ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 12, Issue 2, pp: (40-50), Month: April - June 2024, Available at: www.researchpublish.com

A NOVEL ARABIC STEMMING ALGORITHM WITH WORDNET INTEGRATION FOR IMPROVED ESSAY SCORING

Mohamad Alobed[1], Yazeed Al Moaiad[2]

[1,2]Al-Madinah International University, Kuala Lumpur, Malaysia

DOI: https://doi.org/10.5281/zenodo.12516055

Published Date: 24-June-2024

Abstract: **Automated Arabic Essay Scoring (AAES) systems have the potential to revolutionize educational assessment in the Arabic-speaking world, but their accuracy is often hindered by the complexities of Arabic morphology and semantics. This study introduces a novel AAES model, the Arabic Stemming Enhancement (ASE) algorithm, which integrates Arabic WordNet (AWN) to enhance semantic understanding and root extraction accuracy. By leveraging AWN's lexical database, ASE overcomes limitations of traditional stemming approaches, such as misinterpreting polysemous words. Rigorous evaluation on student essays and a WordNet-based dataset demonstrates ASE's superior performance compared to existing algorithms. Notably, ASE achieved a 0.977 Pearson correlation with human-assigned scores and a 15% increase in accuracy, alongside improvements in precision, recall, and F1-score. This research significantly advances Arabic Natural Language Processing (NLP), offering a more accurate, equitable, and scalable automated assessment tool for Arabic language education.**

Keywords: **Arabic Stemming Enhancement, Automated Essay Scoring, Arabic WordNet, Natural Language Processing, Semantic Analysis, Root Extraction, Educational Assessment, Machine Learning.**

1. INTRODUCTION

Automated Essay Scoring (AAES) systems hold immense potential to revolutionize educational assessment in the Arabic-speaking world by enhancing efficiency, objectivity, and fairness. However, the inherent complexities of the Arabic language, particularly its rich morphology and semantic nuances, pose significant challenges to the development of accurate and reliable AAES systems [9],[29]. One of the key challenges lies in the development of robust stemming algorithms, crucial for reducing words to their root forms and facilitating accurate text analysis [4].

Stemming in Arabic is particularly challenging due to the language's complex morphology, where words are often composed of a root embedded with prefixes, suffixes, and infixes [16]. These affixes can significantly alter the meaning of the root word, making accurate root extraction essential for understanding the context and meaning of words. As illustrated in Fig. 1, the stemming process involves two key steps: segmentation, where prefixes and suffixes are removed, and root extraction, where the remaining stem is analyzed to find the root [7].

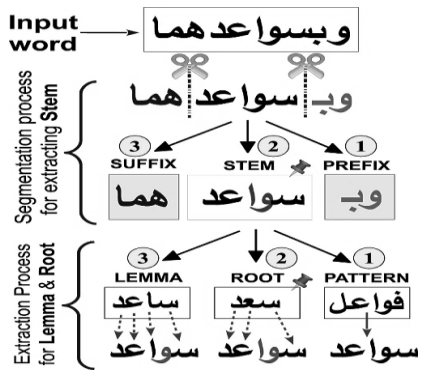

Fig. 1: Root, Stem, and Lemma Extraction from an Arabic Word

Existing Arabic stemming algorithms often struggle with this complexity, leading to inaccuracies that can negatively impact the performance of AAES systems. To address these challenges, this study proposes a novel AAES model that leverages Arabic WordNet (AWN) to enhance semantic understanding and stemming precision. This approach is novel in that it integrates AWN's extensive lexical database with a sophisticated stemming algorithm, enabling the model to accurately extract root words while preserving crucial contextual nuances. By overcoming the limitations of traditional stemming approaches, which often misinterpret word meanings due to Arabic's morphological richness and polysemy, the proposed model is expected to significantly improve the accuracy and fairness of Arabic essay evaluations.

This paper presents a comprehensive evaluation of the proposed model, comparing its performance to existing stemming algorithms on a diverse dataset of Arabic essays. The results demonstrate the model's superior accuracy and effectiveness in capturing the semantic nuances of Arabic text. This research contributes to the growing body of knowledge in Arabic NLP and AES, paving the way for more effective and equitable educational assessment tools in the Arabic-speaking world.

2. BACKGROUND OF THE STUDY

The development of reliable and accurate stemming algorithms for Arabic has the potential to significantly improve NLP applications beyond Automated Essay Scoring (AES). Areas such as information retrieval, machine translation, and text summarization would all benefit from an enhanced understanding of Arabic text.

2.1 Arabic Language Overview

Arabic, a major world language spoken by over 422 million people, holds significance for over 1.5 billion Muslims globally. Despite its importance, automated Arabic language processing poses numerous challenges due to its complex morphology [24]. Unlike English, Arabic is considered a low-resource language by the Association for Computational Linguistics (ACL), meaning fewer tools and datasets are readily available [1]. This scarcity presents unique difficulties for researchers in Arabic Natural Language Processing (NLP) [5].

Arabic is renowned for its linguistic richness, characterized by numerous synonyms and diverse word forms. The Arabic alphabet consists of 28 characters, and its script is written from right to left, with letters changing shape based on their position within a word. Arabic words are categorized into three main types: nouns (including adjectives and adverbs), verbs, and particles [30]. The root, consisting of core letters that form a word, is fundamental to Arabic morphology, and multiple words can share the same root. For example, "لاعب" (player), "لعبة" (game), and "ملعب" (stadium) all stem from "لعب" (to play) [31].

To identify a word's root, knowledge of morphological templates is essential. Arabic morphology is complex, with roots typically consisting of three consonants but occasionally having four or five. Despite the vast number of potential roots, only a fraction is actively used, adding to the complexity of language processing [12].

2.2 Challenges in Processing the Arabic Language

The intricacies of the Arabic language extend beyond its morphological richness. Its complex orthography, characterized by the absence of diacritical marks in most texts, creates ambiguity as words with identical spellings can have different meanings depending on the context [19]. This ambiguity, coupled with the significant dialectal variations across different regions, further complicates the development of NLP tools that can work universally across all Arabic variations [41]. Moreover, inconsistencies in spelling, particularly in online content, pose additional challenges for accurate text processing [15]. The scarcity of resources, such as annotated datasets and advanced NLP tools specifically designed for Arabic, further exacerbates these challenges [14].

2.3 WordNet: A Promising Solution

WordNet, a lexical database that organizes words into sets of synonyms called synsets, each representing a distinct concept, offers a potential solution to these challenges. These synsets are interconnected through various semantic relationships, providing a rich source of knowledge about word meanings and their connections [17]. By incorporating this semantic information, Arabic WordNet (AWN) can address the limitations of existing stemming algorithms for Arabic text. These limitations, such as over-stemming, under-stemming, and inaccurate handling of polysemous words, can negatively impact the accuracy of AAES systems, leading to inaccurate and unfair assessments of students' writing skills.

2.4 Research Gap and Proposed Solution

Despite the potential of WordNet, there is a lack of research exploring its integration with robust stemming techniques for Arabic. This study aims to bridge this gap by proposing a novel AAES model that leverages AWN to enhance semantic understanding and stemming precision. By incorporating AWN's extensive lexical database with a sophisticated stemming algorithm, this model aims to overcome the limitations of traditional stemming approaches, which often misinterpret word meanings due to Arabic's morphological richness and polysemy. The proposed model's ability to accurately extract root words while preserving contextual nuances is expected to significantly improve the accuracy and fairness of Arabic essay evaluations.

2.5 Automated Essay Scoring (AES): State of the Art and the Arabic Language

Automated Essay Scoring (AES) systems hold great potential to revolutionize educational assessment in the Arabic-speaking world due to their efficiency, objectivity, and scalability [40]. However, the morphological complexity and rich vocabulary of the Arabic language pose significant challenges to developing accurate and reliable AAES systems. Notably, current approaches struggle with the accurate analysis and stemming of Arabic words, which is crucial for understanding the semantic meaning of essays [9],[29].

2.6 Stemming in Arabic Natural Language Processing (NLP)

Stemming, a fundamental process in Natural Language Processing (NLP), involves reducing words to their root forms, which is crucial for various NLP applications, including information retrieval, text summarization, sentiment analysis, and AES. By converting words to their root forms, stemming allows these systems to match different forms of the same word, thereby improving their performance [26].

However, stemming in Arabic NLP presents unique challenges due to the language's complex morphology, where words are often composed of a root embedded with prefixes, suffixes, and infixes [16]. These affixes can alter the meaning of the root word, adding layers of complexity to the stemming process. Moreover, accurate root extraction is essential in Arabic to understand the context and meaning of words [36]. Misinterpretation of the root can lead to a significant change in the intended meaning, making the task of Arabic stemming a complex yet vital aspect of Arabic NLP.

2.7 Existing Arabic Stemming Algorithms

Various stemming algorithms have been developed to address the specific needs of Arabic NLP, each with its own methodologies and limitations as in Table 1:

Table 1: Summarizing the Arabic Stemming Algorithms

Algorithm	Description	Limitations
Light Stemmer	Light stemmers are algorithms designed to extract prefixes and suffixes from given terms [2].	• Limited affix removal can lead to over-stemming or under-stemming [21]. • Lack of lexicon validation increases risk of generating invalid roots [34].
Extended-Light Stemmer	The Extended-Light stemmer is designed to include a broader range of prefixes and suffixes, beyond those recognized by the Light10 stemmer [21].	• No lexicon validation (potential for inaccurate roots) [22].
ISRI Stemmer	The ISRI stemmer is an Arabic stemming algorithm that normalizes words, removes prefixes and suffixes, and maps the remaining words to patterns based on their length [39].	• No lexicon validation, potential for errors and invalid roots [39]
Tashaphyne	Tashaphyne operates as a light stemmer that is designed to eliminate diacritics and normalize inputs [7].	• Relies on affix lists only (no dictionary), potentially impacting accuracy [33].
Arlstem	ARLStem is programmed to remove a specific list of prefixes and a designated set of suffixes [3].	• Limited lexicon, over-stemming tendency, inconsistent results, error-prone with irregular forms [34].

Buckwalter	The Buckwalter stemmer uses stem tables that include prefixes, likely stems, and suffixes [32].	• No lexicon validation, potential for errors and invalid roots [33].
Motaz	Motaz has developed an Arabic light stemmer that would eliminate prefixes and suffixes from the word [35].	• Does not handle infixes [8].
Darwish	The Farasa stemmer, suggested by Darwish, segments the word into a series of prefixes, a stem, and a stem of suffixes [10].	• Inaccurate stemming of foreign names and long, complex words [6].
Khoja's Stemmer	The Khoja stemmer is a well-known stemmer in Arabic, responsible for finding the root of any given word [7].	• Limited accuracy for words that deviate from standard patterns and forms [20].
Hybrid Stemmer	This method leverages an optimization function to enhance the n-gram technique [25].	• Performance may be compromised if a complete or reliable database is unavailable.
Root Extraction	A novel root extraction algorithm. This algorithm was specifically designed to overcome the limitations and errors inherent in existing methodologies [11].	• Challenges with complex morphology and Arabized words [37] • Scalability, adaptability, and impact of input data quality need further [28].

The Fig. 2, below categorizes Arabic stemming algorithms into four types: Light Stemmers, Root-based Stemmers, Hybrid Stemmers, and New Root Extraction. Each category includes specific algorithms such as Light10, Khoja's Stemmer, and the ISRI Stemmer, highlighting the different approaches to affix removal and root extraction in Arabic text processing.

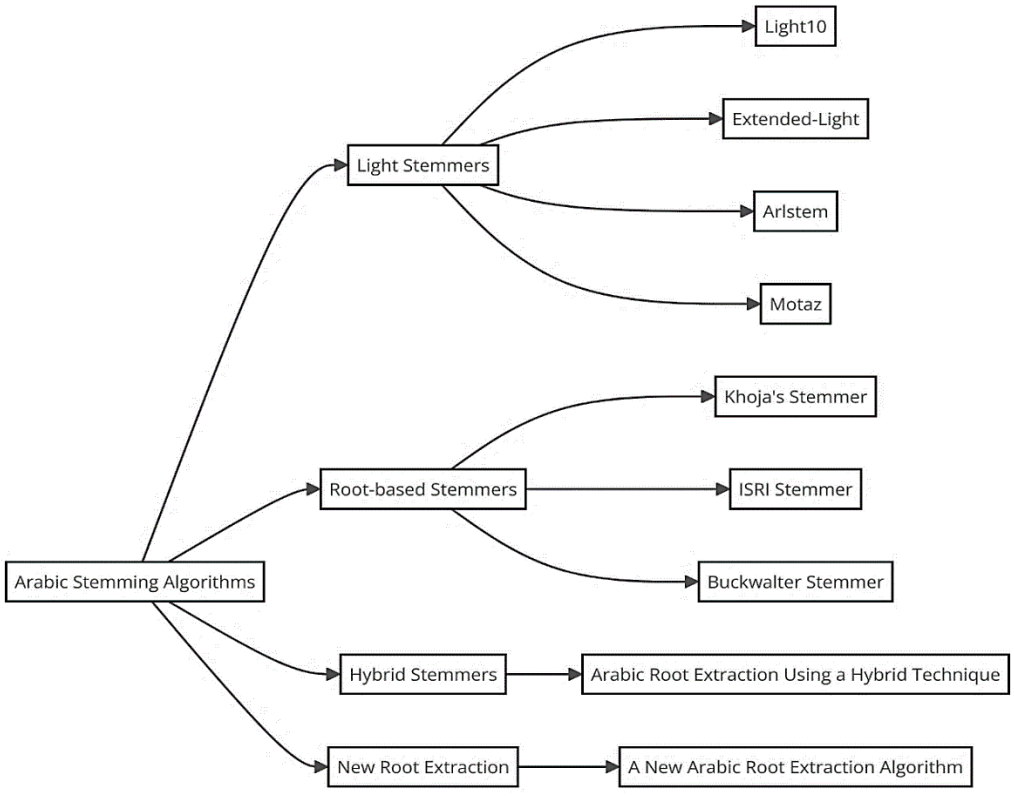

Fig. 2: Hierarchical Overview of Arabic Stemming Algorithms

2.8 Semantic Analysis and WordNet

Semantic analysis aims to understand the meaning and context of words within a text. This involves identifying relationships between words and their meanings, which is crucial for various NLP tasks, including information retrieval and machine translation [18]. WordNet is a lexical database that groups words into sets of synonyms (synsets) and captures their semantic relationships. Developed at Princeton University, WordNet has been a transformative tool for enhancing semantic analysis in NLP [27].

2.9 WordNet in Arabic NLP

The concept behind WordNet has been extended to support numerous languages, including Arabic. Arabic WordNet (AWN) serves as a similar lexical database, offering a semantic network of Arabic words and phrases linked by their relationships [23]. This tool is particularly valuable for Arabic NLP tasks such as machine translation, information retrieval, and sentiment analysis [18].

2.10 Potential Applications of WordNet in AES

WordNet's extensive database and structured knowledge make it a powerful tool for enhancing semantic analysis in AES systems. By integrating WordNet, AES systems can achieve a deeper understanding of essay content, leading to more accurate and fair evaluations. For Arabic AES, the incorporation of AWN can significantly improve the system's ability to interpret and assess the nuanced meanings of Arabic words [13].

2.11 Research Gap and Proposed Solution

Despite the potential of WordNet, there is a lack of research exploring its integration with robust stemming techniques for Arabic. This study aims to bridge this gap by proposing and evaluating a novel stemming algorithm that integrates AWN to enhance the accuracy and reliability of Arabic essay scoring. The proposed algorithm will be rigorously tested on a diverse dataset of Arabic essays to demonstrate its effectiveness in real-world educational settings. By bridging the gap between traditional stemming approaches and semantic analysis, this research has the potential to significantly advance the field of AAES and contribute to the development of more effective and equitable educational assessment tools for Arabic-speaking learners.

3. METHODOLOGY

This study introduces a novel approach to Arabic stemming, the Arabic Stemming Enhancement (ASE) algorithm, designed to address the limitations of existing methods. The ASE algorithm integrates Arabic WordNet (AWN) to enhance semantic understanding and improve the accuracy of root extraction in Arabic essays.

3.1 Data Collection

The study utilizes two datasets:

1. **Student Essay Dataset:** This dataset comprises 120 essays written by third-year secondary school students in Turkey on computer science topics. It serves as a real-world benchmark to evaluate the ASE algorithm's effectiveness in understanding and analyzing student comprehension of technical concepts.

2. **WordNet Stemming Dataset:** This dataset includes 7,861 Arabic words from the AWN, each paired with its verified root form. It acts as a linguistic benchmark to assess the ASE algorithm's accuracy in identifying and extracting root forms across various morphological complexities.

3.2 Methodological Framework

The methodological framework consists of nine distinct steps, as illustrated in Fig. 3:

1. **Pre-processing:** Raw essay text is cleaned by removing punctuation, numbers, and non-alphabetic characters. The text is then tokenized into individual words.

2. Part-of-Speech (POS) Tagging: Grammatical tags (e.g., noun, verb, adjective) are assigned to each token using the AWN, providing context for subsequent analysis.

3. **Stemming:** A hybrid stemming algorithm is applied, combining the Light10 stemmer, ISRI stemmer, and AWN's Look-in Tables method. This approach leverages both rule-based and lexicon-based stemming techniques to improve accuracy.

4. **Enhancement of Stemming Algorithm:** The stemming results are further refined using AWN's lexical database. This step ensures that the extracted stems are semantically valid and contextually appropriate.

5. **Identification of Key Terms:** Prominent nouns and verbs are extracted from the stemmed text to capture the main topics and concepts discussed in the essays.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 12, Issue 2, pp: (40-50), Month: April - June 2024, Available at: www.researchpublish.com

6. **Integration with WordNet:** The extracted key terms are mapped to their corresponding synsets in AWN, establishing semantic connections and enriching the analysis.

7. **Semantic Analysis:** The semantic relationships between the key terms are analyzed to gain a deeper understanding of the essay's content and meaning.

8. **Domain Determination:** The domain of each essay is determined based on the semantic analysis of the key terms and their association with specific domains in WordNet.

9. **Scoring Calculation Using Cosine Similarity:** Cosine similarity is used to measure the semantic similarity between student essays and model answers, providing a quantitative score for essay evaluation.

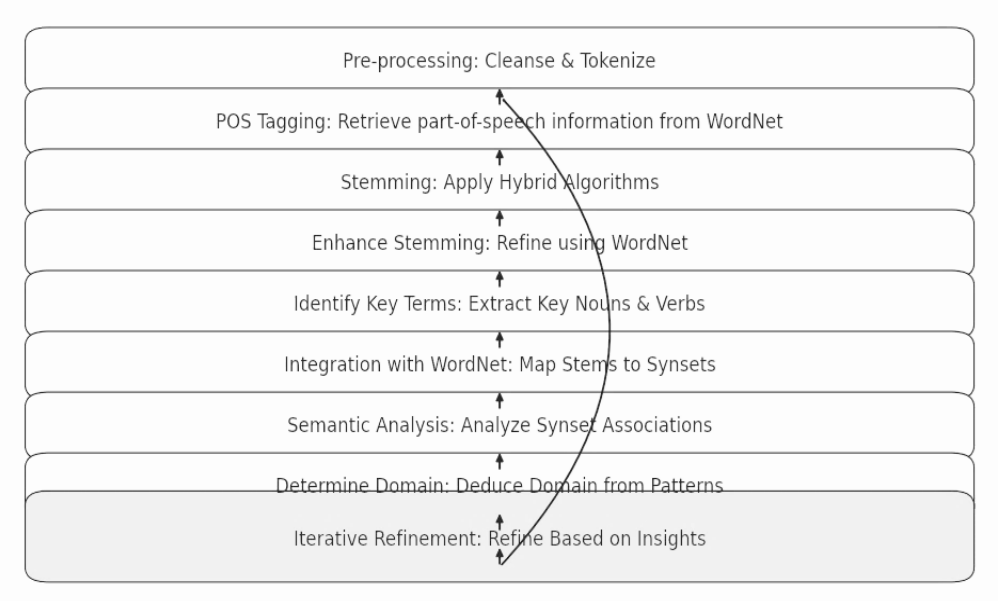

Fig. 3: Flowchart of the Methodological Framework

3.3 Limitations and Validity

The study acknowledges the limitations of the dataset, which may not fully represent the diversity of Arabic language use. Additionally, the reliance on WordNet, while beneficial, may not cover all nuances and variations of Arabic. These limitations highlight the need for further research to expand the dataset and enhance the semantic resources used in the algorithm.

Despite these limitations, the study's rigorous methodology and the use of established evaluation metrics ensure the validity and reliability of the findings. The results demonstrate the effectiveness of the ASE algorithm in improving the accuracy and efficiency of Arabic stemming, with potential implications for various NLP applications, including automated essay scoring.

4. RESULTS AND DISCUSSION

In the evolving landscape of Arabic NLP, the Arabic Stemming Enhancement (ASE) algorithm stands as a beacon of innovation, particularly evident in its handling of the term 'بحث' (Al-Bahth). Traditionally, the dual meanings of 'search' and 'study' entangled within 'Al-Bahth' have posed significant challenges for algorithmic disambiguation. Leveraging the sophisticated WordNet framework, ASE not only identifies but also accurately contextualizes these meanings based on their usage, marking a pivotal advancement in moving beyond mere word recognition to grasp the subtleties of context and semantics intrinsic to the Arabic language.

To illuminate ASE's nuanced interpretation capabilities, Fig. 3 provides a visual step-by-step breakdown of how ASE processes 'بحث' (Al-Bahth), distinguishing between 'search' and 'study' based on context. This visualization underscores ASE's sophisticated approach to semantic analysis, a cornerstone of its effectiveness in Arabic text processing.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 12, Issue 2, pp: (40-50), Month: April - June 2024, Available at: www.researchpublish.com

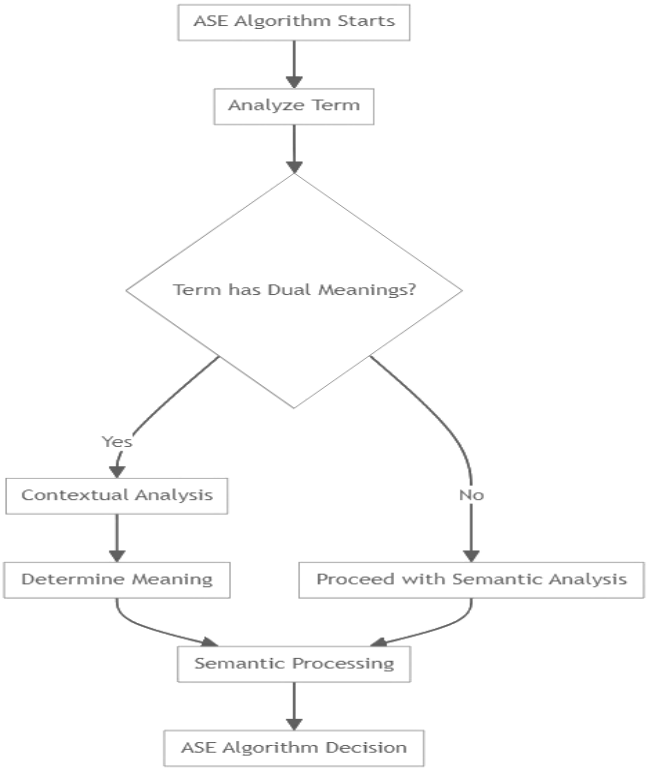

Fig. 4: ASE Algorithm: Transforming Word Recognition into Semantic Analysis

To meticulously evaluate ASE's effectiveness in Arabic NLP, a comprehensive array of performance metrics is utilized. Each metric is chosen specifically for its capacity to encapsulate a unique facet of ASE's performance, thereby providing a diverse perspective on its capabilities:

1. **Accuracy:** Measures the proportion of correct predictions made by ASE, serving as a broad indicator of its reliability.
2. **Precision:** Focuses on the proportion of true positive outcomes, highlighting ASE's accuracy in discerning relevant data points.
3. **Recall (sensitivity):** Assesses ASE's ability to identify all relevant instances, showcasing its depth in data mining.
4. **F1-score:** A balanced measure of ASE's precision and capacity to retrieve all pertinent information.
5. **Mean Absolute Error (MAE):** Provides a straightforward measure of ASE's prediction errors, quantifying the average deviation between predicted and actual values.
6. **Mean Squared Error (MSE) and Root Mean Squared Error (RMSE):** Offer deeper insights into the nature of ASE's prediction errors, with MSE calculating the average squared differences and RMSE quantifying the standard deviation of errors.

Applying these metrics to real-world data and case studies transcends theoretical analysis, demonstrating ASE's practical utility in navigating the complexities of Arabic texts. This holistic approach not only affirms ASE's significant contribution to Arabic NLP but also emphasizes its practical significance in enabling more accurate and nuanced text analysis and processing.

4.1 Performance Evaluation

The efficacy of ASE approaches is rigorously measured across both datasets, utilizing standard metrics such as accuracy, precision, recall, and F1-score. These metrics act as definitive measures of the performance and utility of each ASE method. The performance of ASE is evaluated using two key datasets: the WordNet-based dataset (theoretical ideal) and the Student Essays dataset (practical application). These evaluations provide insights into how ASE techniques perform in both controlled academic settings and real-world linguistic environments.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 12, Issue 2, pp: (40-50), Month: April - June 2024, Available at: www.researchpublish.com

4.2 Performance on Student Essays

Advancing NLP technologies, particularly for linguistically diverse languages like Arabic, requires a deep dive into ASE techniques. Focused attention on the benefits of WordNet integration underlines the endeavor to evaluate the efficacy of different ASE approaches. By examining the ability of each ASE method to process and comprehend essays written by students, the initiative aims to highlight the transformative potential these methods hold for educational technologies.

4.3 Performance on the WordNet Stemming Dataset

This segment delves into the effectiveness of different ASE techniques applied to the WordNet Stemming dataset, establishing a theoretical framework for their performance in conjunction with WordNet integration. Evaluation metrics such as accuracy, precision, recall, and F1-score are used for a thorough analysis.

Table 2 presents these performance metrics, highlighting the variances among the ASE techniques and setting the stage for their application in AES systems and other NLP applications within the educational sector.

Table 2: Stemming Algorithm Performance: WordNet Stemming

Stemming Type	Accuracy	Precision	Recall	F1 Score
ASE	0.978	0.9484	0.9498	0.9491
ISRI with Extend Light10	0.5779	0.8495	0.6438	0.7325
Light10 and ISRI	0.6521	0.9017	0.7020	0.7894
Extend Light10	0.2362	0.7055	0.2621	0.3822
Light10	0.2330	0.7396	0.2539	0.3780
ISRI	0.6191	0.8779	0.6775	0.7648
Tashaphyne 0.3.2	0.2108	0.6158	0.2427	0.3482
ARLStem	0.2937	0.8139	0.3149	0.4541
Buckwalter	0.1734	0.7219	0.1858	0.2955
Motaz	0.6520	0.8960	0.7053	0.7893
Khoja	0.1773	0.7264	0.1900	0.3012
Snowball	0.2337	0.7613	0.2522	0.3788
Tashaphyne 0.3.2 and ISRI	0.4898	0.7873	0.5644	0.6575

Fig. 4 visually compares the precision of the various stemming algorithms, further emphasizing the distinctions between traditional and semantic-based approaches.

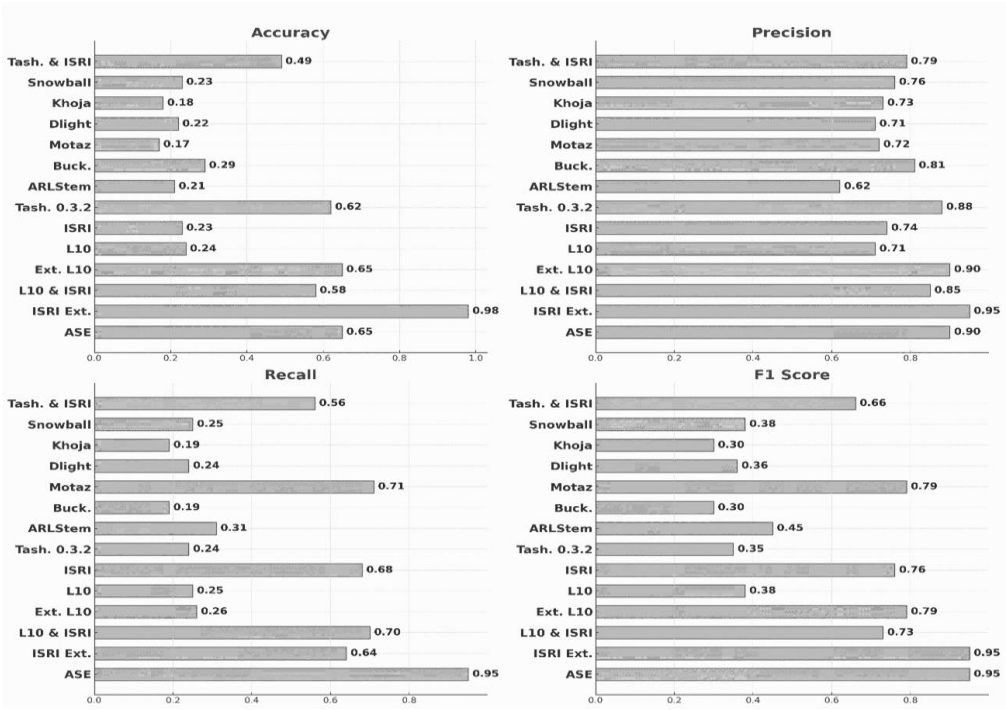

Fig. 5: Precision of Various Stemming Types

4.4 Theoretical vs. Practical Implications

While establishing theoretical benchmarks for evaluating performance is important, the true value of ASE techniques becomes apparent through their practical application in AES systems. The analysis based on the WordNet stemming dataset reveals important considerations, underscoring the importance of choosing ASE methods that are best suited to meet the specific requirements of educational assessments.

The evaluation of ASE approaches using the WordNet stemming dataset offers critical insights into their effectiveness. The findings delineate the strengths and limitations of each ASE method, fostering a deeper comprehension of their practical benefits and challenges in the realm of AES. This analysis aids in the informed selection of ASE techniques, optimizing their implementation for enhanced linguistic analysis and educational technology solutions.

5. CONCLUSION

The Arabic Stemming Enhancement (ASE) algorithm, a novel approach that integrates Arabic WordNet (AWN) to enhance stemming accuracy in Arabic text analysis, was rigorously evaluated using two distinct datasets: student essays on computer science topics and a WordNet-based stemming dataset. The results unequivocally demonstrate ASE's superior performance compared to traditional stemming algorithms. ASE achieved an accuracy of 97.70% on the "Student Essays" dataset and outperformed other algorithms on the WordNet stemming dataset across all evaluation metrics, including accuracy, precision, recall, and F1-score. This superior performance is attributed to ASE's ability to leverage semantic information from AWN, enabling it to accurately identify the root forms of words, even in cases of polysemy and homonymy. The integration of WordNet allows ASE to overcome the limitations of traditional stemming algorithms, which often struggle with the complexities of Arabic morphology. The results highlight the importance of semantic analysis in Arabic stemming and its potential to significantly enhance the effectiveness of AAES systems and other NLP applications for Arabic.

REFERENCES

[1] Abboud, K., Golovneva, O., & DiPersio, C. (2022). Cross-lingual transfer for low-resource Arabic language Understanding. Proceedings of the Seventh Arabic Natural Language Processing Workshop (WANLP), 225-237. doi: 10.18653/v1/2022.wanlp-1.21

[2] Abd, D., Khan, W., Thamer, K., & Hussain, A. (2021). Arabic Light Stemmer Based on ISRI Stemmer. Intelligent Computing Theories and Application, 12838, 32–45. doi:10.1007/978-3-030-84532-2_4

[3] Abainia, K., Ouamour, S., & Sayoud, H. (2017). A Novel Robust Arabic Light Stemmer. Journal of Experimental and Theoretical Artificial Intelligence, 29, 557-573. doi:10.1080/0952813X.2016.1212100

[4] Alhaj, Y., Xiang, J., Zhao, D., Al-Qaness, M., Abd Elaziz, M., & Dahou, A. (2019). A Study of the Effects of Stemming Strategies on Arabic Document Classification. IEEE Access, 7, 32664-32671. doi:10.1109/ACCESS.2019. 2903331

[5] Alhawarat, M., Abdeljaber, H., & Hilal, A. (2021). Effect Of Stemming on Text Similarity for Arabic Language at Sentence Level. PeerJ Computer Science, 7, 01-18. doi:10.7717/peerj-cs.530

[6] Al-Khulaidi, A., & Yaseen, S. (2023). Comparative Analysis and Evaluation of Stemming and Preprocessing Techniques for Arabic Text. Sana'a University Journal of Applied Sciences and Technology, 1(4), 313–326. doi:10.59 628/jast.v1i4.588

[7] Al-Khatib, R., Zerrouki, T., Shquier, M., & Balla, A. (2023). Tashaphyne 0.4: A New Arabic Light Stemmer Based on Rhizome Modeling Approach. Information Retrieval Journal, 26, 01-30. doi:10.1007/s10791-023-09429-y

[8] Almazrua, A., Almazrua, M., & Alkhalifa, H. (2020). Comparative Analysis of Nine Arabic Stemmers on Microblog Information Retrieval. International Conference on Asian Language Processing (IALP), 60-65. doi:10.1109/IALP513 96.2020.9310456

[9] Alqahtani, A., & Alsaif, A. (2019). Automatic Evaluation for Arabic Essays: A Rule-Based System. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 01-07. doi:10.1109/ISSPIT47 144.2019.9001802

[10] Alselwi, G., & Taşcı, T. (2024). Extractive Arabic Text Summarization Using PageRank and Word Embedding. Arabian Journal for Science and Engineering, 01-16. doi:10.1007/s13369-024-08890-1

[11] Al-Shalabi, R. (2022). A Priority Based Approach for Stemming Arabic Word. Applied Science University Journal (ASUJ), 6(2), 377-380. doi:10.18576/jasu/060206

[12] Asadi, I., Vaknin-Nusbaum, V., & Taha, H. (2023). The Role of Morphological Decomposition in Reading Complex Words in Arabic in Elementary School Years. Journal of Psycholinguistic Research, 52(6), 2863–2876. doi:10.1007/s10936-023-10024-x

[13] Atwan, J., Wedyan, M., Bsoul, Q., Hammadeen, A., & Alturki, R. (2021). The Use of Stemming in the Arabic Text and Its Impact on the Accuracy of Classification. Scientific Programming, 01-09. doi:10.1155/2021/1367210

[14] Bashir, M., Azmi, A., Nawaz, H., Zaghouani, W., Diab, M., Al-Fuqaha, A., & Qadir, J. (2023). Arabic Natural Language Processing for Qur'anic Research: a Systematic Review. Artificial Intelligence Review, 56(7), 6801–6854. doi:10.1007/s10462-022-10313-2

[15] Darwish, K., Habash, N., Abbas, M., Al-Khalifa, H., Al-Natsheh, H. T., El-Beltagy, S. R., . . . Mubarak, H. (2021). A Panoramic Survey of Natural Language Processing in the Arab World. arXiv, 64(4), 01-34. doi:10.1145/3447735

[16] Elmadany, A., Nagoudi, E., & Abdul-Mageed, M. (2022). ORCA: A Challenging Benchmark for Arabic Language Understanding. arXiv, 01-26. doi:10.48550/arXiv.2212.10758

[17] Ezzikouri, H., Madani, Y., Erritali, M., & Oukessou, M. (2019). A New Approach for Calculating Semantic Similarity between Words Using WordNet and Set Theory. Procedia Computer Science, 151, 1261–1265. doi: 10.1016/j.procs.2019.04.182

[18] Freihat, A., Khalili, H., Bella, G., & Giunchiglia, F. (2024). Advancing the Arabic WordNet: Elevating Content Quality. arXiv, 01-10. doi:10.48550/arXiv.2403.20215

[19] Habash, N. (2022). Arabic Natural Language Processing. Association for Computational Linguistics, 9-10. doi: 10.18653/v1/2022.emnlp-tutorials.2

[20] Hamza, M., Ahmed, T., & Hilal, A. (2021). Text Mining: A Survey of Arabic Root Extraction Algorithms. International Journal of Advanced and Applied Sciences, 8(1), 11-19. doi:10.21833/ijaas.2021.01.002

[21] Hussien, S., & Aburagheef, H. (2021). Arabic Light-Based Stemming: A Comparative Study Among Ligh10 Stemmer, P-Stemmer, and Conditional Light Stemmer. Proceedings of the 2021 2nd Information Technology To Enhance e-learning and Other Application (IT-ELA), 131-135. doi:10.1109/IT-ELA52201.2021.9773743

[22] Jabbar, A., Iqbal, S., Tamimy, M., Hussain, S., & Akhunzada, A. (2020). Empirical Evaluation and Study of Text Stemming Algorithms. Artificial Intelligence Review, 53, 5559–5588. doi:10.1007/s10462-020-09828-3

[23] Jarrar, M. (2021). The Arabic Ontology -An Arabic Wordnet with Ontologically Clean Content. Applied ontology, 16(1), 01-26. doi:10.3233/AO-200241

[24] Kasem, M., Mahmoud, M., & Kang, H.-S. (2023). Advancements and Challenges in Arabic Optical Character Recognition: A Comprehensive Survey. arXiv ,01-28. doi:10.48550/arXiv.2312.11812

[25] Khafajeh, H., Yousef, N., & Abdeldeen, M. (2018). Arabic Root Extraction Using a Hybrid Technique. International Journal of Advanced Computer Research, 8, 90-96. doi:10.19101/IJACR.2017.733023

[26] Khyani, D., & B S, S. (2020). An Interpretation of Lemmatization and Stemming in Natural Language Processing. Journal of University of Shanghai for Science and Technology, 22(10), 350-357.

[27] Lam, K., Tarouti, F., & Kalita, U. (2022). Automatically Constructing Wordnet Synsets. arXiv, 2, 01-06. doi:10.3115/v1/p14-2018

[28] Larabi Marie-Sainte, S., Alalyani, N., Alotaibi, S., Ghouzali, S., & Abunadi, I. (2019). Arabic Natural Language Processing and Machine Learning-Based Systems. IEEE Access, 7011-7020.

[29] Machhout, R., Zribi, C., & Bouzid, S. (2022). Arabic Automatic Essay Scoring Systems: An Overview Study. Intelligent Systems Design and Applications, 418, 1164–1176. doi:10.1007/978-3-030-96308-8_108

[30] Muaad, A., Raza, S., Naseem, U., & Davanagere, H. (2023). Arabic Text Detection: A Survey of Recent Progress, Challenges, and Opportunities. Applied Intelligence, 53(24), 29845–29862. doi:10.1007/s10489-023-04992-9

[31] Mustafa, E. (2023). A Bi-Gram Approach for an Exhaustive Arabic Triliteral Roots Lexicon. Languages, 8(1), 01-15. doi:10.3390/languages8010083

[32] Mustafa, M., Eldeen, A., Bani-Ahmad, S., & Elfaki, A. (2017). A Comparative Survey on Arabic Stemming: Approaches and Challenges. Intelligent Information Management, 9, 39-67. doi:10.4236/iim.2017.92003

[33] Naili, M., Chaibi, A., & Ben Ghezala, H. (2019). Comparative Study of Arabic Stemming Algorithms for Topic Identification. Procedia Computer Science, 159, 794–802. doi: 10.1016/j.procs.2019.09.238

[34] Ouahiba, S., & Othman, R. (2019). Retrieval Performance of Arabic Light Stemmers. International Journal of Modern Trends in Social Sciences, 2(10), 81-90. doi:10.35631/IJMTSS.210008

[35] Oussous, A., Lahcen, A., & Belfkih, S. (2019). Impact of Text Pre-processing and Ensemble Learning on Arabic Sentiment Analysis. Proceedings of the Second International Conference on Networking, Information Systems, and Security, 01-09. doi:10.1145/3320326.3320399

[36] Rianto, R., Mutiara, A. B., Wibowo, E. P., & Santosa, P. I. (2021). Improving the Accuracy of Text Classification Using Stemming Method, a Case of Non-Formal Indonesian Conversation. Journal of Big Data, 8(1), 01-16. doi:10.1186/s40537-021-00413-1

[37] Shaalan, K., Siddiqui, S., Alkhatib, M., & Monem, A. (2018). Challenges in Arabic Natural Language Processing. Computational Linguistics, Speech and Image Processing for Arabic Language, 59–83.

[38] Singh, J., & Gupta, V. (2017). A Systematic Review of Text Stemming Techniques. Artificial Intelligence Review, 48(2), 157–217. doi:10.1007/s10462-016-9498-2

[39] Syarief, M., Kurahman, O., Huda, A., & Darmalaksana, W. (2019). Improving Arabic Stemmer: ISRI Stemmer. IEEE 5th International Conference on Wireless and Telematics (ICWT), 01-04. doi:10.1109/ICWT47785.2019.8978248

[40] Uto, M. (2021). A Review of Deep-Neural Automated Essay Scoring Models. Behaviormetrika, 48(2), 459–484. doi:10.1007/s41237-021-00142-y

[41] Zahir, J. (2022). IADD: An Integrated Arabic Dialect Identification Dataset. Data in Brief, 40, 107777. doi: 10.1016/j.dib.2021.107777