

2) *The Decision Tree* model offers a distinct method for decision-making, represented as a tree structure, to categorize sentiment [2].
3) *Linear Support Vector Machine (SVM):* A customized version of SVM that is most effective when dealing with data that can be separated by a straight line, providing advantages in terms of efficiency and scalability [3].
4) *Logistic Regression:* A commonly used model for binary classification, which utilizes a logistic function to estimate probabilities [4].
5) *Random Forest:* A technique that enhances the accuracy of categorization by merging many decision trees [5].
6) *Naive Bayes is a probabilistic classifier* that utilizes Bayes' theorem, assuming that the features are independent of each other [6].

E. Data Preprocessing Efficient

Data preprocessing improves model accuracy by removing noise and standardizing text data, while also tackling the specific difficulties posed by Arabic script and dialects.

F. Enhancing Data

Improving training datasets using methods such as replacing synonyms and translating text back and forth enhances the resilience of the model and effectively tackles the problem of limited data availability.

G. The topic of this text is "Deep Learning in Sentiment Analysis".

1) *Deep learning models*, characterized by their hierarchical structure, are highly effective at collecting intricate patterns in data, surpassing conventional algorithms in numerous natural language processing (NLP) applications.
2) *Convolutional Neural Networks (CNN)* are a type of neural network that are specifically designed for processing and analysing visual data. They are widely used in computer vision tasks such as image classification, object detection, and image segmentation [7].
3) *CNNs*, renowned for their exceptional ability to interpret images, have been successfully modified to analyse text, finding patterns that reveal the sentiment expressed in sequences of words.
4) *LSTM (Long Short-Term Memory Networks)* LSTMs effectively manage extended connections in textual data, which are crucial for comprehending the overall meaning and subtle intricacies in sentiment analysis [8].
5) *Arabic BERT* utilizes the transformer architecture to achieve profound contextual representation, hence improving the precision of sentiment categorization in Arabic [9].

H. Methods that use many models to make predictions, known as ensemble methods

Ensemble approaches strive to enhance the accuracy and stability of sentiment classification by merging predictions from multiple models.

1) Lexical Embedding

Word2Vec and GloVe are techniques that transform words into vectors, which capture both semantic and syntactic similarities. These similarities are crucial for machine learning models [10].

2) Tokenization refers to the process of breaking down a text into smaller units called tokens.

Tokenizing text is an essential process in NLP, which can be particularly challenging for Arabic because of its unique alphabet and morphology.

3) TF-IDF stands for Term Frequency-Inverse Document Frequency.

This statistic demonstrates the significance of words in papers, aiding in the identification of keywords that are highly suggestive of sentiment [11]. Utilizing the TF-IDF method to ascertain the significance of words in document queries.

4) Bag of Words (BoW)

Bag-of-Words (BoW) represents text as a disorganized assortment of words, which is a straightforward yet effective method for sentiment classification.

This project seeks to develop a complete framework for Arabic sentiment analysis by including several approaches. The objective is to improve the comprehension and processing of sentiment in Arabic text, thereby making a valuable contribution to the broader field of natural language processing (NLP)

II. CHALLENGE ARABIC LANGUAGE WITH NLP

The Arabic language poses unique challenges for Natural Language Processing (NLP) and machine learning due to its intrinsic linguistic traits and extensive usage.

The challenges significantly affect the development and functioning of computer models specifically designed for processing Arabic text. Acquiring a thorough comprehension of these issues is crucial for researchers and practitioners to develop more effective algorithms and tools for Arabic natural language processing (NLP) tasks.

A. Significant morphological complexity

Arabic is a highly inflected language, meaning that words are constructed by adding prefixes, suffixes, and infixes to root patterns. The profusion of morphological variants in words poses significant challenges for the tokenization, stemming, and lemmatization processes in natural language processing (NLP).

B. Geographical Variations in Linguistic Patterns

The Arabic language consists of numerous dialects that display significant diversity across different regions. These dialects often diverge considerably from Modern Standard Arabic (MSA) in terms of vocabulary, sentence structure, and word formation. Natural Language Processing (NLP) models trained on Modern Standard Arabic (MSA) may not demonstrate enough performance when used with dialectal Arabic. This emphasizes the necessity of developing models that are specific to different dialects or constructing models that are capable of properly handling these variations in language.

C. Li 3.3. Lack of Diacritics in Written Text

Diacritics in Arabic are necessary for accurately representing pronunciation and differentiating between various word significations. However, most written texts, especially those available on digital media, do not include diacritics. The absence of presence can lead to ambiguities and misconceptions, as numerous words may share the same spelling yet differ in meaning and pronunciation.

D. Ambiguity can arise within a certain situation.

Arabic words possess a significance that is significantly shaped by the surrounding circumstances in which they are employed. This is a challenge for machine learning models to accurately capture and comprehend the intended meaning, especially in tasks like sentiment analysis that require understanding subtle distinctions.

E. Limited resources and datasets with annotations

Despite some advancements in the development of Arabic natural language processing (NLP) resources, the availability of comprehensive and high-quality annotated datasets, especially for dialectal Arabic, is still limited compared to the resources available for English. This limitation hinders the advancement and training of robust machine learning models for the Arabic language.

F. Code-switching in language

Code-switching, the act of shifting between Arabic and other languages (often English or French), is prevalent in Arabic-speaking societies. This strategy introduces additional complexity for NLP systems, which must handle many languages inside a single document.

G. Idioms and expressions that are distinctive to a specific language or culture.

Arabic speakers frequently use informal expressions, sayings, and cultural allusions, which can be difficult for NLP systems to understand without a deep understanding of the cultural context.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

H. Addressing the challenges

To overcome these challenges, it is crucial for the NLP community to give priority to developing sophisticated models that specifically account for the complex morphology of the Arabic language. Furthermore, it is important to exert efforts to create and manage comprehensive and varied databases. Moreover, it is necessary to employ innovative techniques to efficiently manage differences in dialects and nuances in context. To advance Arabic Natural Language Processing (NLP) and effectively leverage machine learning for Arabic language processing, it is imperative to sustain continuous collaboration among linguists, data scientists, and domain specialists.

III. LITERATURE REVIEW

The field of sentiment analysis in Arabic text showcases the complex interaction between language complexities and computer approaches. With a focus on word embedding and finding irony in Arabic sentiment analysis, this paper looks at the basic ideas that support this research. These include traditional machine learning methods, advanced deep learning methods, and the new field of hybrid models.

A. Utilizing Traditional Machine Learning Techniques for Arabic Sentiment Analysis

Based on the latest research, the utilization of traditional machine learning techniques for Arabic sentiment analysis has been a subject of numerous studies. Here's a synthesis of the related work in this area, structured to fit into a new paper's related work section:

1) Traditional machine learning techniques have been extensively employed in Arabic sentiment analysis, with various studies demonstrating their effectiveness across different datasets and contexts. Basabain et al. explored the prevalent techniques in Arabic topic-based sentiment analysis, highlighting the use of traditional approaches alongside limited deep learning methods for building high-performance models ("A Survey of Arabic Thematic Sentiment Analysis Based on Topic Modeling") [12].
2) Mazari and Djeffal (2021) applied both traditional machine learning algorithms and deep learning models to analyze sentiments in dialectal Arabic texts, particularly focusing on the Algerian dialect during the Hirak_19 social movement, with CNN and RNN achieving notable accuracy scores in cross-validation tests ("Deep Learning-Based Sentiment Analysis of Algerian Dialect during Hirak 2019") [13].
3) Further extending the application of traditional techniques, Ouchene and Bessou employed machine learning methods like SVM and Naive Bayes, along with deep learning models, to analyze sentiment in Algerian Dialect tweets, highlighting the blend of traditional and modern approaches in sentiment analysis ("Sentiment analysis for Algerian Dialect tweets") [14].
4) Abdelwahab et al. demonstrated the use of the sentiment keywords co-occurrence measure (SKCM) algorithm, a traditional machine learning technique, to enhance sentiment analysis accuracy in Arabic, showcasing the adaptability of these methods to language-specific challenges ("Enhancing the Performance of Sentiment Analysis Supervised Learning Using Sentiments Keywords Based Technique") [15].
5) In the context of the Saudi stock market, Alazba et al. reported an accuracy of 79.08% using SVM with TF-IDF feature extraction method, illustrating the efficiency of traditional methods in sector-specific sentiment analysis ("Saudi Stock Market Sentiment Analysis using Twitter Data") [16].
6) Al-Twairesh discussed the performance improvement in Arabic sentiment analysis using traditional machine learning methods and BERT models, emphasizing the significant strides made with these techniques ("The Evolution of Language Models Applied to Emotion Analysis of Arabic Tweets") [17].
7) Hicham et al. evaluated various machine learning techniques, including ensemble approaches, showcasing their superior performance in accuracy and other metrics, underscoring the potential of traditional methods in Arabic sentiment analysis ("Customer sentiment analysis for Arabic social media using a novel ensemble machine learning approach") [18]. Comparatives are shown Table 1.

Table 1
Analytical comparison between papers

Study Reference	Techniques Used	Focus Area	Data Context	Notable Outcomes
Basabain et al. [12]	Traditional ML techniques & limited	Topic-based sentiment analysis	General Arabic text	Highlighted the effectiveness of traditional approaches in

	deep learning			conjunction with deep learning.
Mazari and Djeffal [13]	ML algorithms & DL models (CNN, RNN)	Sentiment analysis in dialectal Arabic	Algerian dialect during Hirak_19	Demonstrated notable accuracy with CNN and RNN in dialectal sentiment analysis.
Ouchene and Bessou [14]	SVM, Naive Bayes, & DL models	Sentiment analysis in Algerian Dialect tweets	Algerian Dialect tweets	Showcased the blend of traditional and modern approaches in sentiment analysis.
Abdelwahab et al. [15]	Sentiment Keywords Co-occurrence Measure (SKCM) algorithm	Enhancing sentiment analysis accuracy	General Arabic text	Used SKCM to improve sentiment analysis accuracy, showing adaptability to Arabic.
Alazba et al. [16]	SVM with TF-IDF	Sentiment analysis in the Saudi stock market	Saudi stock market-related tweets	Achieved 79.08% accuracy, illustrating the efficiency of SVM with TF-IDF in a specific sector.
Al-Twairesh [17]	Traditional ML methods & BERT models	Performance improvement in Arabic sentiment analysis	Arabic tweets	Discussed the significant strides made with traditional techniques alongside BERT.
Hicham et al. [18]	Various ML techniques including ensemble approaches	Customer sentiment analysis	Arabic social media	Showcased superior performance with ensemble methods in accuracy and other metrics.

These studies collectively demonstrate the robustness and versatility of traditional machine learning techniques in the field of Arabic sentiment analysis, offering a comprehensive backdrop for ongoing and future research in this area.

B. Utilizing Deep Learning Methods for Arabic Sentiment Analysis

In the realm of Arabic sentiment analysis, deep learning methods have shown significant advancements, offering robust solutions to the nuanced challenges of processing Arabic text. Here's an overview of the relevant studies utilizing deep learning for Arabic sentiment analysis, suitable for a related work section in a new paper:

1) Alqarni and Rahman utilized Convolutional Neural Networks (CNN) and Bi-directional Long Short Memory (BiLSTM) to classify sentiments in Arabic tweets related to COVID-19, achieving accuracies of 92.80% and 91.99%, respectively ("Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach") [19].

2) Saleh et al. developed a Heterogeneous Ensemble Deep Learning Model that combined RNN, LSTM, GRU with LR, RF, and SVM meta-learners, showing enhanced performance in Arabic sentiment analysis compared to traditional techniques ("Heterogeneous Ensemble Deep Learning Model for Enhanced Arabic Sentiment Analysis") [20].

3) Mhamed and Noja discussed enhancing Arabic sentiment analysis using a hybrid deep learning approach, demonstrating the effectiveness of combining various deep learning models ("Enhancing Arabic Sentiment Analysis Through a Hybrid Deep Learning Approach") [21].

4) Elhassan et al. explored Arabic sentiment analysis using word embeddings and deep learning, particularly highlighting the effectiveness of CNNs, LSTM, and hybrid CNN-LSTM models with fastText and Word2Vec embeddings ("Arabic Sentiment Analysis Based on Word Embeddings and Deep Learning") [22].

5) Ombabi et al. implemented a deep learning CNN-LSTM framework for sentiment analysis on Arabic social network data, showcasing the synergy between convolutional and recurrent neural networks in this context ("Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks") [23].

6) Alhumoud et al. used deep learning models like SGRU, SBi-GRU, and AraBERT for sentiment analysis on Arabic vaccine-related tweets, noting the ensemble model's superior accuracy ("ASAVACT: Arabic sentiment analysis for vaccine-related COVID-19 tweets using deep learning") [24]. Comparatives are shown Table 2.

Table 2
Analytical comparison between papers

Study Reference	Deep Learning Techniques Used	Focus Area	Data Context	Achievements
Alqarni and Rahman [19]	CNN, BiLSTM	Sentiment classification	Arabic tweets related to COVID-19	Achieved high accuracies of 92.80% (CNN) and 91.99% (BiLSTM).
Saleh et al. [20]	RNN, LSTM, GRU, LR, RF, SVM (Heterogeneous Ensemble)	Enhanced performance	General Arabic text	Demonstrated enhanced performance over traditional techniques with a deep learning ensemble.
Mhamed and Noja [21]	Hybrid deep learning approach	Enhancing Arabic sentiment analysis	General Arabic text	Showed effectiveness in combining various deep learning models for improved analysis.
Elhassan et al. [22]	CNN, LSTM, CNN-LSTM hybrid with fastText, Word2Vec	Using word embeddings and deep learning	General Arabic text	Highlighted the success of hybrid models and the impact of word embeddings on performance.
Ombabi et al. [23]	CNN-LSTM framework	Sentiment analysis on social network data	Arabic social network data	Showcased the synergy between CNN and LSTM in analyzing social media sentiments.
Alhumoud et al. [24]	SGRU, SBi-GRU, AraBERT (ensemble model)	Sentiment analysis on vaccine-related tweets	Arabic vaccine-related tweets	Noted the superior accuracy of ensemble deep learning models in sentiment analysis.

These studies exemplify the diverse and effective use of deep learning methods in Arabic sentiment analysis, indicating a strong trend towards more sophisticated, nuanced, and accurate sentiment analysis in this language domain.

C. Hybrid Approaches in Arabic Sentiment Analysis

Hybrid approaches in Arabic sentiment analysis blend various methodologies to enhance accuracy and adaptability to the Arabic language's intricacies. Here's a compilation of studies focusing on hybrid methods in this domain:

1) Essam et al. explored hybrid classifiers in their study on Arabic tweets, achieving an accuracy of 75% and an F-measure of 74.1%, demonstrating the effectiveness of hybrid methods in analysing sentiments from Arabic tweets ("Arabic Tweets Sentiment Analysis using Hybrid Approaches") [25].
2) In the context of sentiment analysis in English and Arabic languages, Guellil, Azouaou, and Valitutti surveyed 100 papers, highlighting the prevalent use of hybrid approaches in Arabic sentiment analysis ("English vs Arabic Sentiment Analysis: A Survey Presenting 100 Work Studies, Resources and Tools") [26].
3) Abdulla et al. focused on lexicon-based sentiment analysis for Arabic, comparing different lexicon construction techniques and achieving 74.6% accuracy, underscoring the potential of hybrid methodologies in this field ("Automatic Lexicon Construction for Arabic Sentiment Analysis") [27].
4) Alhumoud et al. implemented a hybrid learning approach using SVM and K-Nearest Neighbours (KNN) for sentiment analysis of Twitter's Saudi dialect tweets, showing superior results compared to a supervised approach ("Arabic sentiment analysis using WEKA a hybrid learning approach") [28].
5) Shahad Abuuznien et al. applied SVM with stemming in a comparative supervised learning approach for Sudanese Arabic Dialect, achieving notable success with a F1-score of 0.71 and accuracy of 0.95 ("Sentiment Analysis for Sudanese Arabic Dialect Using Comparative Supervised Learning approach") [29].
6) Al-Rubaiee, Qiu, and Li employed a mix of natural language processing and machine learning techniques for sentiment analysis of Arabic tweets regarding Mubasher products, classifying sentiments into positive, negative, and neutral categories ("Identifying Mubasher software products through sentiment analysis of Arabic tweets") [30]. Comparatives are shown Table 3.

Table 3
Analytical comparison between papers

Study Reference	Hybrid Techniques Used	Focus Area	Data Context	Notable Outcomes
Essam et al. [25]	Hybrid classifiers	Arabic tweets sentiment analysis	Arabic tweets	Achieved 75% accuracy and an F-measure of 74.1%, demonstrating the hybrid method's efficacy.
Guellil, Azouaou, & Valitutti [26]	Review of hybrid approaches	Survey on sentiment analysis	English and Arabic languages	Highlighted the prevalence and success of hybrid approaches in Arabic sentiment analysis across 100 studies.
Abdulla et al. [27]	Lexicon-based techniques & comparison of lexicon construction methods	Lexicon-based sentiment analysis	Arabic text	Attained 74.6% accuracy, illustrating the potential of hybrid lexicon-based approaches.
Alhumoud et al. [28]	SVM and K-Nearest Neighbours (KNN)	Sentiment analysis of Twitter's Saudi dialect	Saudi dialect tweets	Showed superior results with hybrid learning over a supervised approach.
Shahad Abuuznien et al. [29]	SVM with stemming	Sudanese Arabic Dialect analysis	Sudanese Arabic Dialect	Reported an F1-score of 0.71 and an accuracy of 0.95, indicating the hybrid approach's success.
Al-Rubaiee, Qiu, & Li [30]	NLP and machine learning techniques	Sentiment analysis of Mubasher products	Arabic tweets about Mubasher products	Classified sentiments into positive, negative, and neutral, showcasing the effectiveness of the hybrid approach.

These studies exemplify the diverse application and effectiveness of hybrid approaches in Arabic sentiment analysis, indicating their potential to address the language-specific challenges and enhance the performance of sentiment analysis systems.

D. Arabic Word Embedding in Natural Language Processing

Arabic Word Embedding plays a crucial role in enhancing the performance of Natural Language Processing (NLP) tasks. Here's a summary of research in this area, ideal for someone looking to understand the current landscape:

1) Guellil et al. provided an overview of Arabic natural language processing, surveying 90 recent research papers to cover resources and tools for various Arabic varieties, highlighting the need for comprehensive tools and resources in this domain ("Arabic natural language processing: An overview") [31].

2) In a different study, researchers focused on the challenges and methodologies specific to Arabic word embedding within NLP, acknowledging the unique characteristics of the Arabic script and language that necessitate specialized approaches ("Arabic Word Processing").

3) The "Proceedings of the Second Workshop on Arabic Natural Language Processing" offered insights into recent advancements and research themes in Arabic NLP, although the abstract does not detail specific findings related to word embedding [32].

4) Elayeb reviewed Arabic word sense disambiguation, a crucial aspect of word embedding, discussing common techniques and challenges in the field ("Arabic word sense disambiguation: a review") [33].

5) Moreover, a recent paper introduced new Arabic BERT-style and T5-style models, showing their effectiveness in NLP tasks, which signifies a significant step forward in Arabic contextualized word embeddings ("Revisiting Pre-trained Language Models and their Evaluation for Arabic Natural Language Processing") [34].

6) Lastly, Elnagar et al. presented a benchmark for evaluating Arabic contextualized word embedding models, an essential tool for researchers and practitioners in Arabic NLP to gauge the effectiveness of their models ("A benchmark for evaluating Arabic contextualized word embedding models") [35]. Comparatives are shown Table 4.

Table 4
Analytical comparison between papers

Study Reference	Focus Area	Key Contributions	Context or Dataset	Notable Findings
Guellil et al. [31]	Overview of Arabic NLP	Surveyed 90 papers to highlight resources and tools for Arabic NLP	Various Arabic varieties	Emphasized the need for comprehensive tools and resources in Arabic NLP.
Research on Arabic Word Embedding	Specific challenges in Arabic word embedding	Addressed unique characteristics of Arabic script and language	Arabic text	Acknowledged the need for specialized approaches in Arabic word embedding.
Proceedings of the Second Workshop [32]	Advances in Arabic NLP	Provided insights into recent research themes, including word embedding	Arabic NLP	Although details on word embedding were limited, the workshop highlighted current research directions.
Elayeb [33]	Arabic word sense disambiguation	Reviewed techniques and challenges specific to Arabic	Word sense disambiguation	Discussed the intricacies of word sense disambiguation in Arabic, critical for effective word embedding.
Recent Paper on Arabic BERT/T5 Models [34]	Arabic contextualized word embeddings	Introduced BERT-style and T5-style models for Arabic	Various NLP tasks	Demonstrated the effectiveness of new Arabic BERT and T5 models in contextualized word embeddings.
Elnagar et al. [35]	Evaluation of Arabic word embedding models	Presented a benchmark for evaluating Arabic contextualized word embedding models	Arabic text	Provided a benchmarking tool, aiding in the assessment of Arabic word embedding models' effectiveness.

These studies collectively underscore the evolving landscape of Arabic Word Embedding in NLP, demonstrating ongoing efforts to tailor NLP tools and methodologies to the nuances of the Arabic language.

E. Detecting Irony in Arabic Sentiment Analysis

Detecting irony in Arabic sentiment analysis is a complex task that has garnered attention in recent research. Here's a synthesis of the related work in this area:

1) Shah et al. addressed the challenge of detecting sarcasm and classifying sentiment in Arabic text data through a Modified Switch Transformer (MST) model, demonstrating the model's effectiveness in understanding nuanced expressions like sarcasm ("Arabic Sentiment Analysis and Sarcasm Detection Using Probabilistic Projections-Based Variational Switch Transformer") [36].

2) Rahma, Azab, and Mohammed highlighted the emergence of Arabic sarcasm detection studies since 2017, noting the significant advancements in AI techniques applied to this field, marking a growing interest and the evolving complexity of Arabic sentiment analysis ("A Comprehensive Survey on Arabic Sarcasm Detection: Approaches, Challenges and Future Trends") [37].

3) El Mahdaouy et al. introduced a deep multi-task model for sarcasm detection and sentiment analysis in Arabic, showcasing the benefits of an end-to-end deep MTL model in enhancing the detection of sarcasm alongside sentiment in Arabic text, outperforming single-task models ("Deep Multi-Task Model for Sarcasm Detection and Sentiment Analysis in Arabic Language") [38].

4) Alhaidari, Alyoubi, and Alotaibi employed deep convolutional neural networks to detect irony in Arabic microblogs, achieving significant results and underscoring the potential of deep learning in identifying complex linguistic constructs like irony ("Detecting Irony in Arabic Microblogs using Deep Convolutional Neural Networks") [39]. Comparatives are shown Table 5.

Table 5
Analytical comparison between papers

Study Reference	Methodology Used	Focus Area	Data Context	Key Findings
Shah et al. [36]	Modified Switch Transformer (MST) model	Sarcasm detection and sentiment classification	Arabic text data	Demonstrated the MST model's capability in effectively understanding sarcasm, enhancing sentiment analysis.
Rahma, Azab, & Mohammed [37]	Survey of techniques	Review of Arabic sarcasm detection	Review since 2017	Highlighted significant advancements in AI for Arabic sarcasm detection, noting the field's growth and complexities.
El Mahdaouy et al. [38]	Deep multi-task model	Sarcasm detection and sentiment analysis	Arabic text	Showcased the advantages of a deep multi-task learning (MTL) model, outperforming single-task models in detecting sarcasm and analysing sentiment.
Alhaidari, Alyoubi, & Alotaibi [39]	Deep Convolutional Neural Networks (CNN)	Irony detection in microblogs	Arabic microblogs	Achieved significant results in irony detection using deep CNNs, indicating the effectiveness of deep learning in this context.

These studies demonstrate the evolving methodologies and technologies in detecting irony within the realm of Arabic sentiment analysis, reflecting the unique challenges and ongoing advancements in this field.

IV. DATASET AND PREPROCESSING

We will utilize around 67,000 Arabic evaluations from the dataset for the purpose of conducting sentiment analysis. A multitude of companies employ web scraping as a means to obtain data. Several online platforms for ordering food include Talabat, Kabiter, Nasla, Swifil, Alsiwidiu, Kilubatra, Dumati, and others. The rating system comprises three choices: 1 for positive, 0 for neutral, and -1 for negative. This dataset is essential in the domain of Arabic Natural Language Processing (NLP), providing a valuable asset for academics and developers to train, assess, and improve sentiment analysis models. The dataset is specifically tailored for the Arabic language and can be used in many applications such as market research, analysis of consumer feedback, and monitoring of social media.

A. Description of the Dataset

The dataset consists of Arabic reviews, each labelled with a sentiment evaluation that categorizes the review as good, negative, or neutral. Figure 1 provides a detailed analysis of the distribution of sentiments within the sample. Dispersion Raw data:

1) There are 23,753 occurrences of neutral reviews with a rating of "0".
2) The number of negative reviews with a rating of "-1" is 23,035.
3) Positive Reviews: There are 20,339 instances.

B. Data preprocessing, also known as data cleaning, is the initial step in preparing raw data for analysis or modelling. The architecture of our data preprocessing pipeline is specifically crafted to improve the quality of the textual data, rendering it acceptable for subsequent analysis or model training. The primary focus is on Arabic text, requiring specific normalization and cleaning techniques to address the unique characteristics of the language. The steps are illustrated in Figure 2, titled "Preprocessing Data."

1) *Letter Normalization*: Our goal is to establish a uniform standard for the Arabic script by addressing the discrepancies arising from the presence of many versions of the same letter. For instance, the several iterations of the letter alif (ا, أ, آ, إ) are standardized to a consistent shape. This normalization also extends to other letter variants, such as the merging of ي and ى. This stage minimizes the disparities in the representation of text, enhancing the consistency of the dataset.
2) *Whitespace Handling*: Any unnecessary spaces, whether they are at the start, finish, or within the text, are removed. This safeguard ensures that the existence of spaces does not lead to any bias or irregularity in the examination.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

3) *Missing Value Handling*: We examine the dataset to detect any missing values in the text columns. If any missing values are identified, we substitute them with empty strings. This method reduces the likelihood of any errors that may arise during the tokenization or vectorization operations.

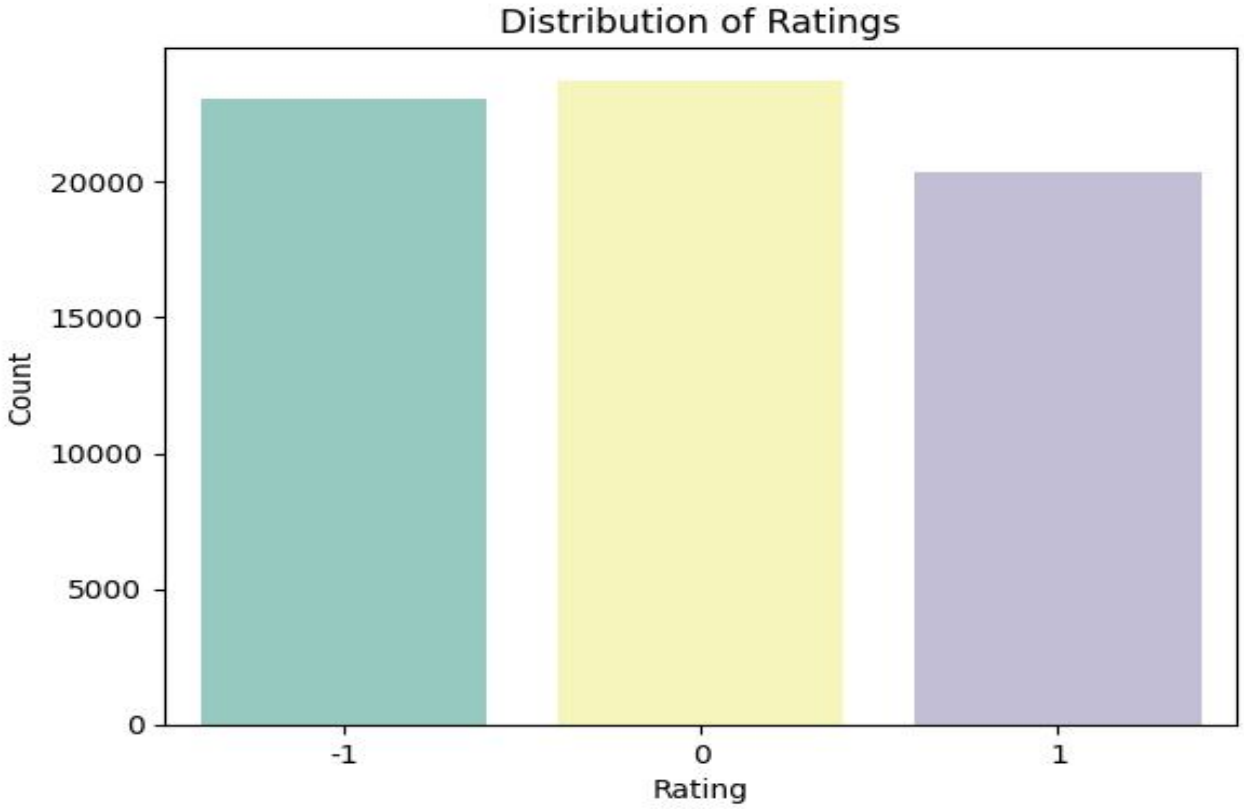


Figure 1 Distribution Original Data

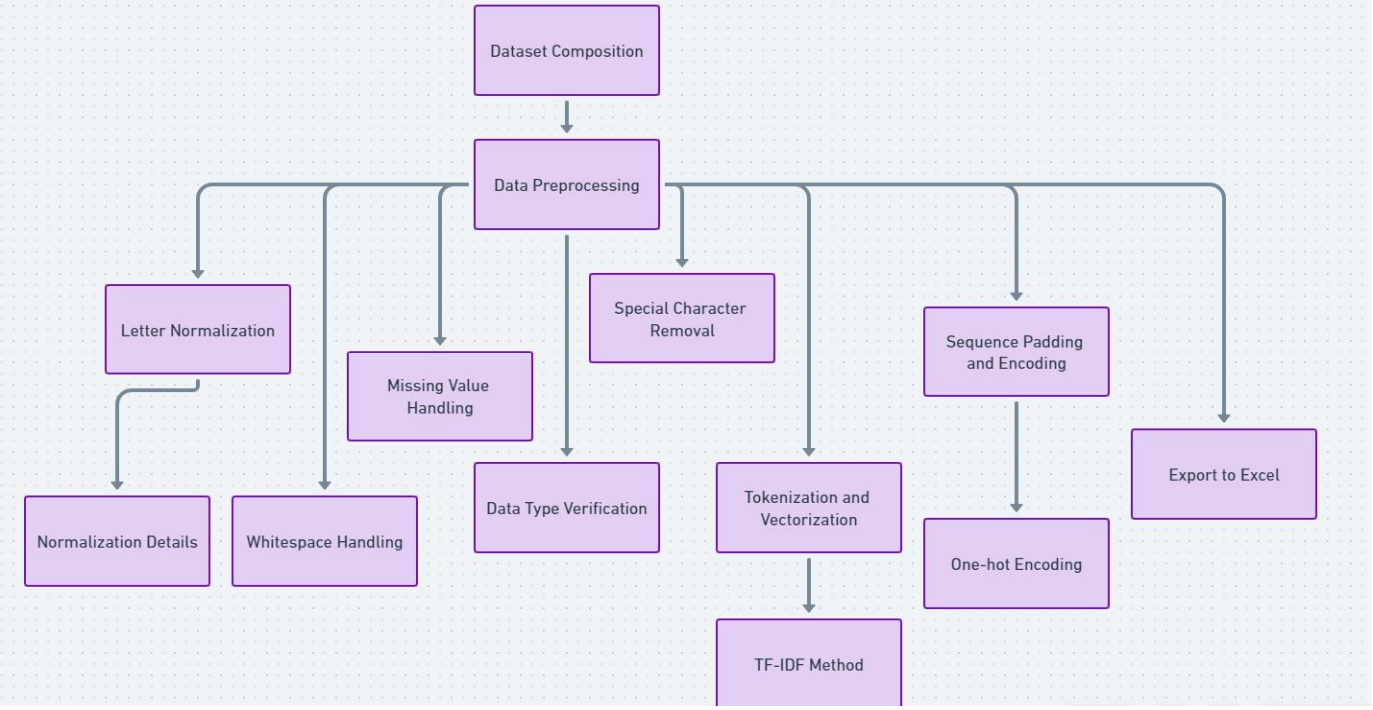


Figure 2 Preprocessing Data

4) *Data Type Verification*: The data type of the text column has been verified and confirmed to be a string (object in Python). If the data is not already in string format, it is converted to ensure that all text data is uniformly represented as strings. This facilitates the application of operations that are specifically designed for text.
5) *Special Character Removal*: We remove special characters, such as punctuation marks and symbols, that could interfere with the tokenization process or model training. Executing this cleaning procedure is crucial to guarantee that the focus remains on the language content rather than on extraneous aspects.
6) *The method of tokenization and vectorization* entails the segmentation of the text into discrete tokens using a specialized tokenizer that is specifically tailored for Arabic language. Following the process of tokenization, the data is converted into a numerical format using the Term Frequency-Inverse Document Frequency (TF-IDF) technique. This enables machine learning algorithms to analyse the text.
7) *Sequence Padding and Encoding*: To simplify modelling tasks, particularly those involving sequences like time series analysis or language modelling, we ensure that all sequences (tokenized texts) have the same length by appending zeros to shorter sequences. Furthermore, labels are encoded using one-hot encoding to align with the anticipated output format of neural network models.
8) *After post-processing*, the processed data is exported to an Excel file, ensuring the long-lasting preservation of the cleaned and processed data for further analysis, or sharing with others, steps shown in Figure. 4.

C. *Data Augmentation Techniques*

Data augmentation techniques can be employed to enhance the robustness and versatility of sentiment analysis systems. These tactics utilize techniques to augment the dataset by generating additional samples derived from the existing ones. This facilitates the acquisition of knowledge by models from a broader spectrum of data. In this article, we offer a thorough elucidation of various augmentation approaches that have been employed on the Arabic sentiment analysis dataset. The process is elucidated in Figure 3, which illustrates Data Augmentation. Figure 3, on the other hand, demonstrates Distribution Augmentation of the data.

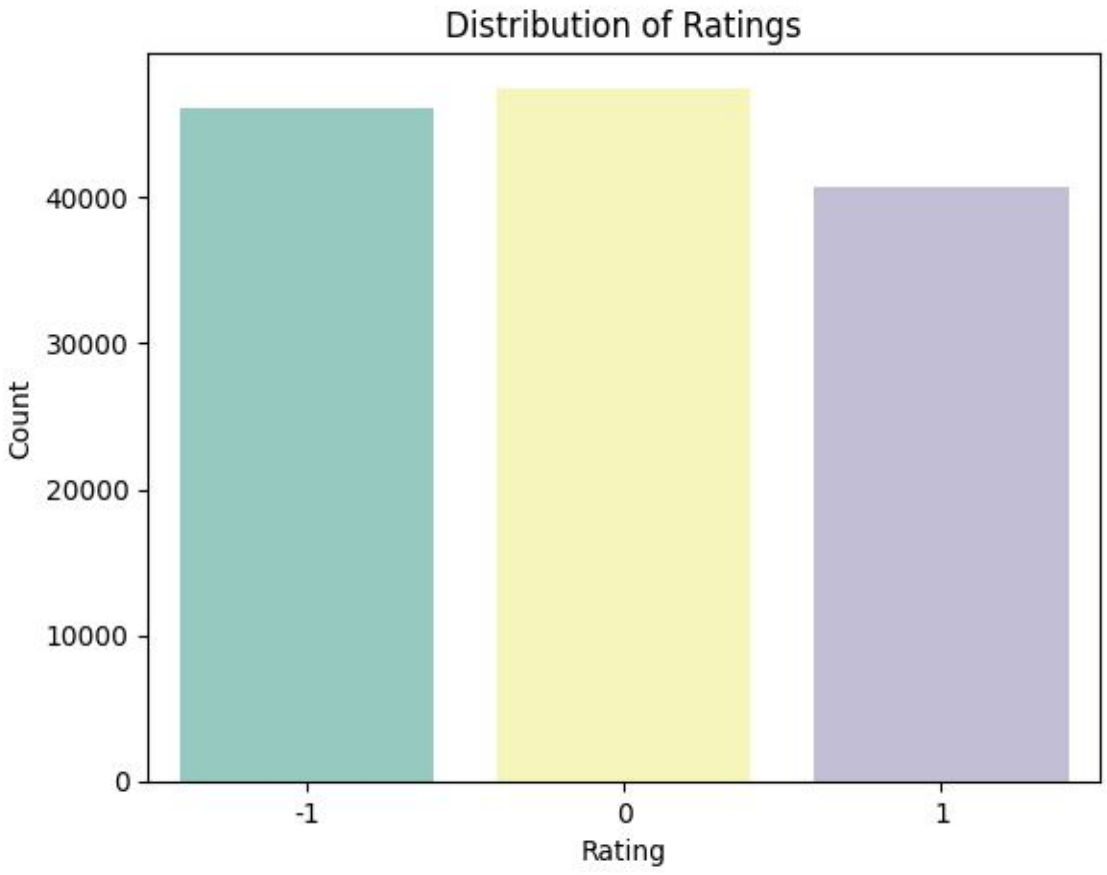


Figure 3 Distribution Augmentation data.

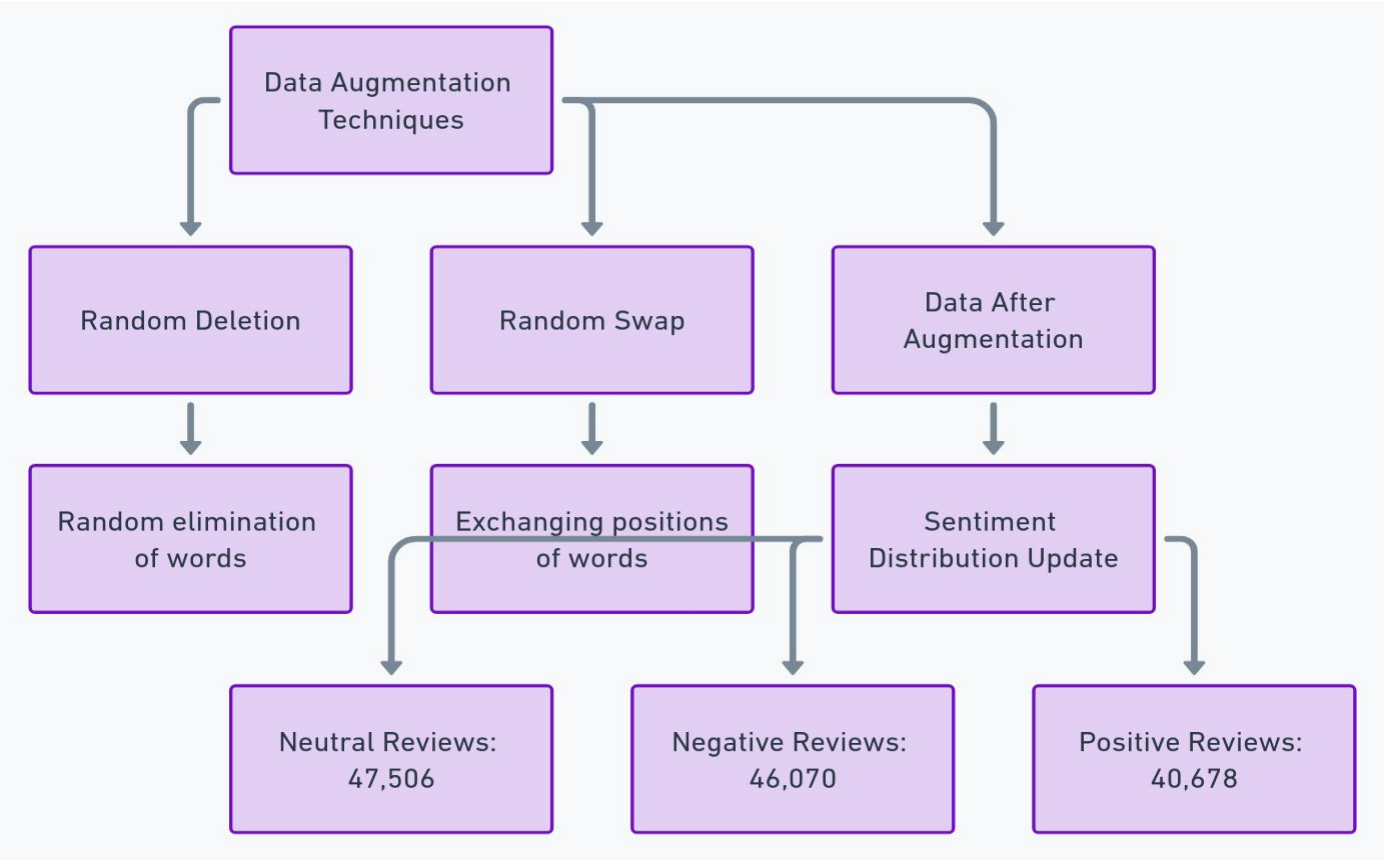


Figure 4 Data Augmentation

1) *Random Deletion* refers to the process of randomly removing elements.
2) Random deletion is a simple and effective technique for data augmentation that involves randomly removing words from a document. This strategy improves the model's capacity to deal with circumstances where words are missing or challenging to detect in the input. The following is the procedure of implementation:
3) *Random Swap* refers to the process of exchanging elements in a random manner.
4) *Random swapping* alters the sentence structure by interchanging the placements of two words inside the phrase, hence diminishing the model's susceptibility to word order.
5) *Data obtained after augmentation in section.*

- The data size grew. Below is a detailed analysis of the distribution of sentiment within the dataset:
- The number of neutral reviews is 47,506.
- There are 46,070 instances of negative reviews with a rating of "-1".
- Positive Reviews ("1"): There are a total of 40,678 instances.

V. EXPERIMENT

A. *Traditional Machine Learning*

This experiment aims to assess the efficacy of conventional machine learning methods in the field of Arabic sentiment analysis. The main goal is to assess the capability of these systems to effectively manage the intricate linguistic components of Arabic text. This will serve as a benchmark for comparing them with more sophisticated models, steps as shown Figure 5.

1) *Application of Traditional Machine Learning Models*

Our research thoroughly evaluates many conventional machine learning models for sentiment analysis in Arabic. These models comprise a range of strategies, each with unique advantages in tasks related to text categorization. The evaluated models include the process is elucidated in Figure 5, illustrating the sequential stages of the Traditional model.

- The Support Vector Machine (SVM) is an effective classifier that excels at managing data with a large number of dimensions. Its resilience is generally acknowledged, making it a popular choice for text classification problems, particularly in the field of sentiment analysis.
- Naive Bayes is a probabilistic classifier that is particularly suitable for applications in natural language processing. It is renowned for its straightforwardness and effectiveness, and it frequently excels in sentiment analysis.
- Logistic Regression is frequently used as a standard for comparison in text categorization jobs. The method is straightforward and effective, producing simply interpretable findings, especially in the context of sentiment analysis.
- Stochastic Gradient Descent (SGD) is a mathematical approach used for optimization. Stochastic Gradient Descent (SGD) is an optimization technique that iteratively adjusts the parameters of a model. It possesses adaptability and can be utilized with different loss functions, rendering it appropriate for diverse categorization assignments.
- Random Forest is an ensemble method that use many decision trees to improve prediction accuracy and reduce overfitting. It is highly beneficial for handling the intricacies associated with sentiment analysis.
- A decision tree is a predictive model that partitions the data into subsets based on the values of particular attributes. Decision trees are useful for understanding the decision-making process, but they may require careful tuning to avoid overfitting in sentiment analysis.

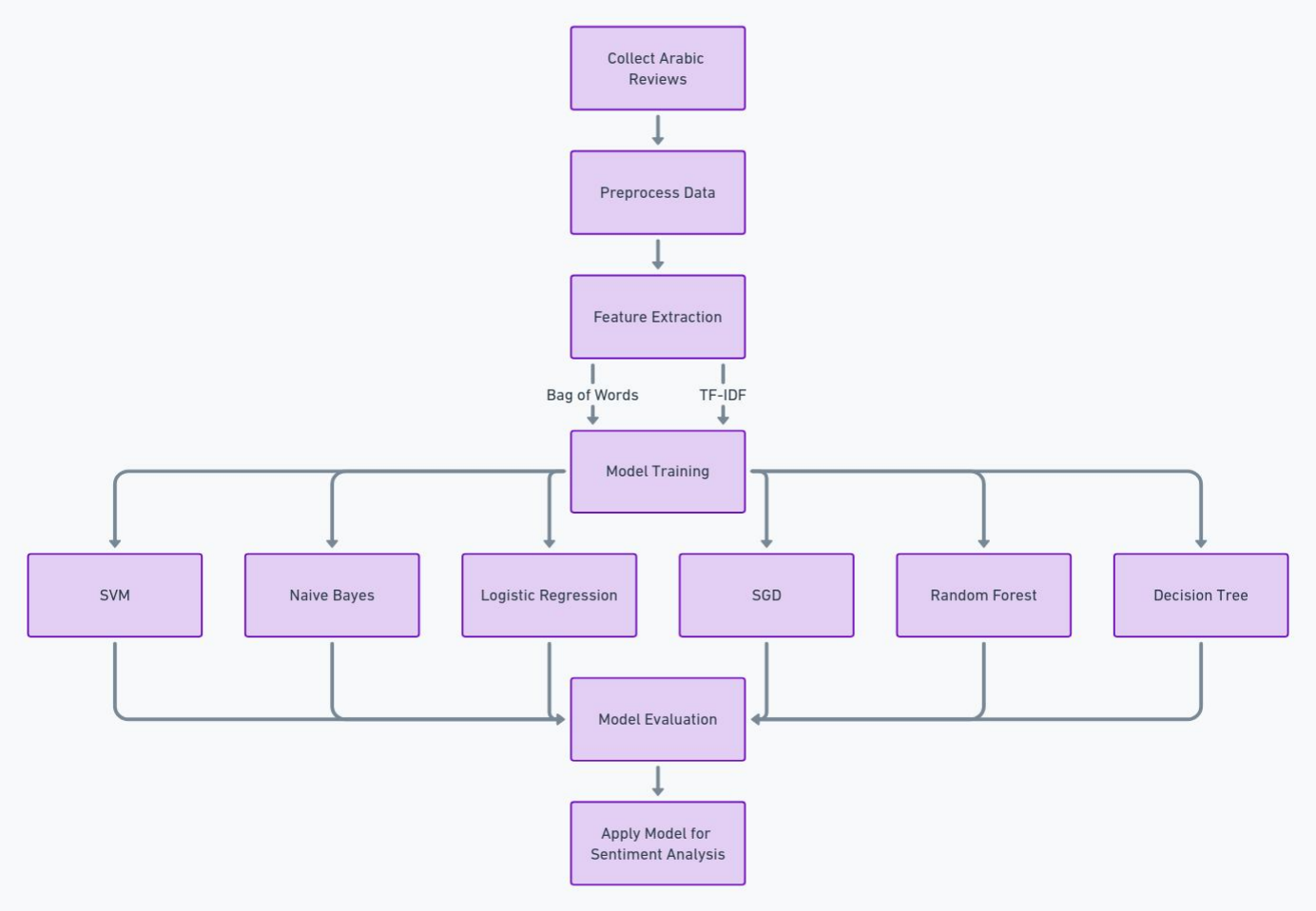


Figure 5 Steps of Traditional model

2) Extraction of Features

The Bag of Words (BoW) methodology is a technique that transforms textual materials into vectors representing the frequencies of words, with a predetermined length. The approach captures the frequency of words in the papers, but it does not consider their sequence or context.

TF-IDF is a statistical metric that quantifies the significance of a phrase in a document relative to a set of documents. The TF-IDF technique measures the significance of a word in a document compared to a larger collection. It helps to counterbalance the prevalence of some terms that have a larger frequency in general.

3) Training and validation of the model

Each model underwent a systematic training and validation process:

- Data Preprocessing: The dataset was subjected to tokenization and vectorization using TF-IDF to convert textual data into a numerical representation suitable for machine learning.
- Dataset Partitioning: The data was divided into separate sets for training, validation, and testing purposes. The models were trained using the training set, the hyperparameters were optimized using the validation set, and the final evaluation of the model's performance was done on the test set.
- Model Training: The models were trained using their respective algorithms, with the main objective of optimizing accuracy and reducing overfitting.
- Validation: The validation set was used during the training process to improve the models and fine-tune the hyperparameters. This ensured that the models could successfully apply their acquired knowledge to fresh and unknown data.

4) Findings

Performance metrics for each model are as follows Table 6 Metrics of traditional model:

TABLE 6
METRICS OF TRADITIONAL MODEL

Model	Accuracy	Precision	Recall	F1-Score
SVM	86.62%	86.76%	86.76%	86.54%
Naive Bayes	84.33%	84.55%	83.85%	84.01%
Logistic Regression	86.88%	86.85%	86.95%	86.80%
SGD	87.08%	87.22%	87.18%	86.99%
Random Forest	80.72%	81.28%	80.70%	80.67%
Decision Tree	68.58%	76.91%	69.80%	68.94%

5) Analysis and Conversation

The examination indicates that linear models outperform ensemble and tree-based models, implying that they are more suitable for sentiment analysis in Arabic because of its textual characteristics. Logistic Regression and Stochastic Gradient Descent (SGD) exhibited remarkable performance, highlighting their effectiveness in handling the intricacies of Arabic sentiment analysis.

B. Experiment 2: Deep Learning Techniques

The aim of this study is to assess the effectiveness of deep learning techniques in the field of Arabic sentiment analysis. The objective is to evaluate the proficiency of deep learning models in effectively capturing the intricate contextual and syntactic aspects of the Arabic language, surpassing traditional machine learning models, Details of steps as shown in Figure 6:

1) Utilized Deep Learning Models

In this study, we employed three sophisticated deep learning models, specifically LSTM (Long Short-Term Memory), CNN (Convolutional Neural Network), and a hybrid LSTM-CNN model, to tackle the task of Arabic sentiment analysis. These models provide the ability to precisely depict the connections between time and space in data, rendering them very appropriate for tasks that involve the examination of text. The process is elucidated in Figure 4, which outlines the phases involved in deep learning.

2) Representation of Features

- The LSTM model utilizes an LSTM layer to process the sequences obtained from the tokenized text input, efficiently capturing the intricate connections within the phrases.

- The CNN model employs convolutional layers to extract spatial features from the tokenized text, allowing for the detection of significant local patterns within the data.
- The Combined LSTM-CNN Model combines the LSTM and CNN layers to use their individual strengths to extract temporal and spatial features, hence enhancing the model's ability to understand complex sentence patterns.

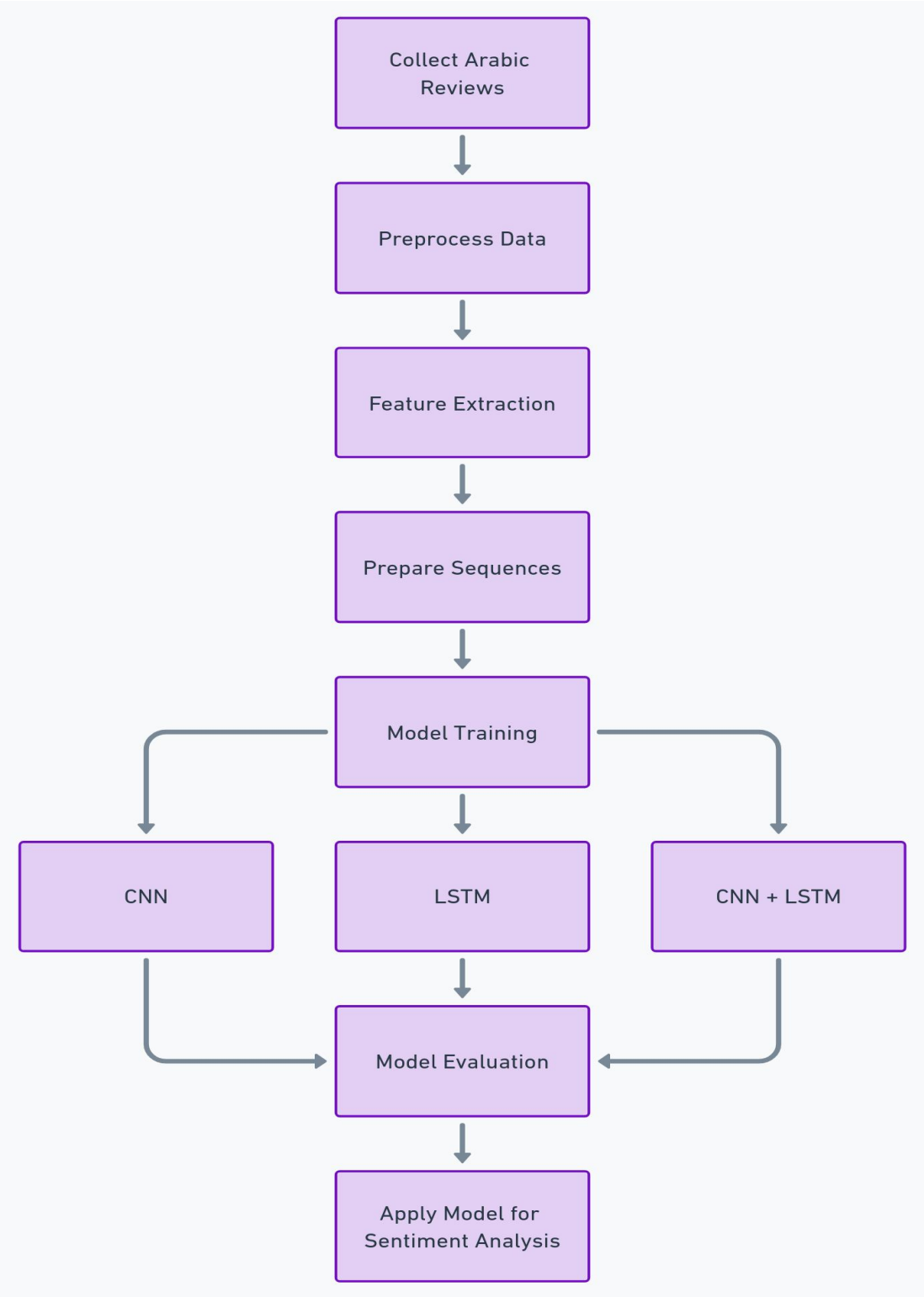


Figure 6 Deep Learning Model

3) Model Training and Validation

- The text data underwent tokenization and padding during the data preparation process, resulting in its transformation into sequences. In addition, TF-IDF vectorization was employed to generate feature representations.
- Partitioning: The dataset was segregated into distinct training and test sets, guaranteeing a resilient assessment framework.
- Training: The models underwent training using the Adam optimizer, and early halting was employed to prevent overfitting.
- Validation: The performance of the model was seen and measured on a separate set of data called the validation set. This allowed for the adjustment and optimization of the model's parameters.

4) Assessment of Performance

- LSTM: Achieved a test accuracy of 77.67%.
- CNN: Exhibited a higher test accuracy of 94.08%.
- Combined LSTM-CNN: Demonstrated a competitive test accuracy of 93.67%, as Showen Table. 7

TABLE 7
TEST ACCURACY FOR DEEP LEARNING

Model	Test Accuracy
LSTM	77.67%
CNN	94.08%
LSTM + CNN	93.67%

5) Findings

The CNN model exhibited superior performance compared to the standalone LSTM, demonstrating the efficacy of convolutional layers in text categorization tasks. Nevertheless, the integrated LSTM-CNN model demonstrated a well-balanced strategy, attaining significant accuracy by utilizing both temporal and spatial characteristics.

6) Analysis and Conversation

The findings indicate that deep learning models, especially those that include multiple architectures, are extremely efficient for sentiment analysis in Arabic. The performance of the combined model highlights the need of utilizing various feature extraction approaches to capture the subtle elements of Arabic text.

C. Experiment 3: Combination Methods

The purpose of integrating Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Bidirectional Encoder Representations from Transformers (BERT), and Ensemble Methods is to augment the model's performance and precision.

The primary objective of integrating Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), Bidirectional Encoder Representations from Transformers (BERT), and ensemble methods is to leverage the unique strengths of each architecture to enhance the overall performance in Arabic sentiment analysis. The following is an examination of the distinct contributions made by each model and the rationale behind creating an ensemble:

Convolutional Neural Networks (CNNs) excel in capturing and analyzing specific details and spatial characteristics within data. They possess the ability to identify patterns and significant signals in sequences of text, which is crucial for the categorization of sentiment.

Contribution: Convolutional Neural Networks (CNNs) provide a robust approach to extract spatial characteristics from textual input by utilizing convolutional layers. These characteristics are essential for discerning patterns that signify sentiment.

LSTM, or Long Short-Term Memory, is adept at understanding and preserving long-term connections, a crucial aspect for accurately capturing sequence context and interpreting sentiment.

Contribution: LSTMs improve the analysis by understanding the order and context of the data, retaining information over longer sections of text, and capturing the time-based relationships in the data.

BERT is a language model acronym for Bidirectional Encoder Representations from Transformers.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

The primary advantage of BERT is in its capacity to understand the complex nuances of the Arabic language through its bidirectional contextual understanding and its pre-trained language model foundation. This feature significantly enhances its predictive precision.

BERT improves sentiment analysis by including a deep contextual understanding, capturing subtle language cues and semantic relationships, hence expanding the usable feature set.

Ensemble Approach Reasoning: The ensemble technique aims to combine the predictive powers of CNN, LSTM, and BERT, overcoming the limits of individual models while capitalizing on the strengths of each. This approach exhibits a proclivity to enhance the overall accuracy, mitigate overfitting, and enhance generalization.

The ensemble technique reduces variance by aggregating predictions from multiple models, incorporating diverse perspectives on the data, and yielding a more equitable and resilient sentiment classification.

Objective of the Integration

The integration of these models into an ensemble framework aims to develop a comprehensive system that:

The ensemble model aims to optimize accuracy by combining multiple analytical perspectives, surpassing the precision achievable by any single model.

Enhances resilience: The ensemble method reduces the likelihood of overfitting by aggregating multiple models and mitigating biases and errors, leading to more reliable predictions.

Harnesses the benefits of related skills: Each model has unique attributes, which, when integrated, provide a thorough examination that incorporates both the specific and broader contexts found in the text.

The objective of this merger is to create a sophisticated system for assessing sentiment in the Arabic language. This initiative aims to set a new industry benchmark by employing sophisticated deep learning architectures.

1) Utilized Deep Learning Models

- The BERT model, short for Bidirectional Encoder Representations from Transformers, has significantly transformed the field of natural language processing (NLP) by utilizing a deep bidirectional representation. This enables the model to comprehend the context more effectively compared to earlier models. The processes involved in this approach are elucidated in Figure 5, titled "Ensemble Approach Steps."
- The CNN model, originally designed for image processing, has been effectively applied to NLP problems by collecting local feature correlations in text.
- LSTM Model: Long Short-Term Memory Model Memory networks are a specific sort of recurrent neural network (RNN) that excel at retaining and utilizing long-term connections between input points. This makes them particularly well-suited for processing sequential data, such as text.
- Ensemble Method: Integrating BERT, CNN, and LSTM predictions to construct a resilient ensemble model.

2) Feature representation

- The BERT model utilizes a pre-trained tokenizer to convert text into tokens, efficiently capturing contextual relationships.
- The CNN and LSTM models employ a Keras tokenizer to transform text into sequences, which are then padded to ensure uniform length.

3) Model Training and Validation

- The BERT model underwent training using a DataLoader, employing the AdamW optimizer, and the training loss was tracked for every epoch.
- The Keras Sequential models were employed to create embedding layers, followed by convolutional and LSTM layers, to make use of CNN and LSTM.
- The training approach consisted of fitting models using training data, implementing appropriate callbacks, and utilizing early stopping based on validation loss to address overfitting.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

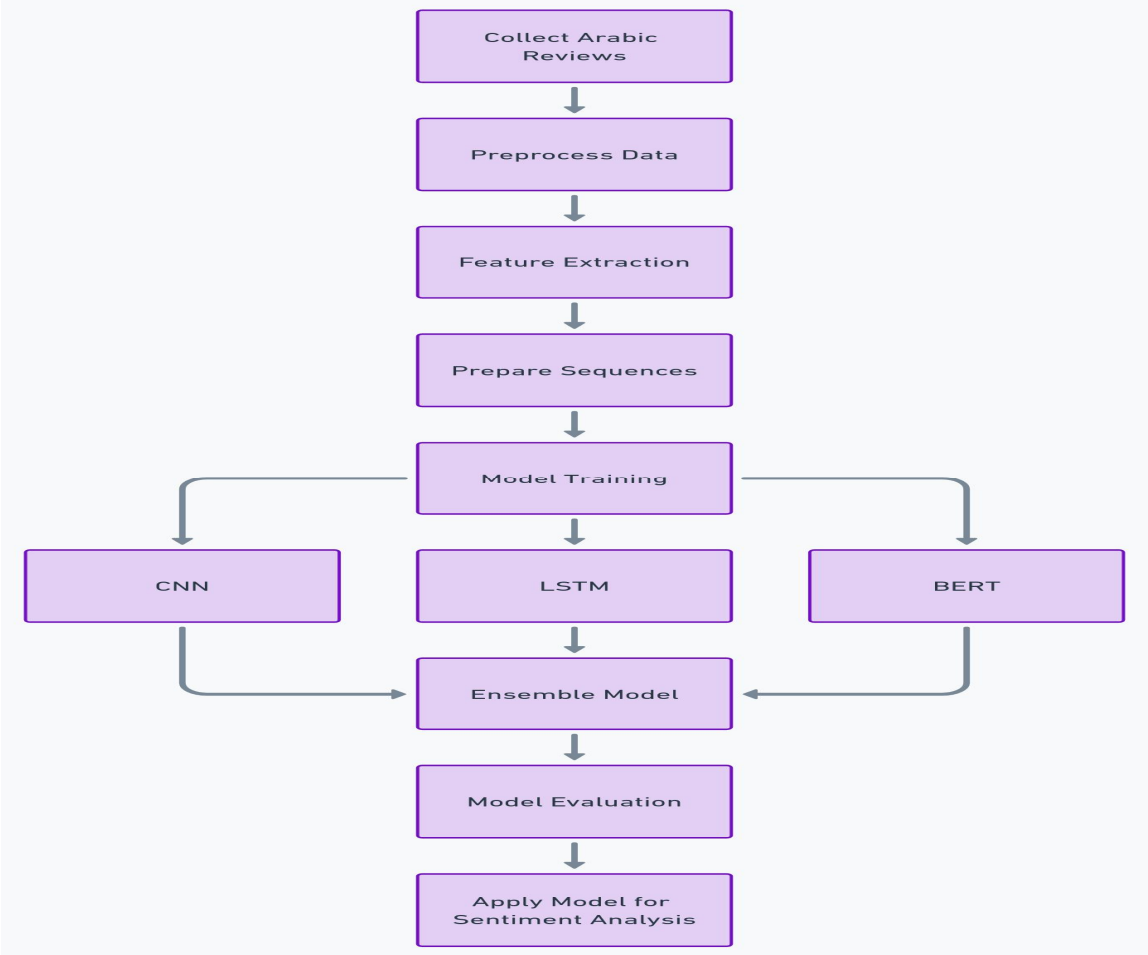


Figure 7 Ensemble Approach model.

4) Assessment of Performance

The performance was evaluated based on accuracy, comparing how each model fared on the test data.

5) Findings

- The BERT model demonstrated favorable outcomes owing to its ability to comprehend context.
- CNN demonstrated proficiency in capturing spatial correlations within textual data.
- LSTM has shown aptitude in comprehending extended relationships.
- The ensemble technique sought to exploit the advantages of each individual model, perhaps providing enhanced accuracy compared to single-model approaches, Details shown in Table. 8
- 7.3.5.4 Combine LSTM and CNN demonstrated proficiency explain in this Figure 8 Combine LSTM and CNN accuracy and loss. As shown in Figure. 8.

TABLE 8
ACCURACY FOR ENSEMBLE APPROACH

Model	Accuracy
LSTM	77.67%
CNN	94.08%
BERT	95.72%
CNN + LSTM	93.67%
ENSEMBLE	95.80%

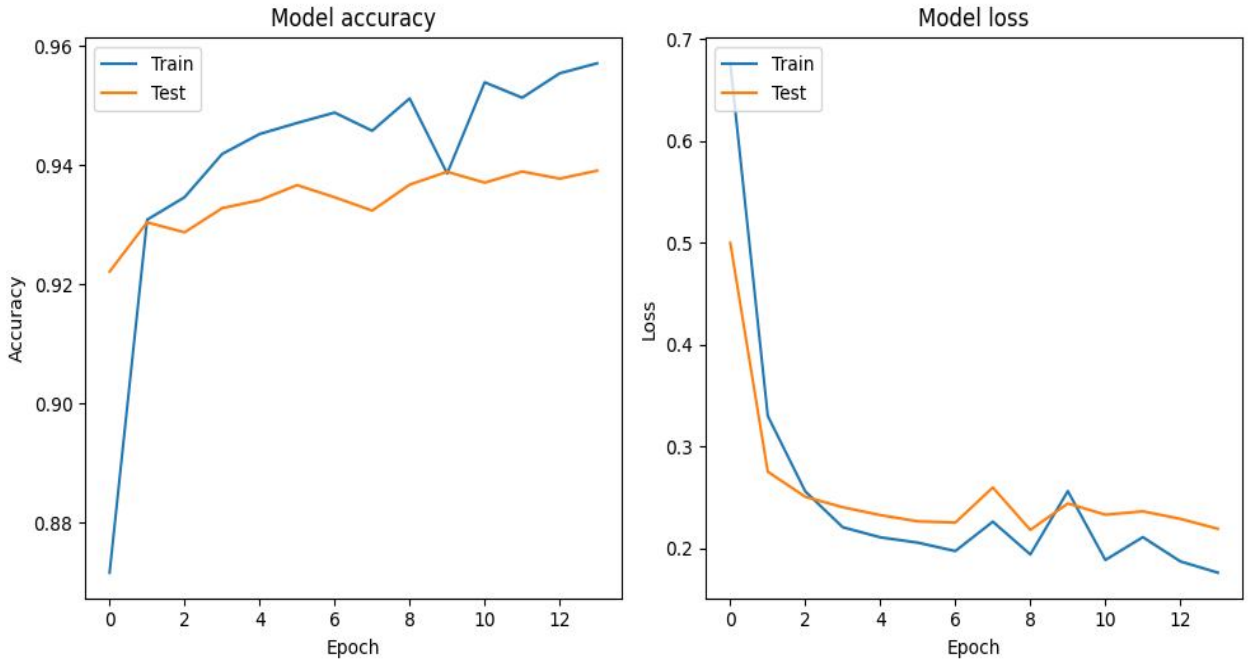


Figure 8 Combine LSTM and CNN accuracy and Loss

6) Analysis and Conversation

Each model's unique structure allows them to excel in different aspects of text comprehension, demonstrating the diverse array of approaches in natural language processing (NLP). The ensemble methodology suggests a strategic method to enhance the performance of a model by integrating many analytical capabilities.

VI. CONCLUSION

This comprehensive study conducted an in-depth examination of various machine learning and deep learning methods to address the complex task of sentiment analysis in Arabic text. Our objective was to utilize classical algorithms, advanced neural networks, and sophisticated ensemble techniques to address the inherent difficulty of Arabic sentiment analysis.

1) Machine Learning techniques: We investigated various traditional machine learning methods, including Support Vector Machines (SVM), Decision Trees, Logistic Regression, Random Forest, and Naive Bayes. Each method provided a unique perspective on sentiment categorization, giving a solid foundation for understanding and analyzing textual sentiment.

2) Our approach prioritized the importance of meticulous preparation and the augmentation of data in the field of Arabic natural language processing (NLP). The stages mentioned are crucial for standardizing the data, addressing the special challenges of the language, and diversifying our training sets. This will enhance the ability of our models to make generalizations.

3) The examination of Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTMs), and Arabic Bidirectional Encoder Representations from Transformers (BERT) has uncovered the vast potential of deep learning in precisely capturing the intricate semantic and contextual complexities of the Arabic language. Due to their ability to gain hierarchical representations, the models outperformed standard machine learning methods in our sentiment analysis tasks.

4) Ensemble Methods: Our objective was to improve the precision and reliability of sentiment categorization by merging the predictive capabilities of different models through ensemble approaches. This strategy enables us to use the advantages of each model while mitigating their distinct shortcomings.

5) By utilizing word embedding approaches like Word2Vec and GloVe, together with meticulous tokenization, we were able to obtain concise and significant word representations, hence enhancing the efficiency of our models.

6) Feature extraction methods, such as TF-IDF and Bag of Words, have been demonstrated to be valuable in turning text into numerical representations. This enabled our models to rapidly analyze and classify textual material.

7) In summary, this paper presents a comprehensive methodology for assessing sentiment in the Arabic language. It amalgamates the benefits of numerous machine learning and deep learning techniques.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Our research suggests that while individual models offer valuable insights and predictive capabilities, employing an ensemble method that integrates the advantages of various models yields a more comprehensive, accurate, and dependable solution for sentiment analysis in Arabic. This study not only improves our understanding of Arabic sentiment analysis but also sets a benchmark for future research in this area, enabling the creation of more sophisticated and nuanced NLP tools tailored specifically for the Arabic language.

VII. FUTURE WORK

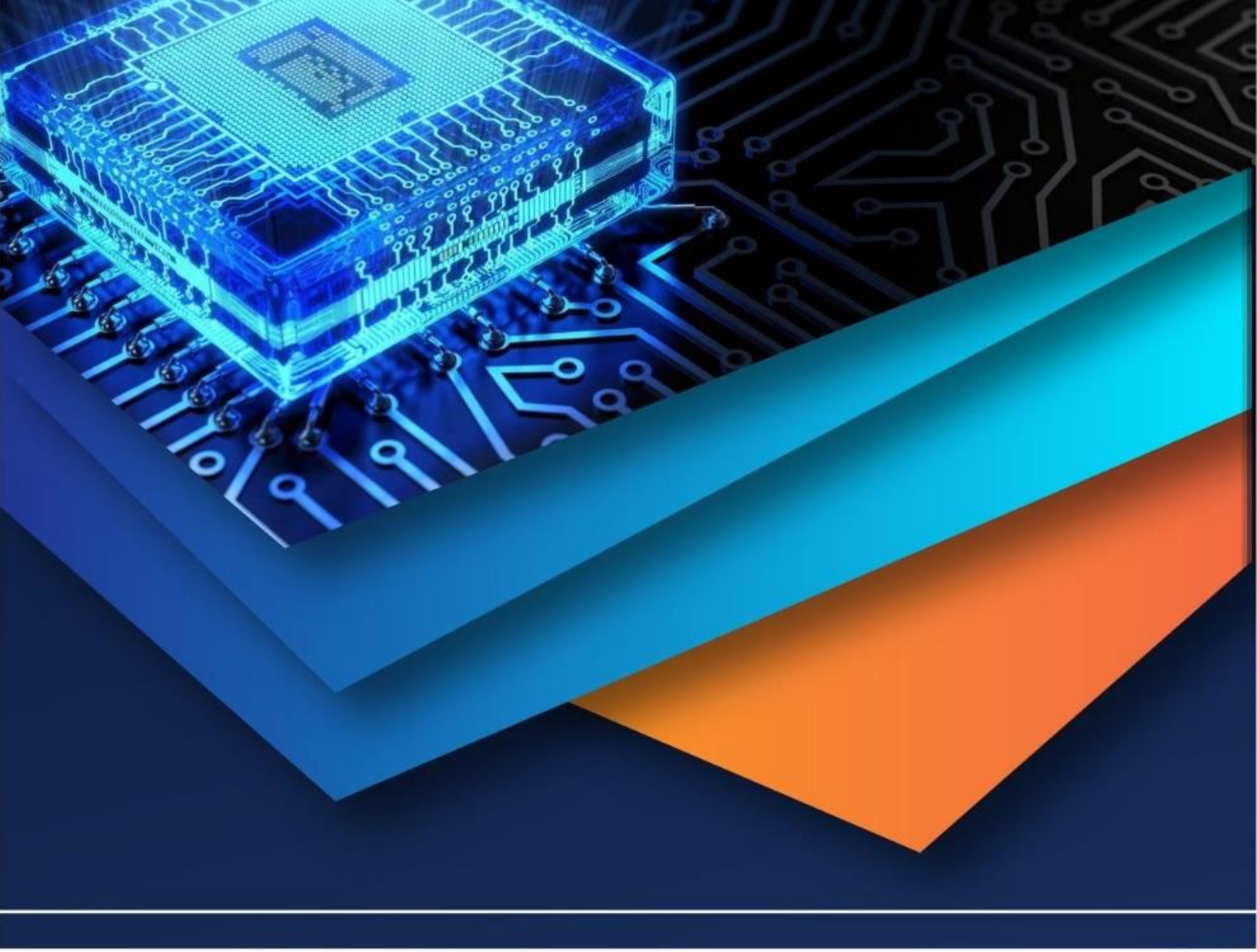
This study has made significant advancements in Arabic sentiment analysis by integrating a variety of machine learning and deep learning techniques. Nevertheless, there exist other domains for prospective investigation that can enhance the understanding and analysis of sentiment in Arabic literature.

1) Future study could explore the development of cross-linguistic models that employ transfer learning to adapt models trained on abundant data in one language to languages with fewer resources, such as Arabic dialects. This methodology has the ability to leverage the fundamental attributes of sentiment that exist in all languages, hence enhancing the precision and efficacy of sentiment analysis models specially designed for the Arabic language.
2) Using Arabic Dialects: Given the wide variety of Arabic dialects, it is essential for future research to concentrate on creating models that have proficiency not only in Modern Standard Arabic but also in understanding various regional dialects. This work involves collecting and annotating a diverse range of dialectical text data, with the aim of developing models that are both comprehensive and precise in reflecting the Arabic-speaking world.
3) Engaging in the study of sophisticated deep learning structures, such as transformers and capsule networks, could perhaps provide a greater understanding. The adaptability of these models in capturing complex patterns and their ability to improve existing methods in Arabic sentiment analysis warrant thorough research.
4) The significance of Explainable AI in Sentiment Analysis becomes paramount when the intricacy of AI models escalates. Future research should prioritize the creation of explainable AI frameworks, which aim to provide a more comprehensive understanding of the decision-making process of sentiment analysis algorithms. Implementing this would improve the reliability and understanding of sentiment forecasts produced by AI systems.
5) Developing models capable of conducting real-time sentiment analysis would offer immediate insights into public sentiments and trends on social media sites. Subsequent investigations should prioritize the enhancement of models' efficiency without compromising their precision, a challenging yet crucial objective.
6) Exploring multimodal sentiment analysis through the integration of audio, video, and textual data has the capacity to offer a holistic understanding of sentiments. Implementing this versatile approach has the capacity to significantly improve areas such as customer service, social media monitoring, and sentiment analysis.
7) Domain-specific models tailored for industries like banking, healthcare, or tourism have the potential to provide more precise sentiment analysis. Future initiatives may involve developing specialized models capable of understanding the unique context and terminology used in various areas. This would allow for the delivery of emotional observations that are more concentrated and relevant.
8) By investigating these potential areas of study, future research could advance the field of Arabic sentiment analysis, resulting in more sophisticated, accurate, and contextually aware technologies that can cater to the diverse needs of users and industries working with Arabic text data.

REFERENCES

[1] V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273--297, 1995.
[2] J. R. Quinlan, "Induction of decision trees," Machine learning, vol. 1, pp. 81--106, 1986.
[3] R.-E. a. C. K.-W. a. H. C.-J. a. W. X.-R. a. L. C.-J. Fan, "LIBLINEAR: A library for large linear classification," the Journal of machine Learning research, vol. 9, pp. 1871--1874, 2008.
[4] D. W. Hosmer, "Applied logistic regression,," John Wiley \& Sons
[5] L. Breiman, "Random forests," Machine learning, vol. 45, pp. 5--32, 2001.
[6] I. a. o. Rish, "An empirical study of the naive Bayes classifier," in IJCAI 2001 workshop on empirical methods in artificial intelligence, 2001, pp. 41--46.
[7] Y. Chen, Convolutional neural network for sentence classification, University of Waterloo, 2015.
[8] S. a. S. J. Hochreiter, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735--1780, 1997.
[9] J. a. C. M.-W. a. L. K. a. T. K. Devlin, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.0480, 2018.
[10] T. a. C. K. a. C. G. a. D. J. Mikolov, "Efficient estimation of word representations in vector space," arXiv preprint arXiv:1301.3781, 2013.

[11] J. a. o. Ramos, "Using tf-idf to determine word relevance in document queries," in Proceedings of the first instructional conference on machine learning, Citeseer, 2003, pp. 29--48.
[12] S. Basabain, "A survey of Arabic thematic sentiment analysis based on topic modeling," International Journal of Computer Science \& Network Security, vol. 21, no. 9, pp. 155--162, 2021.
[13] A. C. a. D. A. Mazari, "Deep learning-based sentiment analysis of algerian dialect during Hirak 2019," in 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH), IEEE, 2021, pp. 233--236.
[14] S. a. f. A. D. tweets, "Sentiment analysis for Algerian Dialect tweets," in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE, 2022, pp. 235--239.
[15] A. a. A. F. a. A. H. Abdelwahab, "Enhancing the performance of sentiment analysis supervised learning using sentiments keywords based technique," in CS \& IT Conference Proceedings, CS \& IT Conference Proceedings, 2017.
[16] A. a. A. N. S. a. A. N. a. A. Z. Alazba, "Saudi Stock Market Sentiment Analysis using Twitter Data," in KDIR, 2020, pp. 36--47.
[17] N. Al-Twairesh, "The evolution of language models applied to emotion analysis of Arabic tweets," Information, vol. 12, no. 2, p. 84, 2021.
[18] N. a. K. S. a. H. N. Hicham, "Customer sentiment analysis for Arabic social media using a novel ensemble machine learning approach," International Journal of Electrical and Computer Engineering, vol. 13, no. 4, pp. 4504--4515, 2023.
[19] A. a. R. A. Alqarni, "Arabic Tweets-based Sentiment Analysis to investigate the impact of COVID-19 in KSA: A deep learning approach}," Big Data and Cognitive Computing, vol. 7, no. 1, p. 16, 2023.
[20] H. a. M. S. a. A. A. a. E.-S. S. a. A. T. Saleh, "Heterogeneous ensemble deep learning model for enhanced Arabic sentiment analysis," Sensors, vol. 22, no. 10, p. 3707, 2022
[21] M. a. N. J. A. Mhamed, "Enhancing Arabic Sentiment Analysis Through a Hybrid Deep Learning Approach".
[22] N. a. V. G. a. A. R. a. G. M. a. D. K. a. A. H. a. E.-A. M. A. a. A.-T. B. N. a. A. F. a. H. A. Elhassan, "Arabic Sentiment Analysis Based on Word Embeddings and Deep Learning," Computers, vol. 12, no. 6, p. 126, 2023.
[23] A. H. a. O. W. a. A. A. M. Ombabi, "Deep learning CNN--LSTM framework for Arabic sentiment analysis using textual information shared in social networks," Social Network Analysis and Mining, vol. 10, pp. 1--13, 2020.
[24] S. a. A. W. A. a. A. L. a. A. L. a. A. A. a. A. R. N. a. A. N. a. A. H. a. A. W. Alhumoud, "ASAVACT: Arabic sentiment analysis for vaccine-related COVID-19 tweets using deep learning," PeerJ Computer Science, vol. 9, p. e1507, 2023.
[25] M. E. M. M. Manal Essam, "Arabic Tweets Sentiment Analysis using Hybrid," International Journal of Computer Applications, vol. 175, no. 36, p. 0975 – 8887, 2020.
[26] I. a. A. F. a. V. A. Guellil, "English vs arabic sentiment analysis: A survey presenting 100 work studies, resources and tools," in 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA), IEEE, 2019, pp. 1--8.
[27] N. a. M. S. a. A.-A. M. a. A.-K. M. a. o. Abdulla, "Automatic lexicon construction for arabic sentiment analysis," in 2014 International Conference on Future Internet of Things and Cloud, IEEE, 2014, pp. 547--552.
[28] S. a. A. T. a. A. M. Alhumoud, "Arabic sentiment analysis using WEKA a hybrid learning approach," in 2015 7th international joint conference on knowledge discovery, knowledge engineering and knowledge management (IC3K), IEEE, 2015, pp. 402--408.
[29] S. a. A. Z. a. A. E. a. A. I. Abuuznien, "Sentiment analysis for Sudanese Arabic dialect using comparative supervised learning approach," in 2020 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), IEEE, 2021, pp. 1--6.
[30] H. a. Q. R. a. L. D. Al-Rubaiee, "Identifying Mubasher software products through sentiment analysis of Arabic tweets," in 2016 International Conference on Industrial Informatics and Computer Systems (CIICS), IEEE, 2016, pp. 1--6.
[31] I. a. S. H. a. A. F. a. G. B. a. N. D. Guellil, "Arabic natural language processing: An overview," Journal of King Saud University-Computer and Information Sciences, vol. 33, no. 5, pp. 497--507, 2021.
[32] N. a. V. S. a. D. K. Habash, "Proceedings of the Second Workshop on Arabic Natural Language Processing," in Proceedings of the Second Workshop on Arabic Natural Language Processing, 2015.
[33] B. Elayeb, "Arabic word sense disambiguation: a review," Artificial Intelligence Review, vol. 52, no. 4, pp. 2475--2532, 2019.
[34] "Revisiting pre-trained language models and their evaluation for arabic natural language understanding},," arXiv preprint arXiv:2205.10687, 2022.
[35] A. a. Y. S. a. M. Y. a. L. L. a. F. S. Elnagar, "A benchmark for evaluating arabic contextualized word embedding models," Information Processing \& Management, vol. 60, no. 5, p. 103452, 2023.
[36] S. M. A. H. a. S. S. F. H. a. U. A. a. R. A. a. A. G. a. A. M. Shah, "Arabic Sentiment Analysis and Sarcasm Detection using Probabilistic Projections Based Variational Switch Transformer," IEEE Access, 2023.
[37] A. a. A. S. S. a. M. A. Rahma, "A comprehensive review on arabic sarcasm detection: Approaches, challenges and future trends," IEEE Access, 2023.
[38] A. a. E. M. A. a. E. K. a. E. M. N. a. B. I. a. K. A. El Mahdaouy, Deep multi-task model for sarcasm detection and sentiment analysis in Arabic language. 2021.
[39] L. a. A. K. a. A. F. Alhaidari, "Detecting irony in arabic microblogs using deep convolutional neural networks," International Journal of Advanced Computer Science and Applications, vol. 13, no. 1, 2022



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)