EDRAAK
Vol. **(2025)**, 2025, **pp**. 94–102
ISSN: 3078-8412

Research Article

Classification Arabic language (Classical Arabic Poetry, Al-Hur Arabic Poetry and Prose) Using Machine Learning

Munef Abdullah Ahmed[1,*], Raed Abdulkareem Hasan [2], Mostafa Abdulghafoor Mohammed[3], Peter Mwangi[4], Tirus Muya[4]

[1] *Electronics Department - Alhawija Technical Institute, Northern Technical University, Iraq*

[2] *Renewable Energy Research Unit, Northern Technical University, Mosul, Iraq*

[3] *Department of Computer science, Imam Aadham university college, Baghdad, Iraq*

[4] *Department of computer science, Murang'a University of Technology, Kenya*

ARTICLEINFO

Article History
Received 20 Mar 2025
Accepted: 13 May 2025
Revised 14 Jun 2025
Published 4 Jul 2025

Keywords
Arabic Text Classification,
Classical Arabic Poetry,
Al-Hur Arabic Poetry,
Machine Learning Algorithms (SVM, LSVC, Naïve Bayes),
Natural Language Processing (NLP).

ABSTRACT

Many languages globally have made significant advances in electronically studying and classifying texts. Making electronic text a great alternative to manual classification by saving time, cost, and effort. However, Arabic has not seen similar progress due to several limitations faced by researchers, such as the complexity of the language, the scarcity of related research, and the use of classical Arabic. Additionally, the poetry presents further challenges, like the reliance on a single activation function. This paper introduces a new method for classifying the Arabic text (Prose, classical Arabic poetry and Al-hur Arabic poetry) based on distinctive features that identify the type of Arabic text. Prepressing data is crucial in this approach as it enhances classification accuracy.

1. INTRODUCTION

Natural language processing is considered one of the most important topics at the present time, as it can save time, especially after the entry of artificial intelligence into this field. In languages that use Latin letters, we see great development in this field, but we have not seen this prosperity in the Arabic language [1-4]. This is due to several reasons, including the lack of research in this field, as well as the difficulty of the Arabic language, especially in the field of poetry, whether Al-hur Arabic poetry, classical Arabic poetry, or prose texts, as each of them has its own rules and requires specialists in this field. It is worth noting that most research relies on verbal content in classification, except for a few researches that relied on feelings and emotions using machine learning [5-11].
This paper is structured into several sections, starting with a review of related literature, which is followed by an exploration of the various forms of Arabic poetry, specifically classical Arabic poetry and Al-Hur Arabic poetry and prose[12]. The subsequent section addresses the datasets utilized in this study, followed by a detailed explanation of the machine learning algorithms employed. The methodology is then presented, leading into the results and discussion section.

*Corresponding author email: m.ahmed1988@gmail.com
DOI: https://doi.org/10.70470/EDRAAK/2025/012

2. STATE OF THE ART

The categorization of emotions in the English language has been examined, and the different ways to express this emotion have been identified. Some classifications are based on simple attributes that cannot be explained in more than one way [13-15]. The automatic classification of Arabic poems according to the overall sentiment of the writing[16]. Text classification is a supervised learning method that automatically categorizes text inputs that lack labels using the derived classifer after utilizing labeled training data to create a system of classification (classifier).[17-20] . With the huge amount of documents and the development of the Internet, document classification has become a necessity. This is done by linking the document to specific categories of documents based on the content. This automatic classification helps solve many problems, such as determining the language of the document and directing and redirecting messages. [21]. Due to the lack of much data that can be used for comparison in classification in the Arabic language, this study provided an amount of data that can be used for comparison.[22-25]. using Naïve Bayes, Linear Support Vector, and Support Vector Machines for classification of modern Arabic poetry [26-28]. Machine learning is used for the classification classical Arabic poetry and Al-Hur Arabic poetry [29].

3. CATÉGORIES OF ARABIC TEXTS

3.1 Classical Arabic Poetry:

In this style of Arabic poetry, each verse is called a "byte" and is composed of two parts: the first part is called the "sadr" and the second part is called the "ajuz". The final word of the "sadr" is resolved as "arud" and the other word is referred to as "hashu sadr." Similarly, the final syllable in "ajuz" is called "darb," the other is designated as "hashu ajuz." Classical Arabic literature includes three different approaches. The initial style is singular: a single row is split into two columns, the first column is dedicated to "sadr" and the second is dedicated to "ajuz," as illustrated in Figure 1.

Verse	H2 or ”ajuz1”	H1 or ”sadr1”
	فَلا خَيرَ في وِدٍّ يَجيءُ تَكَلُّفا	إِذا لَم يَكُن صَفوُ الوِدادِ طَبيعَةً

Fig. 1. The initial form of classical Arabic poetry.

The second type, seen in figure 2, has one column and two rows, with the first row designating "sadr" and the second row designating "ajuz.".

.

Verse	H1 or ”sader1”	اذا لَم يَكُن صَفوُ الوِدادِ طَبيعَةً
	H2 or ”ajuz1”	فَلا خَيرَ في وِدٍّ يَجيءُ تَكَلُّفا

Fig. 2. The second form of classical Arabic poetry.

Third style: As seen in Figure 3, this style similarly consists of a single column and two rows, but the first row, which represents the "sadr," is positioned to the right of the others, while the second row, which represents the "ajuz," is set up to the left.

.

verses	Verse one ”sader1”	اذا لَم يَكُن صَفوُ الوِدادِ طَبيعَةً
	Verse two ”sader1”	فَلا خَيرَ في وِدٍّ يَجيءُ تَكَلُّفا

Fig. 3. The third form of classical Arabic poetry.

AlKhalil bin Ahmed Al Farahidi is a renowned Arab scholar that dedicated himself to the study of Arabic literature and left a significant impact on the field. He was the first to establish the rules of Arabic poetry, as he considered the "tafilah" to be the primary unit of measurement for the "buhur" of Arabic poetry. Based on his investigation, he discovered that Arabic poetry has 16 "buhur" (Mutadariq, Mutaqarib, Munsarih, Sari, Rajaz, Muqtadib, Tawil, Mujtath, Mudari, Madid, Kamil, Basit, Wafir, Hazaj, Khafif, Hamal). To determine the type of "buhur" associated with the poem, the state of the letter, whether it's consistent or mobile, and knowing that all verses in a single poem have the same "buhur" type.

3.2 Al-hur Arabic Poetry:

It depends on the unit of "tafileh" without restricting “tafileh” in each line. Interacts with poetic images, does not adhere to the unity of rhyme; the unity of the subject and the coherence of the poem are more important than the rhyme in Al-hur Arabic poetry verse, because following a certain rhyme may limit the creativity of poets. It does not contain the letter

"RAWI," which is the last letter of the verse. but adheres to unity and objectivity. Relies on pagan and philosophical linguistic books, which makes some of his poems go beyond ambiguity.

verses	Verse 1 ”sader1”	الشمس اجمل في بلادي من سواهـــــــــــــــــــــا
	Verse 2 ”sader2”	حتى الظلام هناك أجملُ ، فهو يحتضن العراق

Fig. 4. Al-hur Arabic Poetry with two verses.

3.3 Prose:

Poetry is considered the first type of Arabic literature, while the second type is prose, which in turn is divided into two sections:

3.3.1 Artistic prose

It is known as beautiful artistic speech in which the writer chooses good and clear words and an effective style to present his idea or topic, such as a story, novel, article, or biography.

3.3.2 Normal prose

It is what people use in their daily lives and dealings, in school books and newspaper reports, which do not require an effective style from the writer.

Artistic prose	Normal prose
عندما تكون المصالح فوق الشرف تسقط القيم ويرتفع النفاق فوق الجبين ويبقى الاحترام سيد كل العلاقات	ذهب طارق مع علي الى المدرسة

Fig. 5. Artistic Prose and Normal Prose

4. DATASET

The natural language processing in Arabic has a different approach than the English language, primarily due to the lack of data that are accessible to researchers. As a result, researchers have depended on data derived from various sources, such as websites, news media, and magazines. This dependence has led to different opinions among scientists regarding the proper way to choose datasets for training and testing, this is an important initial step in machine learning. The limited research in this area, especially regarding classical Arabic literature, prose, and Al-Hur Arabic literature, has increased the difficulty of these issues.

5. DATA PRE-PROCESSING

The Arabic language is considered one of the difficult languages and consists of twenty-nine letters. Some of these letters are called vowels, and there are three of them, and they have their own rules that differ from the rest of the letters, which are called consonants. Another characteristic of the Arabic language is that some letters are written in more than one way depending on the meaning and are called equivalent letters, shown in Table 1, and there are letters that have the same sound for some diacritics and are called equivalent vowels, shown in Table I.

Equivalent Letters	
Ta Marbuta, Ta	ه, ة
Ta Marbuta ,Ta	ت, ة
Equivalent Vowels	
/i/, /i:/	ي, ◌ِ, ◌ٍ
/u/, /u:/	و, وا, ◌ُ , ◌ٌ
/a/, /a:/	ا, ى, ◌َ, ◌ً

TABLE. I. EQUIVALENT LETTERS AND VOWELS

In the Arabic language, there are diacritics that may change the pronunciation and meaning, as shown in Table II.

The short vowel	The Sign	Applied to the letter	Prononciation
“Fatha”	◌َ	دَ – زَ	Ra - Da
“dammah”	◌ُ	دُ – زُ	Ru - Du
“Kasra”	◌ِ	دِ – زِ	Ri - Di
“sukoon”	◌ْ	دْ – زْ	R- D
“tanween fatha”	◌ً	دً – زً	Ran - Dan
“tanween dammah”	◌ٌ	دٌ – زٌ	Ron - Don

“tanween kasra”	◌ٍ	دٍ – زٍ	Rin - Din
“shadde”	◌ّ	دّ – زّ	Rr – Dd
Mad	~	آ	Aa

TABLE. II. ARABIC LANGUAGE DIACRITICS

Some letters in the Arabic language may take several forms depending on their position in the word. We can see an example of this by looking at Table III.

The Arabica letter	The Word	Meaning
ـه	أبدله	Replace him
ـهـ	الهشيم	Broken
هـ	هنا	Her
ه	صغيره	Small

TABLE. III. EXAMPLE FOR LETTER FORMS

The Arabic language has gender-specific distinctions, which are governed by their own set of regulations. This language is categorized into three varieties: singular, plural, and dual, each of which has its own set of rules pertaining to gender equality across the three varieties. To create a system of classification that utilizes machine learning, several pre-requisites are necessary to reduce noise and focus on the features that have the greatest influence on classification. This procedure is important for preserving the memory during the classification process, particularly with large datasets, it also increases the accuracy and speed of the classification. Key tasks involved in data preprocessing include:

Tokenization: This process involves segmenting the data into components based on attributes and recognizing delimiters such as punctuation and whitespace. Removal of non-Arabic numerals, words, terms, symbols, and punctuation.

The creation or addressing of a vector of elements is a crucial and fundamental aspect of machine learning that fundamentally influences the outcomes of artificial intelligence calculations. Each issue should be addressed using its own components.

$D = d1d2d3d4 \ldots\ldots\ldots\ dn$

$d = w1w2w3w4 \ldots\ldots\ldots wn$

Where(D, w)represent document

$d = \ddot{g}(d)$ This relationship can be linear or non-linear.

Where g represent the relation between dome and the feature.

$\text{length of features vector} = \mathrm{C} * \mathrm{K}$

$\mathrm{C} = \text{Class numbers for the vector}$

$\mathrm{K} = \text{features for the vector}$

The likelihood of each characteristic within the various class types (c) can be expressed in the following manner: $n_c=(fi)$; therefore, the primary focus that is typically acknowledged is: .

$dn_c(f_i) = (n_c(f_i) - n_d(f_i))$ where $(d \neq c)$

The text refers to the frequency of occurrences of specific features or elements within a given classification, derived from the frequency of similar features in other classifications. There are two categories of vector models: boolean and count. This study employs the first category, as it is more appropriate than the latter [9].

6. MACHINE LEARNING ALGORITHMS

Three categories of machine learning techniques were employed in this study: linear support vector machines, support vector classification, and Naïve Bayes. These methods are predominantly utilized for the classification of English text and have demonstrated significant efficacy in categorizing the English language.

6. 1 Support Vector Machine algorithm

The field of machine learning (ML) is considered to be of significant importance that affects various aspects of our daily lives in multiple disciplines, including natural language processing, healthcare, industry, and agriculture. Among its numerous uses, classification is particularly important in the context of ML. Support Vector Machines (SVM) are primarily utilized for classification and regression purposes because of their effectiveness and precision when dealing with categorical data. Classification can be categorized into two types: binary classification, which involves the distinction of two classes, and multi-class classification, which encompasses the classification of multiple classes. This algorithm has gained prominence, particularly in the areas of pattern diagnosis, image processing, and the recognition of handwritten digits. For straightforward data, linear classification is applicable; however, in cases where the data is intricate and multi-dimensional, the kernel method of the algorithm is utilized [30].

Classification accuracy is assessed by determining tree characteristics, specifically recall, precision, and the F1-measure. [31-35]

For class *Ci* the precision:

$Pi = ({}^{TPi}/_{(TPi + FPi)})$

Where (TPi, FPi) refer to true and false positive.
For class Ci the recall:
$Ri = \left({}^{TPi}\!/_{(TPi + FNi)}\right)$
Where FNi refer to false negative.
The following function used for calculating Fi [36]
$$F1 = \left(\frac{(2 * Recall * Precision)}{(Recall + Precision)}\right)$$
$$F1 = \left(\frac{2TP}{(2TP + FP + FN)}\right)$$

6. 2 Linear Support Vector Classification

It is regarded as one of the key machine learning algorithms, much like SVM, but it is implemented using liblinear rather than libsvm. The ability to select and lose functions is one of this algorithm's most crucial features; it is also utilized to determine the properties of numerous samples. Numerous tests have revealed that the LSVC is based on the one-versus-rest approach, whereas SVM is based on the one-versus-one approach. It has a variety of uses, but in the area of natural language classification, it stands out. [37-43]

6. 3 Naïve Bayes

The Bayes theorem, which holds that there is no correlation between any particular trait in a class and any other trait, is the foundation of this machine learning algorithm. According to this theory, the document class is provided as [14]
$C^* = (argmax_c * P(c|d))$
Where c represent the class and d the document.
$$C^* = argmax_c * P(d|c) * \left({}^{P(c)}\!/_{P(d)}\right)$$
This theory states that p (d) had no effect, and the resulting equations are as follows:
$C^* = argmax_c * P(d|c) * p(c)$
Since the feature in this theory was independent of the other feature
$$C^* = argmax_c * P(d|c) * \prod_i^n p(f_i/c) * p(c)$$

7. METHODOLOGY

Figure 6 illustrates the framework of our research. Initially, we selected the dataset utilized in our paper; subsequently, we divided it into individual words and implemented all necessary data preprocessing steps, which included feature extraction. We employed three machine learning algorithms (SVM, L SVC, and NB) for both training and testing purposes.

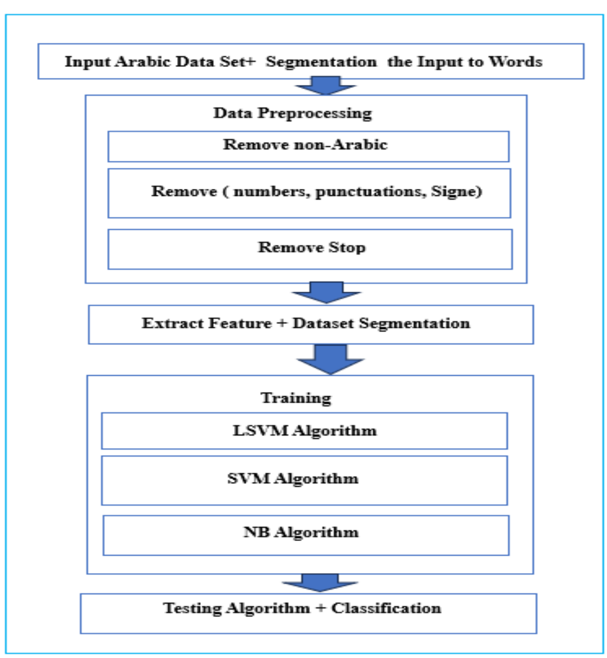

Fig. 6. Methodology for our paper

8. RESULT

The results in Table IV by using the SVM algorithm in (prose, Al-hur Arabic poetry, classical Arabic poetry) to find precision, recall, and f-measure

SVM Algorithm	precision	recall	f-measure
Prose	0.72	0.42	0.39
Al-hur Arabic Poetry	0.08	0.15	0.33
classical Arabic poetry	0.42	0.38	0.35
Average	0.407	0.317	0.357

TABLE. IV. PRECISION, RECALL AND F-MEASURE FOR DATA SET USING SVM ALGORITHEM.

The results in Table V are the same as in Table 4, but we used LSVM instead of SVM.

L..SVC Algorithm	precision	recall	f-measure
Prose	0.79	0.57	0.73
Al-hur Arabic Poetry	0.66	0.69	0.58
classical Arabic poetry	0.81	0.61	0.71
Average	0.753	0.623	0.673

TABLE. V. PRECISION, RECALL AND F-MEASURE FOR DATA SET USING SVM ALGORITHEM.

The results are in Table VI. By using the NB algorithm in our dataset.

Naive Base Algorithm	precision	recall	f-measure
Prose	0.82	0.86	0.61
Al-hur Arabic Poetry	0.71	0.52	0.64
classical Arabic poetry	0.53	0.79	0.53
Average	0.687	0.723	0.593

TABLE. VI. PRECISION, RECALL AND F-MEASURE FOR DATA SET USING LSVC ALGORITHEM.

The result in Table VII presents the average by using algorithms in our dataset.

Algorithems	precision	recall	f-measure
SVM Algorithm	0.407	0.317	0.357
L.SVC Algorithm	0.753	0.623	0.673
Naive Base Algorithm	0.687	0.723	0.593
Average			

TABLE. VII. PRECISION, RECALL AND F-MEASURE FOR DATA SET USING NAÏVE BAYES ALGORITHEM.

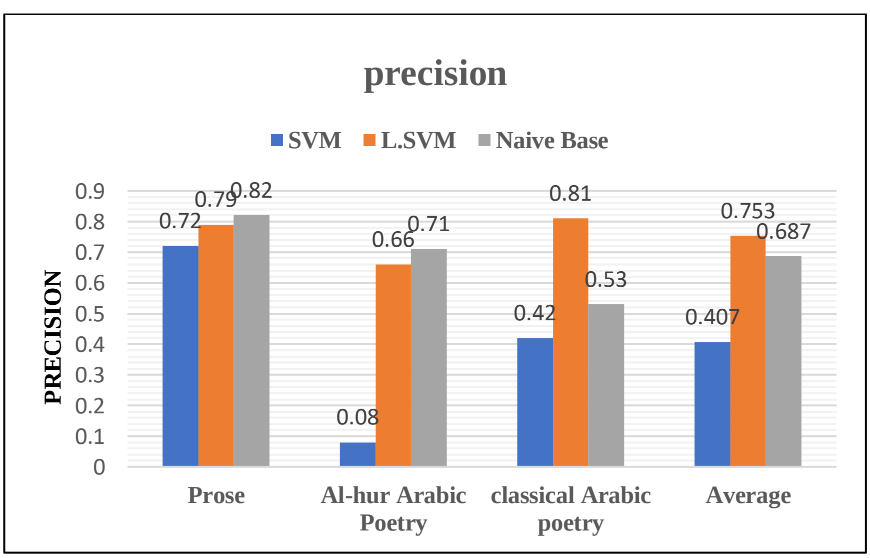

Fig. 7. Precision for SVM, LSVM and Naïve Base Algorithms

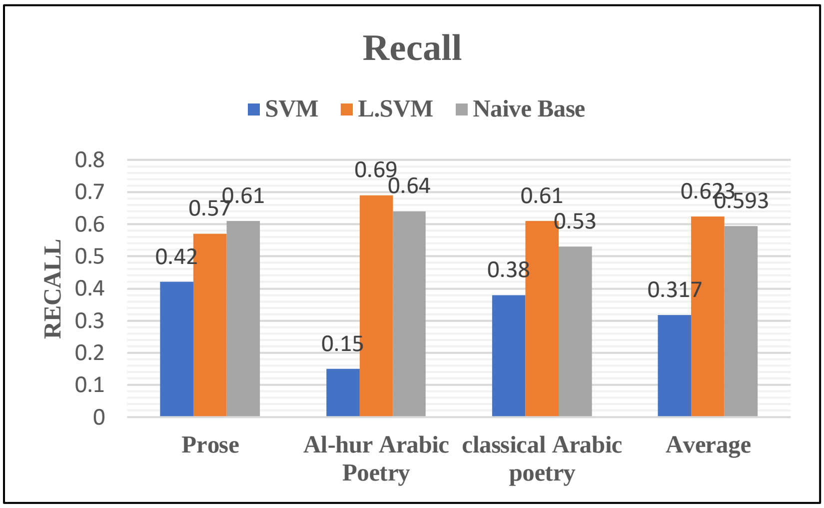

Fig. 8. Recall for SVM, LSVM and Naïve Base Algorithms

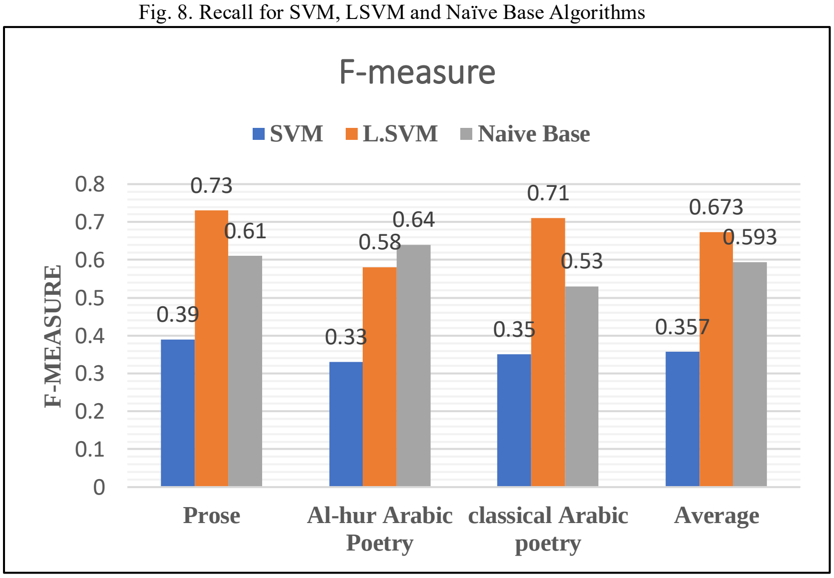

Fig. 9. F-measure for SVM, LSVM and Naïve Base Algorithms

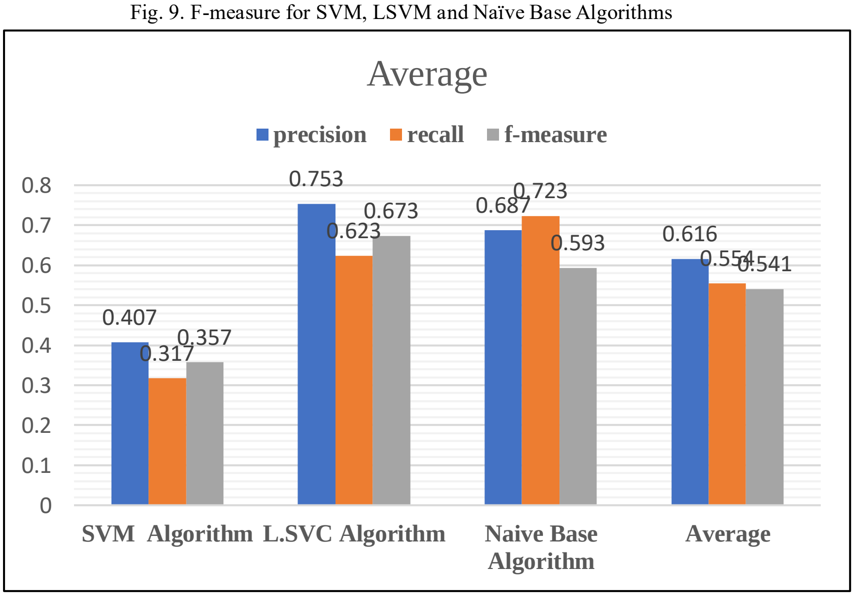

Fig. 10. Average Precision, Recall and F-measure for Algorithms

9. CONCLUSIONS

In this paper, we used support vector machine, linear support vector classification, and Naïve Bayes to classify Arabic text. By comparing the precision, recall, and f-measure results for each type of Arabic text. We discovered that the maximum average precision, recall, and F-measure were when we used the LSVM algorithm; the large precision was for the Arabic prose when we used Naïve Bayes, while the maximum recall was for Al-Hur Arabic poetry and the maximum F-measure was for the Arabic prose with the LSVM algorithm. The dataset's size is a significant factor contributing to this performance disparity, as some machine learning algorithms are capable of performing effectively with a smaller number of datasets. Furthermore, our dataset underwent preprocessing to eliminate essential components that could enhance classification accuracy while also reducing the memory requirements. This method can be used to classify other types of Arabic text.

Funding:

The authors declare that no specific financial aid or sponsorship was received from governmental, private, or commercial entities to support this study. The research was solely financed by the authors' own contributions.

Conflicts of Interest:

The authors declare that there are no conflicts of interest in this study.

Acknowledgment:

The authors express their heartfelt appreciation to their institutions for the essential support and motivation provided throughout the research period.

References

[1] M. A. Ahmed and S. Trausan-Matu, “Using natural language processing for analyzing Arabic poetry rhythm,” in *Networking in Education and Research (RoEduNet), 2017 16th RoEduNet Conference*, 2017, pp. 1–5.

[2] M. Abdul-Mageed, M. T. Diab, and M. Korayem, “Subjectivity and sentiment analysis of modern standard Arabic,” in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*,pp.587-591, 2011, https://aclanthology.org/P11-2103.pdf

[3] G. L. Clore, A. Collins, and A. Ortony, *The Cognitive Structure of Emotions*, Cambridge, UK: Cambridge University Press, 1990.

[4] H. Liu, H. Lieberman, and T. Selker, “A model of textual affect sensing using real-world knowledge,” in *Proceedings of…*, pp. 125–132, 2011 https://dl.acm.org/doi/abs/10.1145/604045.604067

[5] O. Alsharif, D. Alshamaa, and N. Ghneim, “Emotion classification in Arabic poetry using machine learning,” *International Journal of Computer Applications*, vol. 65, no. 16, 2013.

[6] S. Alsaleem, “Automated Arabic text categorization using SVM and NB,” *International Arab Journal of e-Technology*, vol. 2, no. 2, pp. 124–128, 2011.

[7] R. Belkebir and A. Guessoum, “A hybrid BSO-Chi2-SVM approach to Arabic text categorization,” in *Proceedings of …*, pp. 1–7.

[8] M. S. Khorsheed and A. O. Al-Thubaity, “Comparative evaluation of text classification techniques using a large diverse Arabic dataset,” *Language Resources and Evaluation*, vol. 47, no. 2, pp. 513–538, 2013.

[9] M. A. Ahmed, R. A. Hasan, A. H. Ali, and M. A. Mohammed, “The classification of the modern Arabic poetry using machine learning,” *Telkomnika*, vol. 17, no. 5, pp. 2667–2674, 2019.

[10] M. A. Ahmed, “A classification of Al-hur Arabic poetry and classical Arabic poetry by using support vector machine, Naïve Bayes, and linear support vector classification,” *Iraqi Journal for Computer Science and Mathematics*, vol. 3, no. 2, pp. 128–137, 2022.

[11] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up?: sentiment classification using machine learning techniques,” in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 79–86, 2002, https://arxiv.org/abs/cs/0205070

[12] X. Zhang, S. Ding, and Y. Xue, “An improved multiple birth support vector machine for pattern classification,” *Neurocomputing*, vol. 225, pp. 119–128, 2017.

[13] T. Joachims, “Text categorization with support vector machines: Learning with many relevant features,” in *Proceedings of the 10th European Conference on Machine Learning*, pp. 137–142, 1988, https://link.springer.com/chapter/10.1007/BFb0026683

[14] W. Jiang, D.-S. Huang, and S. Li, “Random walk-based solution to triple level stochastic point location problem,” *IEEE Transactions on Cybernetics*, vol. 46, no. 6, pp. 1438–1451, 2016, https://doi.org/10.1109/TCYB.2015.2446198

[15] F. Debole and F. Sebastiani, “An analysis of the relative hardness of Reuters-21578 subsets,” *Journal of the Association for Information Science and Technology*, vol. 56, no. 6, pp. 584–596, 2005.

[16] A. Saleh *et al.*, “Green building techniques: Under the umbrella of the climate framework agreement,” *Babylonian Journal of Machine Learning*, pp. 1–14, 2024, https://doi.org/10.58496/BJML/2024/001

[17] R. A. Hasan, M. M. Akawee, and T. Sutikno, “Improved GIS-T model for finding the shortest paths in graphs,” *Babylonian Journal of Machine Learning*, pp. 7–16, 2023, https://doi.org/10.58496/BJML/2023/002

[18] A. H. Ali *et al.*, "Big data classification based on improved parallel k-nearest neighbor," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 21, no. 1, pp. 235–246, 2023.

[19] S. A. Ahmed *et al.*, "Advancements in UAV image semantic segmentation: A comprehensive literature review," *Multidisciplinary Reviews*, vol. 7, no. 6, art. 2024118, 2024.

[20] S. I. Jasim, M. M. Akawee, and R. A. Hasan, "A spectrum sensing approaches in cognitive radio network by using cloud computing environment," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 750–757, 2022.

[21] M. A. Mohammed *et al.*, "A focal load balancer based algorithm for task assignment in cloud environment," in *2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI*), pp. 1–4, Jun. 2018

[22] R. A. Hasan and M. N. Mohammed, "A krill herd behaviour inspired load balancing of tasks in cloud computing," *Studies in Informatics and Control*, vol. 26, no. 4, pp. 413–424, 2017.

[23] R. A. Hasan *et al.*, "A comprehensive study: Ant Colony Optimization (ACO) for facility layout problem," in *2017 16th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, Sep. 2017, pp. 1–8.

[24] M. A. Mohammed and R. A. Hasan, "Particle swarm optimization for facility layout problems FLP—A comprehensive study," in *2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 93–99, 2017.

[25] R. A. Hasan *et al.*, "HSO: A hybrid swarm optimization algorithm for reducing energy consumption in the cloudlets," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 16, no. 5, pp. 2144–2154, 2018.

[26] R. A. Hasan, S. N. Shahab, and M. A. Ahmed, "Correlation with the fundamental PSO and PSO modifications to be hybrid swarm optimization," *Iraqi Journal for Computer Science and Mathematics*, vol. 2, no. 2, pp. 25–32, 2021.

[27] R. A. Hasan *et al.*, "Hybrid spotted hyena based load balancing algorithm (HSHLB)," *Mesopotamian Journal of Big Data*, p. 199–210, 2024, https://doi.org/10.58496/MJBD/2024/014

[28] W. A. Hammood *et al.*, "Data analysis of an exploring the information systems success factors for early warning systems adoption," *Babylonian Journal of Machine Learning*, pp. 102–111, 2024, https://doi.org/10.58496/BJML/2024/010

[29] H. R. Ibraheem *et al.*, "A new model for large dataset dimensionality reduction based on teaching learning-based optimization and logistic regression," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 3, pp. 1688–1694, 2020.

[30] A. H. Ali *et al.*, "Large scale data analysis using MLlib," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 19, no. 5, pp. 1735–1746, 2021.

[31] A. H. Ali, R. A. Alhayali, M. A. Mohammed, and T. Sutikno, "An effective classification approach for big data with parallel generalized Hebbian algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 6, pp. 3393–3402, 2021.

[32] M. Aljanabi, M. G. Yaseen, A. H. Ali, and M. A. Mohammed, "Prompt engineering: Guiding the way to effective large language models," *Iraqi Journal for Computer Science and Mathematics*, vol. 4, no. 4, pp. 151–155, 2023.

[33] M. Aljanabi *et al.*, "Distributed denial of service attack defense system-based auto machine learning algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 1, pp. 544–551, 2023.

[34] N. M. Hussien *et al.*, "A smart gas leakage monitoring system for use in hospitals," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 2, pp. 1048–1054, 2020.

[35] M. Aljanabi, M. A. Ismail, R. A. Hasan, and J. Sulaiman, "Intrusion detection: A review," *Mesopotamian Journal of CyberSecurity*, pp. 1–4,2021, https://doi.org/10.58496/MJCS/2021/001

[36] R. A. Hasan, S. N. Shahab, and M. A. Ahmed, "Correlation with the fundamental PSO and PSO modifications to be hybrid swarm optimization," *Iraqi Journal for Computer Science and Mathematics*, vol. 2, no. 2, pp. 25–32, 2021.

[37] A. H. Ali *et al.*, "Big data classification based on improved parallel k-nearest neighbor," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 21, no. 1, pp. 235–246, 2023.

[38] A. S. Abdalzahra, H. W. Abdulwahid, and R. A. Hasan, "Using ideal time horizon for energy cost determination," *Iraqi Journal for Computer Science and Mathematics*, vol. 2, no. 1,2024, https://doi.org/10.52866/ijcsm.2021.02.01.002

[39] M. A. Mohammed *et al.*, "The effectiveness of big data classification control based on principal component analysis," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 1, pp. 427–434, 2023.

[40] S. I. Jasim, M. M. Akawee, and R. A. Hasan, "A spectrum sensing approaches in cognitive radio network by using cloud computing environment," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 750–757, 2022.

[41] H. D. K. Al-janabi, and R. A. H. Al-Bukamrh, "Impact of light pulses generator in communication system application by utilizing Gaussian optical pulse," in *2019 22nd International Conference on Control Systems and Computer Science (CSCS)*, May 2019, pp. 459–464, May 2019.

[42] M. M. Akawee, M. A. Ahmed, and R. A. Hasan, "Using resource allocation for seamless service provisioning in cloud computing," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 2, pp. 854–858, 2022.

[43] A. S. T. Hussain *et al.*, "Unlocking solar potential: Advancements in automated solar tracking systems for enhanced energy utilization," *Journal of Robotics and Control*, vol. 5, no. 4, pp. 1018–1027, 2024.