Classification of Arabic text according to feature vectors and labels using several models

Researchers:

Sara Mahmoud Alsukhni (2023751005)

Abstract --This paper investigates the realm of Arabic text classification, aiming to evaluate and compare the efficacy of various machine learning models in handling this linguistic domain. The study undertakes a systematic exploration, testing multiple models including Naive Bayes, Logistic Regression, and Linear SVC, on a corpus of Arabic texts. Through meticulous experimentation and evaluation, it is revealed that different models exhibit diverse performances in accurately categorizing Arabic texts. Logistic Regression and Linear SVC emerged as the top-performing models, each achieving an accuracy of 98%, signifying their robustness in effectively classifying Arabic texts. The findings underscore the significance of model selection in Arabic text classification tasks, shedding light on the models that excel in this field and paving the way for enhanced natural language processing applications tailored to the Arabic language.

Keywords-- Natural Language Processing (NLP), Arabic text classification, text classification, Machine Learning Models, Naive Bayes, Linear SVC (Support Vector Classifier).

1. Introduction

Arabic, as one of the world's prominent languages, holds significant importance in the digital era, especially considering the exponential growth of online Arabic content. The classification of Arabic texts stands as a pivotal research domain within natural language processing (NLP) and data mining. It involves organizing, categorizing, and assigning labels to vast volumes of Arabic textual data, a task fraught with complexities due to the unique linguistic characteristics inherent in the Arabic language.

Arabic texts classification is one of the crucial topics in the field of mining Arabic texts on a large scale. This is attributed to the continuous increase in the growth of Arabic content on the internet. The classification of Arabic texts is a significantly important research topic as it involves extracting high-quality information from texts and identifying the topics to which those texts belong, particularly when these texts are large in volume and cannot be manually classified. Current research problems revolve around the vast number of Arabic texts available on the internet, which continues to grow, and researchers are working to solve these problems and utilize this data by applying data mining techniques specific to their classification.

Text classification is considered as one of the earliest techniques used in statistical approaches based on Natural Language Processing (NLP) [1]. Text classification is the process of grouping texts into one or more predefined categories based on their content [2]. Arabic text classification using AI models falls within the realm of natural language processing (NLP), dedicated to organizing, categorizing, and labeling Arabic textual data. The emergence of artificial intelligence technologies, notably machine learning and deep learning, has significantly propelled this field by automating the analysis of large volumes of Arabic text. AI models specialized in Arabic text classification employ algorithms to comprehend linguistic patterns and meanings, striving to automatically categorize Arabic text into predetermined groups.

The preprocessing stage involves segmenting Arabic text into coherent segments, with labeled datasets serving as the training material for models. This training allows models to discern connections between text attributes and categories. However, the complexity of Arabic language, limited labeled datasets, and the necessity for precise model training using relevant Arabic text pose substantial challenges.

most studies have focused on common text genres, like SMS, book reviews, social media, and so forth [10]. This research aims to develop AI models specifically tailored for Arabic text classification. The study's objectives encompass an examination of existing techniques, the refinement of models to address Arabic language intricacies, and an evaluation of model performance against established benchmarks, emphasizing accuracy, scalability, and adaptability.

The applications of AI-driven Arabic text classification span various domains, including sentiment analysis, topic modeling, information retrieval, and spam detection. Its implementation has notably enhanced the efficiency of search engines, recommendation systems, and the automated processing of Arabic text. Ongoing advancements suggest continual improvement in navigating and comprehending the landscape of Arabic text across digital platforms.

2. Literature Review

Text classification has found extensive application across various Natural Language Processing (NLP) scenarios. Despite the considerable body of research in the field, there has been a notable dearth of work specifically addressing the Arabic language. This section will delve into a review of the limited research conducted in Arabic literature, shedding light on the existing gaps and opportunities for further exploration in this linguistic context.

Sara A. Aboalnaser. [1] give an overview of the text classification procedure, especially for documents in Arabic. Additionally, it displays the most significant methods employed in the classifying process of documents. Moreover, the outstanding problems pertaining to Arabic they talk about text classification and thereby emphasize the future suggestions in this particular scenario.

Authors in [3] used a big dataset, The purpose of this work is to present experimental analyses of six popular models for classifying a sizable Arabic corpus. Nave Bayes (NB), Random Forest, Support Vector Machine (SVM), Logistic Regression, Decision Tree (DT), and Stochastic Gradient Descent (SGD) are these models. They made use of a corpus of 111,728 Arabic documents divided into five groups: news, sports, culture, economics, and diverse. To assess each model's experimental outcomes, three performance indicators were used. According to the experimental findings, the models with the best weighted F1 score include the Logistic Regression model, SGD, SVM, NB, Random Forest, and DT. Additionally, the results demonstrate that feature size and corpus size both have high.

Another study in [2] Fawaz AL Zaghoul and Sami Al-Dhaheri showcase and examine the application's outcomes Utilizing Artificial Neural Networks (ANN) to categorize Arabic textual materials. The principal supply source of understanding is an Arabic text categorization (TC) corpus created locally at the University of Jordan and accessible serves as the main knowledge source for the system. The ANN model is built and tested using this corpus. Techniques for allocating feature reductions and weights that correspond to the relative importance of each phrase are covered. The term weighting method represents each Arabic document. The most pertinent features for the categorization have been chosen using feature reduction techniques due to the large number of unique words in the collection set. The experimental findings demonstrate that an ANN model employing features reduction techniques outperforms a simple ANN in terms of classification performance.

Leen Al Qadi et al. [4] goal is to automatically classify a document according to its linguistic characteristics. In order to accomplish this, they created a brand-new dataset with 90k Arabic news items from Arabic news portals tagged with relevant information. The Arabic computational linguistics research community will have unrestricted access to the dataset. There are four primary categories in the dataset: Middle East, Technology, Sports, and Business. Every article that was gathered was free of stop words, punctuation, Latin characters, and numbers. They employed a variety of traditional supervised machine learning classifiers to examine the dataset's efficacy. Specifically, the ten most widely used classifiers were as follows: XGBoost classifier, Random Forest classifier, Multinomial classifier, K-nearest neighbors (KNN), Logistic Regression, Nearest Centroid, Decision Tree (DT), Support Vector Machines (SVM), Multi-Layer

Perceptron (MLP) and Ada-Boost Classifier. They used an ensemble model to merge the best classifiers into a majority-voting classifier to achieve high accuracy. test results demonstrated strong performance, with Ada-Boost achieving a minimum F1-score of 87.7% and SVM achieving a maximum performance of 97.9%. They present the experiment results in terms of accuracy, F1-scores, and confusion matrices. Arabic Text Classification; Shallow Learning Classifiers; Arabic Dataset; Single-Label Classification are index terms.

Authors in [5] suggested the use of artificial neural networks for the classification of Arabic-language documents. Automatically categorizing ANN for Arabic documents has not been investigated.in depth thus far. This work uses an Arabic corpus as used in the ANN model's construction and testing. Techniques a representation of the document, allocating weights that show how important each term is discussed. The term weighting is used to represent each Arabic document. As the quantity of distinct words in The Singular Value is large in the collection set. Decomposition (SVD) has been employed to determine which attributes that are most pertinent to the classification. The. According to experimental findings, an ANN model using The SVD attains 88.33%, surpassing the basic ANN's performance, which results in 85.75%.

Another study in [6] pointed to the challenge of Arabic multi-label text classification by proposing and evaluating deep learning models. Specifically, Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) are utilized to build two models in Python. The objective is to classify Arabic news, enabling users to access and display the most relevant news based on their interests. The research involves cleaning the data to enhance experimental data quality. The results indicate that the LSTM model achieved a test data accuracy of 82.03%, while the MLP model achieved a slightly lower accuracy of 80.37%. Evaluation of both models was conducted using the F1 score. This study showcases the effectiveness of employing deep learning techniques, specifically LSTM and MLP, in addressing the challenges of Arabic multi-label text classification.

Authors in [7] focused on feature selection as a crucial step to enhance classification performance. The study introduces a new feature selection method, referred to as ImpCHI, particularly when employing light stemming. ImpCHI is an enhancement of the chi-square method, known for its effectiveness in feature selection. The evaluation utilizes a corpus comprising 250 Arabic documents classified into five classes: art and culture, economics, politics, society, and sport. The experimental results reveal that Arabic text classification, incorporating ImpCHI as a feature selection method, outperforms the use of chi-square, specifically in terms of recall measures. This suggests that ImpCHI holds promise as an improved feature selection technique for enhancing the accuracy of Arabic text classification.

Another study in [8] discussed the Arabic text classification (ATC), also known as Arabic text categorization, focusing on the assignment of categories to Arabic documents based on their content. The study empirically evaluates five classification models, namely Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbor (KNN), and Logistic Regression (LR), utilizing two Arabic datasets (CNN Arabic and osac_uft8). Three feature vectorization methods—word count, Terms Frequency-Inverse Document Frequency (TF-IDF), and word embedding using word2vec—are applied to convert text into numeric vectors. The experimental results reveal that SVM and LR classifiers demonstrate the highest performance, followed by RF, KNN, and DT. Additionally, the study indicates that feature vectorization methods and dataset size significantly impact the performance of RF, KNN, and DT, while SVM and LR maintain stable outcomes. This comprehensive examination of classification models and feature vectorization methods contributes valuable insights into optimizing Arabic text classification systems.

Authors in [9] focused on improving the accuracy of Arabic text classification through innovative methods. The proposed approach involves utilizing an Arabic stemming algorithm for feature extraction, selection, and reduction. The Term Frequency-Inverse Document Frequency technique is then applied as a feature weighting method. For the classification step, the study employs Convolutional Neural Networks (CNN), a powerful algorithm in image processing and pattern recognition, but less commonly used in text mining. Through this combination and hyperparameter tuning in the CNN algorithm, the paper achieves excellent results on multiple benchmarks, addressing the scarcity of studies in categorizing and classifying Arabic text.

Another study in [10] Arabic poetry texts are categorized. To increase the efficiency of the model, a clustering technique is combined with a customized feature selection. Two well-known machines have been used in experiments with deep learning. strategies, such as decision trees and support vector machines. The suggested feature extraction technique has produced good accuracy using each of the three methods.

Study Reference	Model/Method	Dataset	Performance/Results
[1]		data is almost labelled by experts to different categories such as politics, health, economy, news…. etc.	Discusses the need for research in Arabic text classification and proposes future research directions.
[2]	Naive Bayes, Random Forest, SVM, Logistic Regression, DT, SGD	collection of Arabic texts, which covers modern Arabic language used in newspapers articles.	Naive Bayes 94%, Random Forest 92%, SVM 95%, Logistic Regression 96%, DT 89%, SGD 95%
[3]	Artificial Neural Networks (ANN)	they have picked a sub set of Arabic articles covering different topics from an Arabic corpus built locally at the University of Jordan	96%,
[4]	XGBoost, Random Forest, Multinomial NB, KNN, Logistic Regression,	They collected the proposed dataset using web scraping	AdaBoost at 87.7% F1-score and SVM at 97.9% performance.

	Nearest Centroid, DT, SVM, MLP, AdaBoost	(Python Scrapy), from seven popular news websites (beIN sports. com, tech-wd.com, skynewsarabic.com, Arabic.rt.com, cnbcarabia.com, arabic.cnn.com and youm7.com)	
[5]	Artificial Neural Networks (ANN)	is a set of prophetic (says of prophet “peace be upon him”) collected from the Prophetic encyclopedia.	ANN with Singular Value Decomposition (SVD) attained 88.33% accuracy, surpassing basic ANN's performance of 85.75%.
[6]	Multilayer Perceptron (MLP), LSTM-based RNN	collected from the following ten categories: Arts & Celebrities, Economy, Kitchen, Technology, Islam & Religions, News, Sports, Health, Weather, and others. Moreover, we used the collected data from Mowjaz Multitopic Labelling Task competition	LSTM model achieved 82.03% test data accuracy, while MLP model achieved 80.37%. F1 score used for evaluation.
[7]	Imp CHI, Chi-square	an Arabic text corpus is collected from the Arabic online newspapers Hespress1 and Hesport2 .	Imp CHI outperformed chi-square in feature selection.
[8]	SVM, DT, RF, KNN, LR, Word Count, TF-IDF, Word Embedding	The dataset "CNN Arabic" is a collection of Arabic texts, it consists of 5.070 documents	SVM and LR demonstrated highest performance, while impact of dataset size varied for RF, KNN, and DT.
[9]	Arabic stemming, CNN	collected from 3 Arabic online newspapers: Assabah (www.assabah.ma),Hespress (www.hespress.com) and Akhbarona (www.akhbarona.com)	Improved accuracy using Arabic stemming, TF-IDF, and CNN in categorizing Arabic text.

[10]	Deep Learning, Decision Trees, Support Vector Machines	Arabic poems	Decision Tree 95.45% Deep Learning 95.56% Support Vector Machine 96.34%

Table 1 Summarize of Related Works.

3. Methodology

In this study, we present a classification system for Arabic language texts that is classified into cultural, political, mathematical, financial, religious, technical, or medical. **Figure 1** shows the architecture of the proposed system.

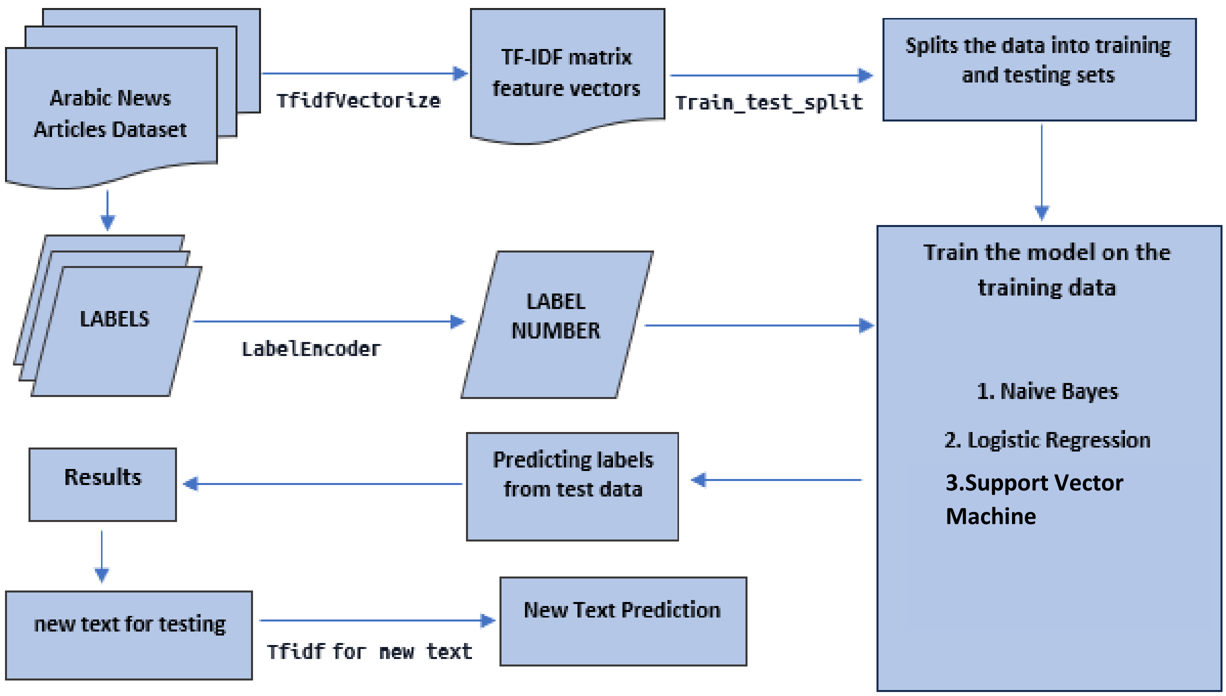

Figure 1. The Architecture of Arabic text classification.

We will explain the stages that the proposed system went through, starting with the data collection and ending with prediction in text classification.

3.1 Dataset

The data set used is a data set published on Kaggle and includes seven classifications, and each classification contains texts as follows (cultural, financial, medical, political, religious, sports, technical), and each classification contains 6,500 documents, a total of 45,500 documents for all classifications.

To explain more about the division of data and Classifications, **Table 1** shows the number of documents in the category, the number of categories and the type of documents in all categories.

Table 1. The number of documents in the dataset classifications

Classifications	Number of documents
Culture	6,500
Culture	6,500
Medical	6,500
Politics	6,500
Religion	6,500
Sports	6,500
Tech	6,500
Total	**45,500**

3.2 TF-IDF matrix

It is an important part of the process of analyzing and classifying texts. The goal of this stage is to transform texts into a matrix of numbers that reflects their representation in the feature space in a way that reflects the importance of words in the document based on their apparent distribution in the text.

3.3 Labels

Through it, categories are converted into numbers, because some models do not deal well with categorical data represented by texts

3.4 Splits the data into training and testing sets

At this stage, the data set is divided into a training set and a test set, the training set to train the models on, and the test set to evaluate the model's performance.

3.5 Train the model on the training data

In this training phase, trends are directed to the training data set, so the links used are six models (Naive Bayes, Logistic Regression, Support Vector Machines).

3.6 Predicting labels from test data

At this stage, the performance of the models is evaluated through a set of tests to reach results and determine the effectiveness of the proposed system.

3.7 New text for testing

At this point, new text is input from outside the dataset, and the system is checked by predicting the classification of the new text based on the labels in the trained dataset.

3.8 Naive Bayes

The Naive Bayes model is a prominent tool in the realm of Arabic text classification. This model relies on the concept of simple Bayesian theory, calculating the probability of classifying a text into a specific category based on the presence of certain words. The model assumes independence among words in the texts, making it straightforward to understand and implement.

Naive Bayes excels in handling linguistic data and demonstrates good performance in classifying Arabic texts. Widely used in applications requiring text classification, such as linguistic analysis of social media and email filtering, Naive Bayes proves to be an effective tool for efficiently examining and categorizing linguistic data.

3.9 Logistic Regression

Logistic Regression is a statistical method used for binary and multiclass classification problems. Despite its name, it's mainly employed for classification tasks. It models the probability of the occurrence of an event by fitting data to a logistic function, allowing predictions between 0 and 1. In the context of text classification, Logistic Regression can be effective in determining the likelihood of a document belonging to a specific category.

3.10 Linear SVC (Support Vector Classification)

Linear SVC is a type of Support Vector Machine designed for classification tasks. It works by finding the hyperplane that best separates classes in the feature space. SVC is effective in high-dimensional spaces, making it suitable for text classification where the number of features (words) can be significant. Linear SVC performs well when there is a clear margin of separation between classes.

4. Results

The evaluation of model performance involves subjecting the dataset to scrutiny using 5 ML classification algorithms. To gauge the effectiveness of these algorithms, a set of comprehensive evaluation criteria will be employed. This section will begin by delving into the intricacies of the performance metrics associated with each algorithm. Subsequently, the outcomes derived from the application of these algorithms will be meticulously presented and analyzed. The results will be displayed as A, which displays the results of testing on the data set, and B, which displays the results of testing texts from outside the data set.

4.1 Performance setrics

In this section, the evaluation criteria used to compare the performance of the classification algorithms will be explained. Information will be given about the criteria that show the accuracy performance of the algorithms and the criteria that show the error performance. The main parameters used to calculate evaluation criteria are True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN).

True positive (TP): The instances where the model correctly predicts the class of a text sample.

True negative (TN): The instances where the model correctly excludes a text sample from being classified into a particular class.

False positive (FP): The instances where the model incorrectly predicts a text sample as belonging to a specific class when it actually belongs to another class.

False negative (FN): The instances where the model incorrectly fails to predict a text sample as belonging to a certain class when it actually does belong to that class.

4.1.1 **Accuracy:** is considered the most common metric, where it represents the number of correct predictions over the total number of observations, as illustrated by the following equation:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}$$

4.1.2 **Precision (p):** It is the number of correctly classified for the positive value divided by the total number of values classified positively.

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

4.1.3 **4.1.3 Recall (r):** is the correctly predicted positive class over all actual classes whether the disease absence or presence.

$$r = \frac{TP}{TP + FN}$$

4.1.4 **F1-score:** It is a combination of recall and precision as illustrated in the following equation:

$$F1 = \frac{2pr}{P + R}$$

a- Test set results

In this study, several machine learning models were trained and evaluated for their performance on a specific task. The Naive Bayes classifier demonstrated an accuracy of 96%, showcasing its effectiveness in classification tasks despite its simplicity. Logistic Regression exhibited a slightly higher accuracy of 98%, indicating its superior performance in this particular context. On the other hand, Linear Support Vector Classifier (Linear SVC) achieved an accuracy of 98%, mirroring the performance of Logistic Regression. The **Table 2** shows the models and accuracy resulting from the testing process

Table 2. Accuracy of the ML models used in the study.

Model	Accuracy
Naive Bayes	96%
Logistic Regression	**98%**
Linear SVC	**98%**

These findings underscore the varying degrees of effectiveness among the employed models, with Logistic Regression and Linear SVC displaying notable strengths in accurate classification based on the metrics evaluated in this study. **Figure 2** shows the accuracy results for the 3 models used in the study.

Figure 2. The difference between the accuracy results through the chart.

Below, Figures 3 - 7 show the performance criteria for the 3 ML models used in the study, showing the criteria for the categories into which the data set is divided. **Figure 3** displays the performance criteria according to the Naive Bayes model, **Figure 4** the logistic regression model, and **Figure 5** the linear SVC model.

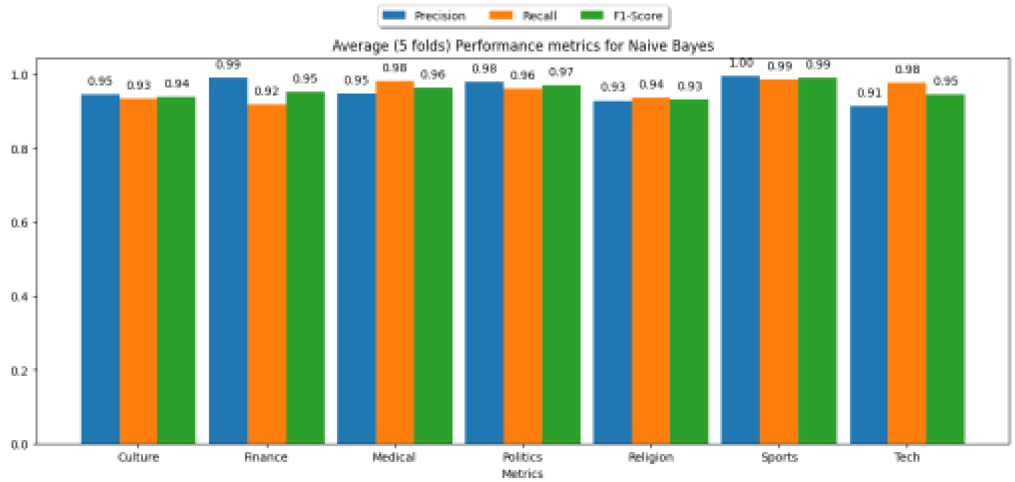

Figure 3. Performance criteria for the Naive Bayes model

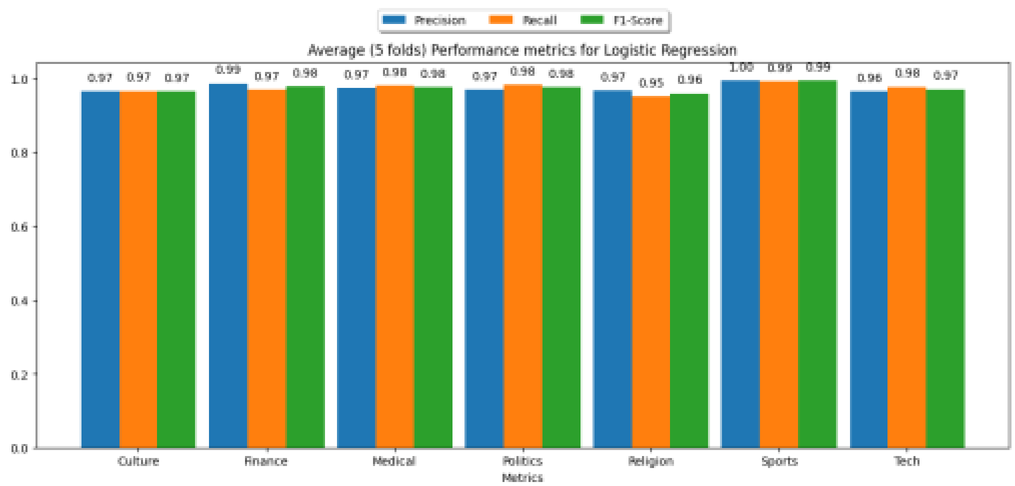

Figure 4. Performance criteria for the logistic regression model.

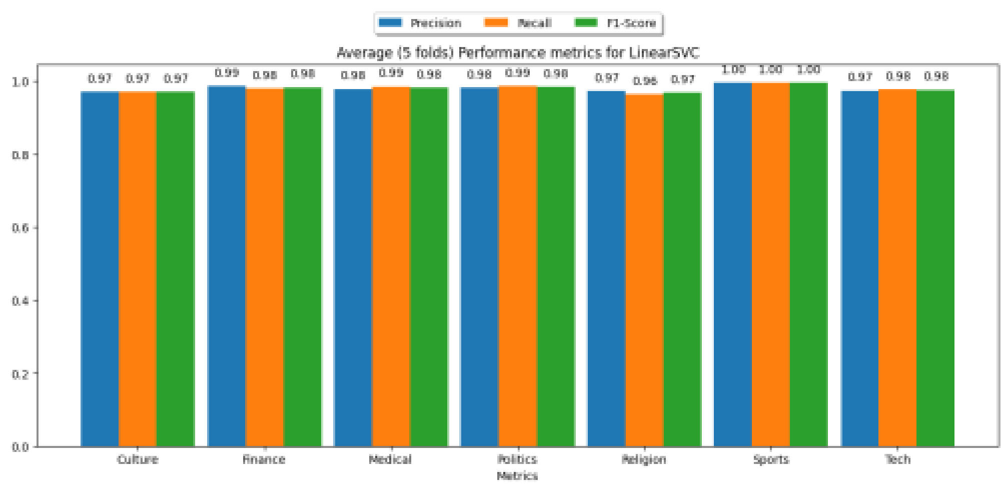

Figure 5. Performance criteria for linear SVC model.

b- Results of texts from outside the data set (New text stage)

At this level, it is for testing texts from outside the data set, which proves that our system can predict texts from outside the data set. We used it as a sample or test data set for the classification process, and here is a sample text for testing purposes, and these were our results:

- **Input testing text regarding medical topics:**

في مستشفى التخصصي االردني يُعنى تخصص طب األطفال بتقديم الرعاية الطبية لألطفال وتشخيص أمراضهم، وعالجهم
حديثي الوالدة أم ال؛ نحو معني بتقديم الرعاية والعناية الطبية والصحية لإلنسان في المرحلة ما بين سن الطفولة، .سوا ًءكانوا "
وحتى بلوغ سن الرشد.

```
التصنيف المتوقع باستخدام نموذج Naive Bayes: ['Medical']
التصنيف المتوقع باستخدام نموذج Logistic Regression: ['Medical']
التصنيف المتوقع باستخدام نموذج LinearSVC: ['Medical']
```

Figure 3 image for new text prediction for 3 models.

- **Input testing text regarding religion topics:**

اقتضت حكمة هللا تعالى أن يرسل في كل أمة نذيرا ً، يوضح لهم ما أنزل هللا لعباده من الدين والهدى، الذي به صالح حالهم في
دنياهم وأخراهم، قال تعالى: { وَ ن أُ مَّة إ لا خَلَا فِيهَا نَذِير }]فاطر: 24[.

```
التصنيف المتوقع باستخدام نموذج Naive Bayes: ['Religion']
التصنيف المتوقع باستخدام نموذج Logistic Regression: ['Religion']
التصنيف المتوقع باستخدام نموذج LinearSVC: ['Religion']
```

Figure 4 image for new text prediction for 3 models.

- **Input testing text regarding sports topics:**

أعلن الفيصلي عن بداية قوية في الموسم الحالي، عندما استحوذ على أول األلقاب المحلية بتتويجه بلقب درع االتحاد األردني لكرة
القدم.وسجل الفيصلي أرقاما ً قياسيا ً في بطولة الدرع أعطت مؤشرات بأن القادم سيكون أفضل، لكن عاصفة تغيير
المدربين سرعان ما أثرت على أداء ونتائج الفريق.

```
التصنيف المتوقع باستخدام نموذج Naive Bayes: ['Sports']
التصنيف المتوقع باستخدام نموذج Logistic Regression: ['Sports']
التصنيف المتوقع باستخدام نموذج LinearSVC: ['Sports']
```

Figure 3 image for new text prediction for 3 models.

5. Conclusion

In conclusion, the exploration and evaluation of various machine learning models for Arabic text classification revealed nuanced performance variations among the models. The study systematically tested and compared models such as Naive Bayes, Logistic Regression, and Linear SVC on an Arabic text corpus. The results demonstrated that while each model exhibited distinct strengths and weaknesses, certain models notably excelled in accurately categorizing Arabic texts.

Logistic Regression and Linear SVC emerged as the standout performers, achieving an impressive accuracy rate of 98%. Their robust performance underscores their suitability and effectiveness in handling Arabic text classification tasks.

These findings emphasize the critical role of model selection in Arabic text classification. The identified top-performing models showcase their potential for enhanced natural language processing applications tailored to the Arabic language. Moving forward, leveraging these models can significantly impact various NLP tasks involving Arabic texts, paving the way for more accurate and efficient language processing solutions in the Arabic linguistic domain.

6. References

[1] S. A. Aboalnaser, "Machine learning algorithms in arabic text classification: A review," *Proc. - Int. Conf. Dev. eSystems Eng. DeSE*, vol. October-2019, pp. 290–295, 2019, doi: 10.1109/DeSE.2019.00061.

[2] F. A. L. Zaghoul and S. Al-Dhaheri, "Arabic text classification based on features reduction using artificial neural networks," *Proc. - UKSim 15th Int. Conf. Comput. Model. Simulation, UKSim 2013*, pp. 485–490, 2013, doi: 10.1109/UKSim.2013.135.

[3] M. A. H. Madhfar and M. A. H. Al-Hagery, "Arabic text classification: A comparative approach using a big dataset," *2019 Int. Conf. Comput. Inf. Sci. ICCIS 2019*, pp. 1–5, 2019, doi: 10.1109/ICCISci.2019.8716479.

[4] L. Al Qadi, H. El Rifai, S. Obaid, and A. Elnagar, "Arabic text classification of news articles using classical supervised classifiers," *2019 2nd Int. Conf. New Trends Comput. Sci. ICTCS 2019 - Proc.*, pp. 1–6, 2019, doi: 10.1109/ICTCS.2019.8923073.

[5] F. Harrag and E. El-Qawasmah, "Neural network for Arabic text classification," *2nd Int. Conf. Appl. Digit. Inf. Web Technol. ICADIWT 2009*, pp. 778–783, 2009, doi: 10.1109/ICADIWT.2009.5273841.

[6] B. Alsukhni, "Multi-Label Arabic Text Classification Based on Deep Learning," *2021 12th Int. Conf. Inf. Commun. Syst. ICICS 2021*, pp. 475–477, 2021, doi: 10.1109/ICICS52457.2021.9464538.

[7] S. Bahassine, A. Madani, and M. Kissi, "An improved Chi-sqaure feature selection for Arabic text classification using decision tree," *SITA 2016 - 11th Int. Conf. Intell. Syst. Theor. Appl.*, pp. 1–5, 2016, doi: 10.1109/SITA.2016.7772289.

[8] T. Sabri, O. El Beggar, and M. Kissi, "Comparative study of Arabic text classification using feature vectorization methods," *Procedia Comput. Sci.*, vol. 198, no. 2021, pp. 269–275, 2021, doi: 10.1016/j.procs.2021.12.239.

[9] A. Elnagar, R. Al-Debsi, and O. Einea, “Arabic text classification using deep learning models,” *Inf. Process. Manag.*, vol. 57, no. 1, pp. 103–114, 2020, doi: 10.1016/j.ipm.2019.102121.

[10] F. Alqasemi, S. Al-Hagree, N. A. Ali Abdu, B. Al-Helali, and G. Al-Gaphari, “Arabic Poetry Meter Categorization Using Machine Learning Based on Customized Feature Extraction,” *Int. Conf. Intell. Technol. Syst. Serv. Internet Everything, ITSS-IoE 2021*, pp. 1–4, 2021, doi: 10.1109/ITSS-IoE53029.2021.9615302.