International Journal of Engineering & Technology, 12 (2) (2023) 126-137

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Automatic Classification of Arabic News and Column Articles using Machine Learning and Deep learning Approaches

Hanen Himdi[1*], Bayan Alotaibi[2], Dania AlSahafy[3], Gharam AlGhamdi[4] and Layan AlGhamdi[5]

[1]*Computer Science and Artificial Intelligence Department, College of Computer Science and Engineering, University of Jeddah, Saudi Arabia; hthimdi@uj.edu.sa*
[2]*bayanalotaibi@gmail.com*
[3]*dhosain0001.stu@uj.edu.sa*
[4]*gharamm1912@gmail.com*
[5]*lalghamdi0058.stu@uj.edu.sa*
[*]*Hanen Himdi E-mail: hthimdi@uj.edu.sa*

Abstract

News is a report of recent events that is distributed from its point of origin to receivers by journalists who work for commercial organizations. Columns, on the other hand, are writings presented in special sections on news platforms that provide content on a variety of subjects and are often written by individuals or small groups of authors. Usually, columns might present information in a more subjective manner compared to the objective constraints in news reports. However, a major concern is that many columns lack the news's editorial oversight process presented in news production, which can have negative effects, such as including incorrect information. In some cases, the presence of column articles on news platforms is mistaken for a news article, causing credibility concerns. To tackle this problem, a wide range of extant studies conducted in the recent past offer suitable techniques for column article classification. However, there are no studies in Arabic for this purpose. In this study, we introduce the first Arabic column article dataset that includes more than 12k articles. Then, we compiled several classification models using machine learning and deep learning approaches. It was found that deep learning models, CNN-LSTM, trained by BERT achieved the highest accuracy, reaching 96.6%. Finally, we propose a web platform that can be used freely to classify news and column articles based on their textual content solely

1. Introduction

A successful column may serve as the foundation for a whole magazine. For example, when Cyrus Curtis established the Tribune and Farmer in 1879, it was a four-page weekly with a 50-cent yearly subscription fee. In 1883, he decided to create a separate monthly supplement, Ladies Journal and Practical Housekeeper, edited by his wife, Louise Knapp Curtis. Curtis sold Tribune and Farmer to focus his efforts on the new journal, which became the Ladies Home Journal, which had 25,000 subscribers by the end of its first year [30]. Nowadays, the Internet has provided a valuable platform for news and column writers to obtain continual data, thus writing well-organized articles [1]. However, these articles may lack factual basis. Though both news and columns might seem to have the same genre, they are written with totally different intentions.

News can be defined as a report of recent events or material reported in newspapers or newscasts [38]. According to [15], 'news is a report of what a news organization has recently learned about matters of some significance or interest to the specific community that news organization serves.' Meanwhile, [16] defined news as 'information which is transmitted from the source to recipients by journalists who are both employees of bureaucratic, commercial organizations and also members of a professional group'.

Unlike news articles, columns refer to a specific section in a publication that provides the space or platform for writers and experts to present their views and opinions. For instance, experts from outside the newspaper or publication house, as well as in-house journalists, both contribute to the columns. Almost any topic can be the subject of a column. It offers analysis and opinions on a topic and usually does so while presenting arguments and supporting data. Normally, a column is published on a regular basis in a publication, such as a newspaper or a periodical. It acts as a platform that allows professionals like scientists, engineers, politicians, writers, or social commentators a forum to

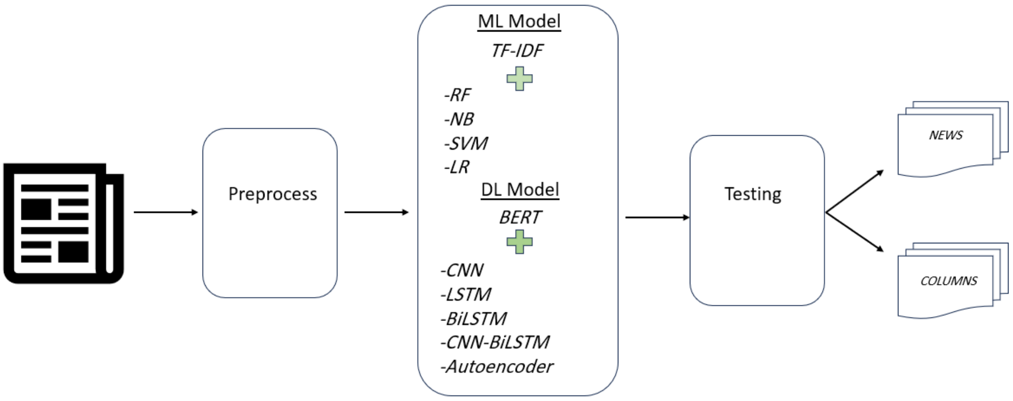


Figure 1: News and columns classification model.

weigh in on a topic. Columns are frequently named "Opinion", "Op-Ed", and "Editorial" pieces because they may simultaneously publish opposing viewpoints if it is intended to portray different points of view [27].

Concurrently, a study by [31] defines a column as a section of a newspaper, magazine, journal, or other serial publication where a particular writer regularly expresses opinions, provides analysis, or discusses a certain subject. Columns are distinct from other types of journalism. Individual columns, however, might frequently mimic and resemble other types of articles. There are no hard and fast rules, but normally, columns describe any of these attributes: those having a singular voice and are well-written and topic-focused or articles published regularly and supervised by an editorial filter from the parent publication. It also may include written pieces by a hired or handpicked writer and form a part of a larger publication where a subscription is possibly required before having access to them.

There are significant differences between news and column articles or content. The most common attributes of news articles are their objectivity and unbiased written form. The aim is to present factual information about recent events or developments [9]. News articles heavily rely on citing credible sources, such as experts, eyewitnesses, or official statements, to support the information provided. Additionally, news articles are time-sensitive and focus on delivering up-to-date information [35]. They also follow a standardized structure, with the most crucial details presented at the beginning and additional information provided in subsequent paragraphs. Editorial neutrality is also maintained in news articles, in that they strive to be free from personal opinions and biases and aim to inform the reader [9, 35].

On the contrary, columns are opinion and commentary pieces, and such articles are subjective in nature, offering the personal opinions, insights, or commentary of the author [25, 9]. The tone and tenor are also identifiable only to the author's perspective on a particular issue or topic [35]. While column articles may reference facts and information, they do not rely as heavily on citing sources as news articles. The emphasis is on the author's viewpoint. The topics also can cover a wide range of subjects, including politics, culture, lifestyle, and personal experiences. They may not always focus on breaking news [9]. As such, opinions in columns should be expected. Moreover, readers should understand that column articles are meant to reflect the author's opinions and may not always present a balanced view of the topic [25, 9].

The majority of people might believe articles that appear in any columns as news for a variety of reasons. The most common being misidentification. For instance, readers may mistakenly identify a column as a news article, especially if the publication does not clearly label the content. This can lead to the incorrect belief that the column represents objective reporting [25]. Trust in the publication often makes people assume that all content from that outlet is objective, even if it includes columns with subjective viewpoints. This also leads to bias. For example, if a column aligns with a reader's preconceived notions, they may accept it as factual news, even if it contains opinions. Moreover, a lack of media literacy can make it difficult and challenging to differentiate between news and opinion pieces because they may not be aware of the distinctive characteristics of each type of content [25, 9, 35]. Moreover, in the age of social media, content can be easily shared without context or verification. If a column is shared on social platforms without a clear indication that it's an opinion piece, it can be mistaken for news by those who see it in their feeds [25, 26]. People only have a limited amount of time, and when they focus more of that time on columns, they may reduce time spent on learning the facts. Without this basic factual grounding, people are less able to think for themselves or to think critically about the opinions and analyses they read. Making things worse, if columns regularly attract more readership than news articles, newspapers and news channels may reallocate time and resources toward opinion pieces, detracting from the quality and quantity of news publication [9].

To avoid believing column content as news, it is crucial for readers to develop strong media literacy skills. This includes checking the source, looking for attribution and citations, and being aware of the publication's overall editorial stance. As stated by [35], clear labeling by media outlets can help differentiate between news articles and opinion pieces. Perhaps drawing a sharper line between news and opinions would increase public scrutiny of news outlets because people would have greater expectations that their newspapers and news channels will report the facts alone [9].

In recognition of the fact that machine learning and deep learning models have substantially improved the efficacy of text classification models, we plan to investigate how they can be utilized in the classification of Arabic news and blogs. There has been a wide range of studies that offer approaches to column classification [8, 23] or extract information from columns in English [13, 19]. However, to the best of our knowledge, there have been no studies that tend to automatically classify news and columns in Arabic. Our study involved the task of automating the classification of an article as either news or column by utilizing machine learning and deep learning approaches with innovative word embeddings such as the pre-trained Arabic BERT model [6]. Unlike previous studies that investigated credibility methods in columns through several metadata associated with the article, the novelty of our work is that we rely only on the textual content of the article to predict the correct class of the article. Several machine learning and deep learning models were trained on this data, with the aim of achieving high accuracy, precision, and recall in classifying news and column articles. Figure 1 shows the models' framework.

The contributions of this paper are as follows:

- Compile the first Arabic columns and news articles dataset, which includes 12k columns and 15k news articles. The dataset can be used to train classification models for similar tasks.
- Compose several classification models that enable to differentiate between news and column articles, using machine learning and deep learning approaches.
- Conduct a comprehensive set of tests and evaluations to compare the performance of the composed classification models.
- Develop a web platform that automates the classification of news and column articles, from the optimal classification model in this study.

The remainder of the paper is organized as follows: Section 2 demonstrates the previous works related to this topic, then Section 3 describes the dataset, pre-processing tasks, and the details of developing the classification models. Moreover, the design of the experiments, along with their results, is described in Section 4. Section 5 presents the developed web platform, while Section 6 describes its evaluation. Lastly, the proposed work's limitations, possible future directions, and conclusion are provided in Section 7 and Section 8.

2. Previous Works

Due to the nature of blogs, which are articles to express one's opinion on a topic, and their close relativity to columns; moreover, due to the lack of studies on column classification, we exhibit previous works conducted for blogs. In terms of blogs' identification, a work by [36] studied the effectiveness of integrating tags in blog classification using 24,247 blogs. They studied the effectiveness of tags in blog classification and tried to answer three questions: (i) Are tags more effective than other types of data, (ii) is it true that more tags lead to more accurate classification, and (iii) does tag expansion help in getting better classification accuracy. Experimental results showed that tags were more effective than features extracted from blog titles and descriptions, but the best classification accuracy was achieved when all these features were used together. Results also showed that tags could lead to better classification accuracy, but more tags did not necessarily lead to better classification accuracy.

Along the same line, work by [21] proposed a semi-supervised learning method for blog classification, using unlabelled data to improve classification accuracy. They assumed that entries from the same blog have the same characteristics and used these characteristics to improve classification accuracy. The proposed method used a huge number of unlabelled blogs to extract useful features and is an instance of Alternating Structure Optimization (ASO). ASO is a machine learning framework for semi-supervised learning and multi-task learning. It requires using auxiliary problems, such as the prediction of frequent words and the prediction of words in sequence, to train a main classifier. Similarly, the automatic classification of blog entries using a semi-supervised machine learning task was proposed by [12]. They assigned blogs to one of a set of pre-defined classes based on features extracted from their textual content. The study included main phases such as pre-processing, feature extraction, classification model, train classifier, test classifier accuracy, and external glossary of terms. The feature set was enriched by adding synonym words and acronyms from external glossaries. The classifier modeling and training phase was followed by binary feature vectors using the Multivariate Bernoulli model. Nave Bayesian Model and Artificial Neural Network Model were used to classify unstructured blog text data using binary variables and high TF-IDF values. Finally, the probability of a blog entry belonging to each class is predicted, and the blog post is assigned to the class with the highest probability. The tested dataset included 3000 blog posts and comments and results showed that the Naive Bayesian classifier gave better overall classification accuracy than basic neural-network-based classification.

In a recent study, [22] aimed to explore using machine learning models to classify Bangla blog posts into categories. Nine supervised learning classification models were tested: support vector machine, multinomial naive Bayes, multi-layer perceptron, k-nearest neighbours, stochastic gradient descent, decision tree, perceptron, ridge classifier, and random forest. The models were used to classify Bangla blog posts into eight categories. Three feature extraction techniques were applied: unigram TF-IDF, bigram TF-IDF, and trigram TF-IDF. The majority of classifiers achieved over 80% accuracy on the classification task.

Generally, Blogs can also be a rich source of information. It has been used to offer information to projects of sentiment analysis [33, 34], gender classification for commercial applications [28], and sentiment analysis through text and images [11], and author identification [24]. In the domain of Arabic blogs, the authors in the study conducted by[3] put out a framework for evaluating the credibility of these blogs. The definition of credibility was derived from existing literature, where it is understood as the degree of believability. It encompasses several qualities, including trustworthiness, quality, authority, persuasiveness, and popularity. The credibility attributes were categorized into two levels, the first is the blog level, which includes the author's name and the number of comments, and the second is the post level, which encompasses factors such as spelling, use of emoticons, presence of spam, punctuation use, length of the post, presence of positive or negative language, and resemblance to verified material. The researchers presented a conceptual framework for a credibility system, which involved the extraction of characteristics from a dataset of blogs. They proceeded to train and test the system, ultimately constructing classifiers to determine the credibility of each blog. The authors additionally highlighted various challenges encountered in Arabic language processing, such as the identification of proper nouns, the retrieval of writings with diacritics, the inherent inflectional and derivational nature of the Arabic language, and the insufficiency of existing tools for Arabic natural language processing in comparison to those available for English. Along the same effort, the article by [4] discusses the lack of research on credibility analysis of Arabic content, specifically in the context of blogs. To solve this issue they present an Arabic blogs corpus. The researchers manually extracted a limited number of blog features, including bias, sentiment, reasonability, and objectivity. These features were then utilized to train several machine learning models, such as Naive Bayes and Decision Tables. The blog corpus finally consisted of 175 Arabic blog posts collected through Google search queries and annotated for credibility. The corpus was further used for testing in a recent study of Arabic blog credibility presented by [20]. They proposed a deep co-learning approach to assess the credibility of Arabic blog posts via deep neural networks. To overcome the lack of sufficient training data, they used a semi-supervised deep learning method called deep co-learning. Deep co-learning is based on co-training, which utilizes multiple views of data to label additional data without human annotation. This approach helped train robust deep learning models for assessing the credibility of Arabic blogs despite the scarce labeled training data available. They trained their model on 20392 blogs and tested its performance on the dataset presented by Zaatari. Their optimal model, using deep co-learning, reached 63% F-score.

[20] introduced a novel method called deep co-learning, which is a semi-supervised end-to-end deep learning strategy designed to evaluate the trustworthiness of Arabic blogs. The authors have put forth two classifiers that utilize a convolutional neural network (CNN) architecture.

Table 1: Dataset statistics.

Genre	News	Columns
Number of Articles	9500	9500
Average Word Length	4.8	4.3
Average Character Count	3485	2546
Word Count Average	684.45	525.16

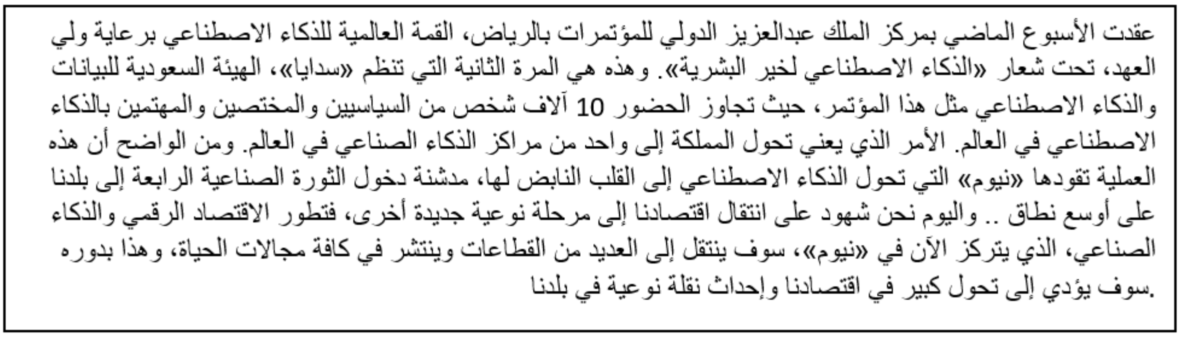

عقدت الأسبوع الماضي بمركز الملك عبدالعزيز الدولي للمؤتمرات بالرياض، القمة العالمية للذكاء الاصطناعي برعاية ولي
العهد، تحت شعار «الذكاء الاصطناعي لخير البشرية». وهذه هي المرة الثانية التي تنظم «سدايا»، الهيئة السعودية للبيانات
والذكاء الاصطناعي مثل هذا المؤتمر، حيث تجاوز الحضور 10 آلاف شخص من السياسيين والمختصين والمهتمين بالذكاء
الاصطناعي في العالم. الأمر الذي يعني تحول المملكة إلى واحد من مراكز الذكاء الصناعي في العالم. ومن الواضح أن هذه
العملية تقودها «نيوم» التي تحول الذكاء الاصطناعي إلى القلب النابض لها، مدشنة دخول الثورة الصناعية الرابعة إلى بلدنا
على أوسع نطاق .. واليوم نحن شهود على انتقال اقتصادنا إلى مرحلة نوعية جديدة أخرى، فتطور الاقتصاد الرقمي والذكاء
الصناعي، الذي يتركز الآن في «نيوم»، سوف ينتقل إلى العديد من القطاعات وينتشر في كافة مجالات الحياة، وهذا بدوره
.سوف يؤدي إلى تحول كبير في اقتصادنا وإحداث نقلة نوعية في بلدنا

Figure 2: Column article.

The initial model uses continuous bag of words (CBOW) word embeddings as its features, whereas the subsequent model utilizes character level embeddings.

Recently, a comprehensive work proposed by [14] proposed an objective to create an entirely automated method for evaluating the credibility of Arabic blog postings. The researchers gathered a dataset consisting of Arabic blog posts. These posts were then annotated, and the important features were extracted and reduced. The researchers then utilized different machine learning models, such as Support Vector Machines, and deep learning models, such as Long Short-Term Memory (LSTM) and CNN, with different input configurations. Based on our analysis, it can be inferred that the LSTM model exhibits superior performance, achieving an accuracy rate of 74%. This outcome is observed when the input consists of whole blog posts, accompanied by a collection of syntactic and morphological data. The collection of extracted attributes comprised reasonably, author expertise, bias, and the perceived overall trustworthiness of the website.

In yet another study by [5], it is argued that extant studies were largely concentrated on utilizing probabilistic data to cluster Arabic content, whereas other studies utilize a sizable corpus of blogs and news in their modeling approach for classification. To address these gaps, there is a need to utilize sentiment analysis (SA) of Arabic content, specifically natural language texts. As a result, big data will benefit from the adoption of a method that automatically creates sentiment lexicons. If the clustering of the news documents is to take into account the existence of natural languages, it is necessary to evaluate the best model for determining semantic word orientation in Arabic websites. It is in these contexts that a recent study by Sherif [32] on Arabic dialectal lexicon annotation for SA offers a comprehensive analysis of present trends and future directions using data annotations and SA for Arabic dialects that were published between 2015 and 2023. For a better classification of data annotation, the study suggests three methods: manual, automatic, and hybrid. These modeling approaches can prove to be vital, especially in a setting where context-specific information on Arabic SA is used to classify columns and news articles or content.

Though the previous works presented added to the literature focused on blogs, the textual content found in blogs may not follow the formal genre found in column articles. The informal language found in blogs can be easily exploited as an indicator that it is different from news articles. However, column articles may follow the formal- journalistic genre found in news articles, causing a misinterpretation of news articles. This study will therefore concentrate on the nuances between news and column articles which may be employed to identify an article as news or column.

3. Methodology

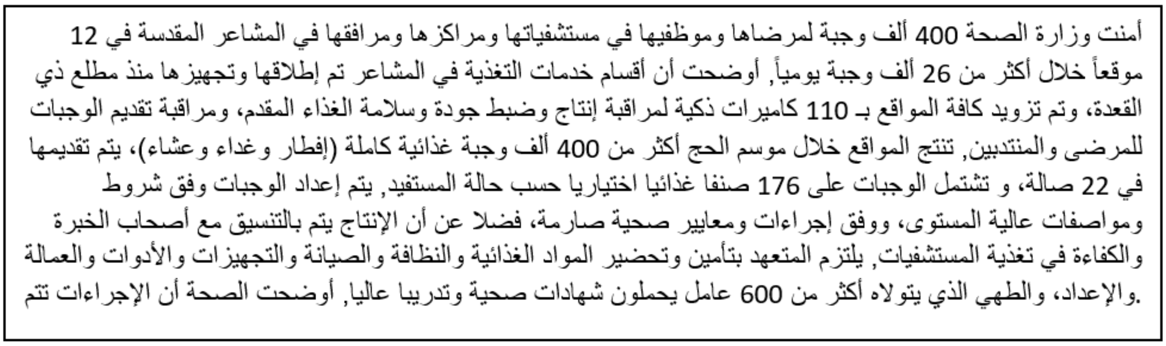
The methodology proposed is demonstrated in three folds. First, we detail the compilation of the dataset. Second, details of the preprocessing approaches are explained. Lastly, a detailed description of the models' development is addressed.

3.1. Data collection

For the data collection, we used Python Application Programming Interface (API) to extract Arabic columns in several Arabic news platforms; Table 1 shows the details. We collected data in two phases. First, we collected 12k columns, 3k from two Saudi and two Egyptian news platforms equally, which are: okaz[1] sabq[2],youm7[3], ahram[4].

In the second phase, we combined the news articles included in the datasets from works of [2] and [39]. The total number of news articles collected was 15 k news articles. To balance the news articles in terms of word length with the column articles, we filtered out any news or column articles with more than 4000 words. We used such an approach to avoid any biases that may appear when using the word embeddings. In total, our dataset included 19k articles divided into 9500 columns and 9500 news articles from various news platforms, where Table 1 shows the dataset's statistics. Figure 2 displays a column article and Figure 3 a news article.

[1] www.okaz.com.sa
[2] www.sabq.org
[3] www.youm7.com
[4] www.ahram.org.eg



أمنت وزارة الصحة 400 ألف وجبة لمرضاها وموظفيها في مستشفياتها ومراكزها ومرافقها في المشاعر المقدسة في 12
موقعاً خلال أكثر من 26 ألف وجبة يومياً, أوضحت أن أقسام خدمات التغذية في المشاعر تم إطلاقها وتجهيزها منذ مطلع ذي
القعدة، وتم تزويد كافة المواقع بـ 110 كاميرات ذكية لمراقبة إنتاج وضبط جودة وسلامة الغذاء المقدم، ومراقبة تقديم الوجبات
للمرضى والمنتدبين, تنتج المواقع خلال موسم الحج أكثر من 400 ألف وجبة غذائية كاملة (إفطار وغداء وعشاء)، يتم تقديمها
في 22 صالة، و تشتمل الوجبات على 176 صنفا غذائيا اختياريا حسب حالة المستفيد, يتم إعداد الوجبات وفق شروط
ومواصفات عالية المستوى، ووفق إجراءات ومعايير صحية صارمة، فضلا عن أن الإنتاج يتم بالتنسيق مع أصحاب الخبرة
والكفاءة في تغذية المستشفيات, يلتزم المتعهد بتأمين وتحضير المواد الغذائية والنظافة والصيانة والتجهيزات والأدوات والعمالة
.والإعداد، والطهي الذي يتولاه أكثر من 600 عامل يحملون شهادات صحية وتدريبا عاليا, أوضحت الصحة أن الإجراءات تتم

Figure 3: News article.

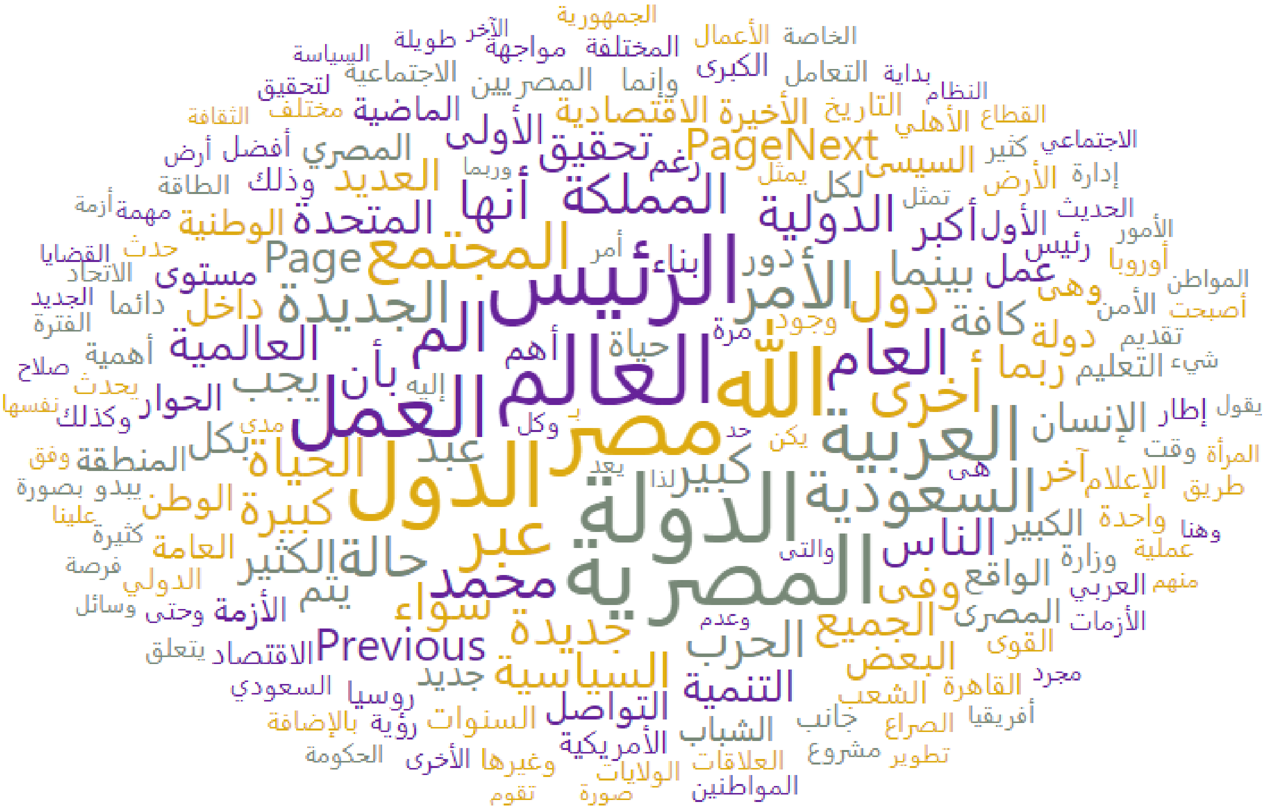


Figure 4: Word Cloud.

3.2. Columns word Analysis

A word cloud is a visual representation of text data. In the context of text analysis, a word cloud is used to depict the frequency of words within a particular text or set of texts. The words are displayed in varying sizes and sometimes colors, where the size (and possibly color) of each word indicates its frequency or importance within the dataset. Here, we applied the word cloud approach to the columns articles. According to Figure 4, we find some interesting insights. First, a dominant number of words related to "Egypt" and "Saudi Arabia" are found, such as Egyptian and Saudi, this can be clearly explained as the column articles are collected from Egyptian and Saudi news platforms. Second, a high number of subjective words are found. Subjectivity, in general, in the context of language and analysis, refers to words that express personal feelings, beliefs, opinions, judgments, assumptions, or interpretations rather than presenting objective, observable facts. Subjective terms in the context include "يجب should," "يبدو seems," "لا يجب should not," and "في الواقع reality," all of which are prevalent. Third, since the columns express the author's opinions about a topic, adjectives in the forms of superiority and comparison are overused in the context. We find words such as "من أفضل better", "أفضل best", "أسوأ worse", and "أكبر larger". Lastly, these findings provide insights into the textual content of column articles.

3.3. Data Pre-processing

The Arabic alphabet consists of 28 letters presenting constants. Moreover, there are three vowels that can be added to any place of a word, forming different meanings. Derived from the vowels, come three diacritical signs that are written below or above the letter, forming different

Table 2: ML Classifiers Parameters.

Model	Parameters
SVM	batchSize 100 kernel linear
NB	batchSize 100
LR	batchSize 100, maxBoosting-Iterations 500
RF	batchSize 100, bagging with num-Iteraions 100, and number of trees 100

orthographical variations to the word. The Arabic words are written from right to left, linked together from the first letter to the last, except for six letters that are joined from the right side only. Arabic is a language with ligatures, which means a letter changes its form based on its location in a word. More than half the letters are written differently depending on their position in the word: the beginning, middle, or last. In addition, some letters have dots on them. Unfortunately, not all the written Arabic text complies with writing these dots in their right positions to the words. For example; the letter *(Y/ي /Ya)* has two small dots written under it, but some omit the two dots under the letter. While many understand the word even without the dots because of the language practice, it is a challenge for Natural Language Processing (NLP) developers to detect letters not written correctly, such as in this case.
To pre-process the Arabic text data, we first defined a set of Arabic stop words and a regular expression pattern to remove Arabic punctuation marks. We then created a Python script, that took the dataset as input and performed the following tasks:

- Remove non-Arabic characters
- Tokenization
- Normalize words to overcome glitches or misplaced dots in a word was performed by replacing letters as follows:
    - آ ، إ ، أ → ا
    - ي → ى
    - ؤ → و
    - ة → ه

3.4. Machine learning Approach

3.4.1. Feature Extraction

In this work, TF-IDF was used for feature extraction in compiling the model using ML approaches, due to its efficiency compared to alternative methods such as bag of words, Term Frequency (TF), and Inverse Document Frequency(IDF). It is mainly used in natural language processing to measure the importance of terms in documents. It combines both TF and IDF as it calculates term frequency and inverse document frequency to assign weights to each term, which improves text classification and information retrieval accuracy.
The formula for TF-IDF is:

$$\text{TF-IDF} = (\text{Term Frequency}) \times (\text{Inverse Document Frequency}) \tag{1}$$

TF-IDF measures term importance in a document or corpus using term frequency and inverse document frequency. This helps to adjust for the fact that some words appear more frequently in general. It can also be used to compare documents and determine which are more relevant. Ultimately, it is a powerful tool for gaining insight into a large amount of text data. The resulting scores can be used to represent documents as numerical vectors for text analysis purposes.
Hence, TF-IDF is widely used and effective for various natural language processing tasks. It can be used for a wide variety of tasks, such as document classification, keyword extraction, clustering, and topic modelling. Researchers have explored modifications and extensions to improve its effectiveness in specific applications [18]. TFIDF has been used in combination with other techniques such as deep learning and word embeddings for higher accuracy in text analysis tasks.

3.4.2. Training

In recent years, machine learning algorithms have been widely used to train models. The complex algorithms are capable of learning from large datasets, making them a popular choice among data scientists. By applying these algorithms to training models, it is possible to quickly identify patterns, trends and correlations in data that may have been missed with more traditional methods. Furthermore, the algorithms can be used to detect anomalies in data and optimize performance. With the help of machine learning algorithms, powerful models can be built to tackle an array of data-related problems. This makes machine learning algorithms the ideal tool for model training and data analysis.The following ML classifiers are used in this study, and Table 2 provides their parameters for the models' compilation:

- **Naive Bayes(NB):** This simple probabilistic classifier works to assume a conditional relation between features of the given data [37]. It is a collection of algorithms based on the Bayes Theorem, assuming that all attributes are strictly independent.
- **Random Forest (RF):** A meta-estimator classification algorithm that builds a 'forest' using decision tree model learning based on bagging techniques [17].
- **Support Vector Machine (SVM):** Classifies input data based on dimensional surfaces by finding the maximum separating hyperplane between different classes [37].
- **Logistic Regression (LR):** A classifier that finds a relation between features and the probability of a certain outcome [29].

Table 3: News and columns confusion matrix.

Type	**Prediction**	
	News	**Columns**
News	True Positive	False Negative
Column	False Positive	True Negative

Table 4: Classification results using ML models.

Model	**Precision**	**Recall**	**F-Score**	**Accuracy**
SVM	96.9	81.0	84.1	85.8
RF	96.0	90.0	92.0	90.0
NB	97.5	87.7	87.6	88.9
LR	97.0	91.4	91.4	91.6

Evaluation Matrix

The performance of each algorithm is based on precision (P), recall (R), and F-Measure (F). All three measures are computed from the confusion matrix created from performing the classification task. The confusion matrix for news and columns classification is shown in Table 3.

Its four probabilities are listed below:

1. True Positive (TP): represents the number of news articles correctly classified as news articles
2. False Negative (FN): represents the number of news articles incorrectly classified as column articles.
3. True Negative (FP): represents the number of columns articles correctly classified as columns articles.
4. False Positive (FP): represents the number of columns articles incorrectly classified as news articles

P, R and F measures are indicated, respectively, in equations (2), (3), (4)

$$\text{Precision (P)} = \frac{\text{TP}}{\text{TP}+\text{TN}} \tag{2}$$

$$\text{Recall (R)} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{3}$$

$$\text{F-Score (F)} = \frac{2 \cdot P \cdot R}{P+R} \tag{4}$$

3.4.3. Testing and Evaluation

The dataset was split into 80% training and 20% testing, resulting in 15200 articles for training and 3800 articles for testing. This ensured that our model was trained on a significant amount of data and also evaluated on an unseen set of data to measure its generalization performance. The training set was used to train the machine learning and deep learning models, while the testing set was used to evaluate the performance of the models.

As shown in Table 4 regarding the SVM, it also achieved impressive performance on a classification task. The evaluation metrics show the accuracy of 85.8%, indicating a high degree of separation between the columns and news, and meaning it correctly classified approximately 85.8% of the total articles in the dataset. The F-score, which balances precision and recall, is 84.1%. This score considers both false positives and false negatives and is particularly useful when dealing with balanced datasets. The good F-score suggests a good balance between precision and recall. Additionally, precision is 96.9%, indicating that when it predicts the positive class, it is correct approximately 96.9% of the time. The recall is 81%, meaning that it captures about 81% of all actual positive articles in the dataset. On the other hand, the RF classifier exhibits strong performance across multiple evaluation metrics. It achieves an impressive accuracy of 90%, which means it correctly classifies 90% of the articles. Moreover, the model demonstrates balanced performance in terms of precision, recall, and the F-score, all of which are in the range between 90% and 96%.This indicates that the model has a high level of precision in correctly identifying positive articles, minimizes false positives, and has a strong ability to find true positive articles, resulting in a balanced trade-off between precision and recall.

Concurrently, the NB exhibits strong performance across several key evaluation metrics. The accuracy attained is 88.9%, which shows that it correctly classifies roughly 88.9% of the articles. This is a respectable accuracy score, especially when the classes are balanced. With an F-score of 92%, the model strikes a good balance between precision and recall. The F-score considers both false positives and false negatives, making it a valuable metric when dealing with balanced datasets. The precision score of 97.5% reflects its ability to make accurate positive predictions. Finally, the LR achieved impressive performance metrics in its evaluation. The high accuracy score of 91.6% indicates it is optimal between ML models. In addition, the F-score, recall, and precision prediction results of 91.4%, 91.4%, and 97.7%, respectively, suggest that the model can effectively identify a large percentage of the correct articles in the dataset.

The ML algorithms show good accuracy in classifying news and columns. However, our next objective is to develop and evaluate the classification models' performance using DL algorithms.

3.5. Deep Learning (DL) Approach

3.5.1. Feature Extraction

Embedding Using BERT
Training models based on embeddings is an essential step in natural language processing tasks as it converts raw text data into a numerical format that can be processed by machine learning models. We use AraBERT [7], which has a vocabulary capacity of 64,000 words, 12 attention heads, 12 hidden layers, 768 hidden sizes, a total of 110 M parameters, and 512 maximum sequence lengths. It was trained using a dataset of 3B Arabic words. Specifically, we used the available version 'paraphrase-multilingual-mpnet-base-v2 model [5]', which is a pre-trained BERT model that is designed to encode multilingual texts into high-quality embeddings.

3.5.2. Compiling models using DL approaches

Details of each DL models' compilation specification re detailed below.

- **Convolutional Neural Network (CNN) :** is an one-dimensional convolutional layer.
  - The Conv1D layer applies 64 filters with a kernel size of 3 to the input data. The activation function used is ReLU.
  - MaxPooling1D layer performs max pooling with a pool size of 2.
  - The flatten layer flattens the output of the previous layer.
  - The dense layer has 2 neurons with a softmax activation function that outputs a probability distribution over the output classes.
  - The model is compiled with categorical cross-entropy loss and the Adam optimizer.
  - The model is trained with the training data and labels for 50 epochs.

- **CNN & Long Short Term Memory(LSTM) :** is a model architecture that combines both LSTM and CNN layers.
  - Bidirectional LSTM layer with 64 units and return_sequences=True, which means the output of each time step will be fed into the next layer.
  - 1D Convolutional layer with 64 filters and kernel size 3, which can extract local patterns from the sequence.
  - MaxPooling1D layer with pool size 2, which reduces the dimensionality of the feature maps.
  - Flatten layer to convert the 3D tensor output from the previous layer to a 1D tensor.
  - Dense layer with 2 units and softmax activation function, which outputs the probability distribution over the classes.

- **Bi-LSTM Architecture**
  - Single Bidirectional LSTM layer with 64 units, followed by a dense layer with a sigmoid activation function that outputs two values for binary classification.
  - Loss function used was categorical_cross entropy
  - optimizer used was Adam
  - The model was trained for 50 epochs.
  - The input shape for the model was (768, 1) which is the size of the BERT embeddings after reshaping.

- **Auto Encoders**
  - The encoding dimension is set to 32.
  - The autoencoder architecture includes an input layer, an encoding layer with ReLu activation function, and a decoding layer with sigmoid activation function.
  - The autoencoder is compiled using Adam optimizer and mean squared error loss function.
  - The model is trained on the training set for 50 epochs with a batch size of 32 and validated using the validation set.
  - The compressed representation of the training and validation sets are extracted using the encoder model.

3.5.3. Testing and Evaluation

DL models have been shown to perform much better than ML models in classifying news and column articles, as given in the Table 5. For example, the LSTM model has demonstrated high performance in classification tasks, with a high accuracy of 92%. The evaluation metrics show a precision of 92%, indicating that the model is excellent at correctly identifying positive articles while minimizing false positives. The recall score of 92.9% implies that the model is highly effective at identifying most of the actual positive articles. The F-Score of 91.9% indicates that the model achieved a strong balance between precision and recall. Along the same line, regarding CNN, the model achieved outstanding performance with precision, F-score, and accuracy. The accuracy reached 92.3%, with minimal difference between its accuracy compared to LSTM's accuracy. Interestingly, when combining CNN with LSTM, the accuracy boosted to a high score of 4%, reaching 96.6%. The precision and recall also gave high results, which made the model an optimal classification model. Finally, the autoencoders model demonstrated commendable performance with accuracy scores of 87%. Precision and recall scores of 89.6% and 84.4%, respectively, imply that when the model makes a positive prediction. The F-score, a harmonic mean of precision and recall, is 88.5%, indicating a balanced performance in terms of both precision and recall and suggesting that the model can effectively manage the trade-off between them. Finally, the Autoencoders model performed well across the evaluation metrics, however, a combined model of CNN-LSTM made it an optimal choice for the classification task.

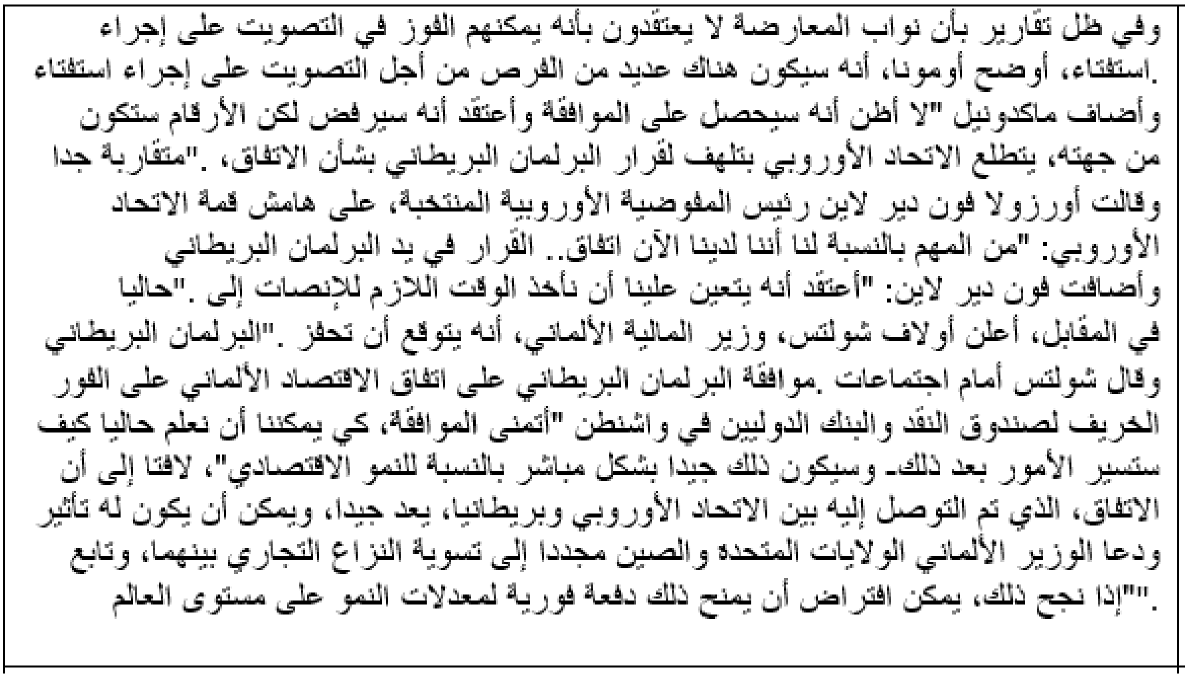
[5] https://huggingface.co/

Table 5: Results using DL models.

Model	Precision	Recall	F-Score	Accuracy
LSTM	92.0	92.9	91.9	92.0
CNN	95.3	91.4	93.4	92.3
CNN-LSTM	96.6	95.2	96.2	96.6
AutoEncoders	89.6	84.4	88.5	87.6

Table 6: News and columns confusion matrix.

Genre	News	Columns
News	8542	958
Columns	348	9152



وفي ظل تقارير بأن نواب المعارضة لا يعتقدون بأنه يمكنهم الفوز في التصويت على إجراء
.استفتاء، أوضح أومونا، أنه سيكون هناك عديد من الفرص من أجل التصويت على إجراء استفتاء
وأضاف ماكدونيل "لا أظن أنه سيحصل على الموافقة وأعتقد أنه سيرفض لكن الأرقام ستكون
من جهته، يتطلع الاتحاد الأوروبي بتلهف لقرار البرلمان البريطاني بشأن الاتفاق، ."متقاربة جدا
وقالت أورزولا فون دير لاين رئيس المفوضية الأوروبية المنتخبة، على هامش قمة الاتحاد
الأوروبي: "من المهم بالنسبة لنا أننا لدينا الآن اتفاق.. القرار في يد البرلمان البريطاني
وأضافت فون دير لاين: "أعتقد أنه يتعين علينا أن نأخذ الوقت اللازم للإنصات إلى ."حاليا
في المقابل، أعلن أولاف شولتس، وزير المالية الألماني، أنه يتوقع أن تحفز ."البرلمان البريطاني
وقال شولتس أمام اجتماعات .موافقة البرلمان البريطاني على اتفاق الاقتصاد الألماني على الفور
الخريف لصندوق النقد والبنك الدوليين في واشنطن "أتمنى الموافقة، كي يمكننا أن نعلم حاليا كيف
ستسير الأمور بعد ذلكـ وسيكون ذلك جيدا بشكل مباشر بالنسبة للنمو الاقتصادي"، لافتا إلى أن
الاتفاق، الذي تم التوصل إليه بين الاتحاد الأوروبي وبريطانيا، يعد جيدا، ويمكن أن يكون له تأثير
ودعا الوزير الألماني الولايات المتحده والصين مجددا إلى تسوية النزاع التجاري بينهما، وتابع
.""إذا نجح ذلك، يمكن افتراض أن يمنح ذلك دفعة فورية لمعدلات النمو على مستوى العالم

Figure 5: News article misclassified as column.

Based on the presented results, CNN-LSTM performed best in classifying news and columns. Hence, it is an optimal tool while dealing with the huge dataset of Arabic text. The following section will discuss misclassified articles using the optimal model, CNN-LSTM trained by BERT.

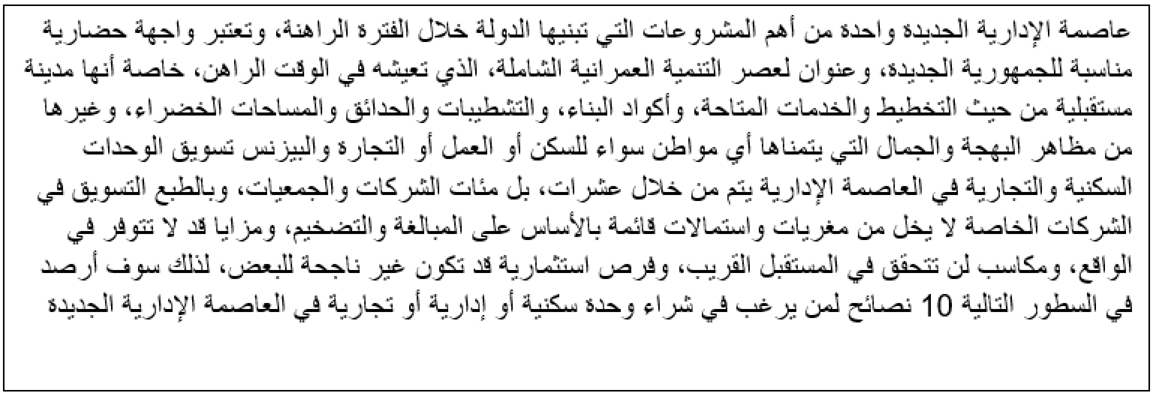
4. Error Analysis

To further enhance the credibility of our work, we examine the misclassified articles in both news and column classes predicted from the CNN-LSTM optimal model. According to Table 6, we found that 90% of news articles were correctly classified. Moreover, around 96% of columns were correctly classified. We investigated the 10% and 4% misclassified news and columns, respectively, and further summarized several of the optimal models' errors. First, most mispredicted news articles as columns included some subjective indicators. Figure 5 shows a sample of news articles misclassified as columns. The article described the Brexit negotiations in the British Parliament. However, the interesting part of the article is that it contains many quotes from key figures in the article. The quotes contained their personal perspective of the topic, thus including subjective indicators such as “أتمنى I wish.” and “لا أعتقد I don't think.” These subjective indicators may have resulted in ambiguity, hence classifying the news article as a column. In contrast, columns misclassified as news articles featured very similar news registers with no subjective indicators. A sample of columns wrongfully classified as news articles is presented in Figure 6. The column announces an innovative housing project and describes some of its facilities and benefits to the community. It focuses on using general descriptions with no personal insights, thus offering no subjective indicators.

.

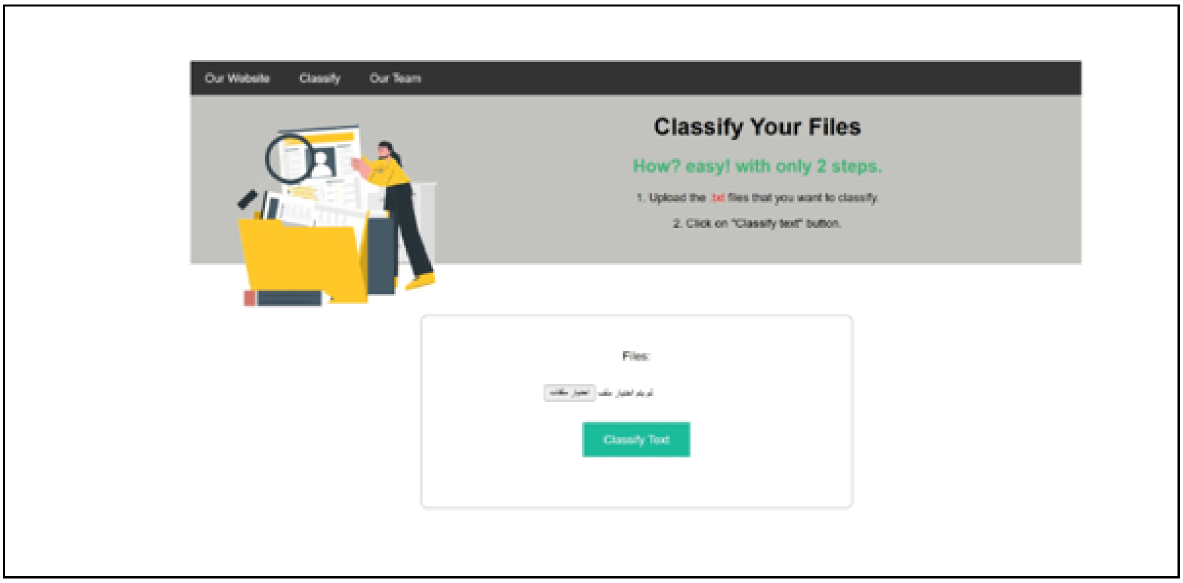
5. Model Platform

We developed a classification model in the form of a web platform by deploying the best perfoming model, CNN-LSTM, using Django software. Figure 7 shows the platform's Graphical User Interface (GUI) which includes a brief description of the platforms' usage and an input box to input multiple articles in the form of text files. Figure 8 shows the output of tested files entered. The output displays the file's name and its classification result. The platform can be obtained upon request.



عاصمة الإدارية الجديدة واحدة من أهم المشروعات التي تبنيها الدولة خلال الفترة الراهنة، وتعتبر واجهة حضارية
مناسبة للجمهورية الجديدة، وعنوان لعصر التنمية العمرانية الشاملة، الذي تعيشه في الوقت الراهن، خاصة أنها مدينة
مستقبلية من حيث التخطيط والخدمات المتاحة، وأكواد البناء، والتشطيبات والحدائق والمساحات الخضراء، وغيرها
من مظاهر البهجة والجمال التي يتمناها أي مواطن سواء للسكن أو العمل أو التجارة والبيزنس تسويق الوحدات
السكنية والتجارية في العاصمة الإدارية يتم من خلال عشرات، بل مئات الشركات والجمعيات، وبالطبع التسويق في
الشركات الخاصة لا يخل من مغريات واستمالات قائمة بالأساس على المبالغة والتضخيم، ومزايا قد لا تتوفر في
الواقع، ومكاسب لن تتحقق في المستقبل القريب، وفرص استثمارية قد تكون غير ناجحة للبعض، لذلك سوف أرصد
في السطور التالية 10 نصائح لمن يرغب في شراء وحده سكنية أو إدارية أو تجارية في العاصمة الإدارية الجديده

Figure 6: Column misclassified as news article.



Classify Your Files

How? easy! with only 2 steps.

1. Upload the .txt files that you want to classify.

2. Click on "Classify text" button.

Figure 7: GUI of platform.

6. Tool Evaluation

In an effort to test the reliability and usability of our tool, we employ 30 participants from our local university, the University of Jeddah. They were given the URL for the deployed application and trained on how to use it in a Zoom meeting. After that, they were asked to answer ten questions on the System Usability Scale (SUS) [10]. An effective “quick and dirty” approach for gauging usability is the SUS. It consists of a 10-item questionnaire with five response alternatives, ranging from Strongly agree to Strongly disagree. The calculation for SUS is as follows:

$$X = \text{Sum of the points for all odd-numbered questions} - 5 \tag{5}$$

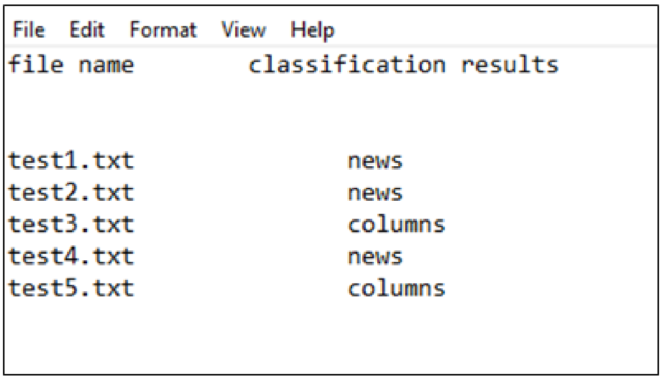
$$Y = 25 - \text{Sum of the points for all even-numbered questions} \tag{6}$$

$$\text{SUS Score} = (X+Y) \times 2.5 \tag{7}$$

Following the equations above for the 30 participants the usability of each participant is depicted in Figure 9. The average usability was 75.15% which is “good” according to the SUS scales.

7. Limitations and Future Work

According to the foregoing error analysis, the present study highlights the challenging task of classifying columns written in a journalistic genre with no subjective sense. Furthermore, news articles that include subjective indicators may be difficult to differentiate from columns. For that, our future objective is to employ more analytical works to identify column articles. Our future target is to compile artificial intelligent models that capture more specific nuances between news and column articles in terms of their textual content for a more accurate prediction.



```
file name        classification results


test1.txt              news
test2.txt              news
test3.txt              columns
test4.txt              news
test5.txt              columns
```

Figure 8: Output.

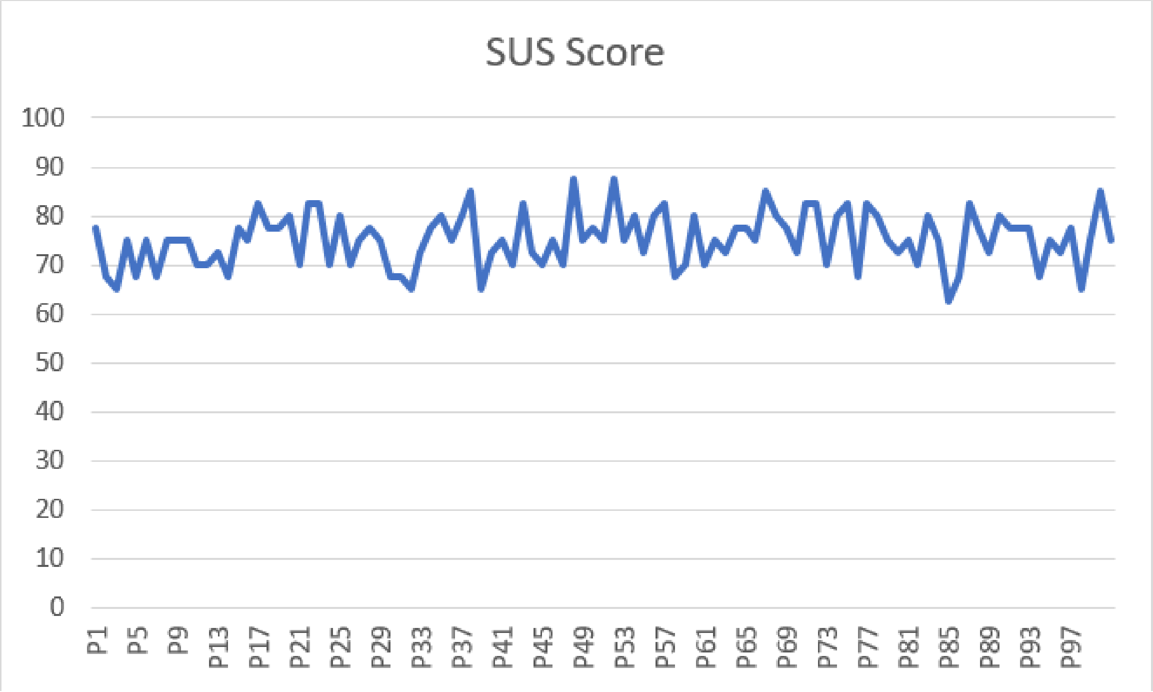


Figure 9: Tool usability scores.

8. Conclusion

In this paper, in an effort to increase credibility measures across news platforms, we present an automated classification model for differentiating between Arabic news and column articles. We compiled a new dataset of posted columns across real-time news platforms. To the best of our knowledge, this is the first and largest Arabic dataset for column article classification. Furthermore, we applied ML and DL approaches and built a classification model that performed with high accuracy. Our findings show that the ML models perform very well with the highest classifier, Random Forest, reaching 90% accuracy. However, DL models achieve higher accuracies when trained by BERT word embedding. Our results show that CNN-LSTM gives the best accuracy score of 96.6% compared to other DL algorithms. We further applied a thorough word analysis for the column articles. By conducting word analysis in column articles, readers and researchers can gain insights into the writer's linguistic choices and how they contribute to the overall effectiveness of the column. It provides a deeper understanding of the author's style, rhetorical strategies, and intended message. These insights can be further used for future works in compiling models for text classification or sentiment analysis. Lastly, we make use of our DL optimal model and develop an automated classification web platform that automatically classifies articles as news or columns. In the future, we plan to enhance our data analysis method by considering linguistic features in the articles' textual content to obtain better results.

Acknowledgement

References

[1] Aggarwal, C. C., and Zhai, C. A survey of text classification algorithms. *Mining text data* (2012), 163–222.

[2] Al-Barhamtoshy, H. M., Himdi, H. T., and Alyahya, M. Arabic pilgrim services dataset: Creating and analysis. In *2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC)* (2023), pp. 1–8.

[3] AL-EIDAN, R. M. B., AL-KHALIFA, H. S., AND AL-SALMAN, A. S. Towards the measurement of arabic weblogs credibility automatically. In *Proceedings of the 11th international conference on information integration and web-based applications & services* (2009), pp. 618–622.

[4] AL ZAATARI, A., EL BALLOULI, R., ELBASSOUNI, S., EL-HAJJ, W., HAJJ, H., SHABAN, K., HABASH, N., AND YAHYA, E. Arabic corpora for credibility analysis. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (2016), pp. 4396–4401.

[5] ALHARBI, A. R., HIJJI, M., AND ALJAEDI, A. Enhancing topic clustering for arabic security news based on k-means and topic modelling. *IET Networks 10*, 6 (2021), 278–294.

[6] ANTOUN, W., BALY, F., AND HAJJ, H. Arabert: Transformer-based model for arabic language understanding. In *LREC 2020 Workshop Language Resources and Evaluation Conference 11–16 May 2020* (2020), p. 9.

[7] ANTOUN, W., BALY, F., AND HAJJ, H. AraBERT: Transformer-based Model for Arabic Language Understanding. *arXiv e-prints* (2020), arXiv:2003.00104.

[8] ARUNA DEVI, K. Two dimensional feature extraction and blog classification using artificial neural network. *International Journal of Applied Engineering Research 13*, 9 (2018), 6536–6544.

[9] BLOCK, A. Why newspapers should not have columnists. https://stanforddaily.com/2014/11/09/why-newspapers-should-not-have-columnists/, November 9 2014. Accessed on December 22, 2023.

[10] BROOKE, J. Sus: a "quick and dirty'usability. *Usability evaluation in industry 189*, 3 (1996), 189–194.

[11] DAL MOLIN, G. P., SANTOS, H. D., MANSSOUR, I. H., VIEIRA, R., AND MUSSE, S. R. Cross-media sentiment analysis in brazilian blogs. In *Advances in Visual Computing: 14th International Symposium on Visual Computing, ISVC 2019, Lake Tahoe, NV, USA, October 7–9, 2019, Proceedings, Part II 14* (2019), Springer, pp. 492–503.

[12] DALAL, M. K., AND ZAVERI, M. A. Automatic classification of unstructured blog text. *Journal of Intelligent Learning Systems and Applications 5*, 02 (2013), 108–114.

[13] DU, Y., YI, Y., LI, X., CHEN, X., FAN, Y., AND SU, F. Extracting and tracking hot topics of micro-blogs based on improved latent dirichlet allocation. *Engineering Applications of Artificial Intelligence 87* (2020), 103279.

[14] EL-HAJJ, W., BRAHIM, G. B., AND ZAATARI, A. Assessing in real-time the credibility of arabic blog posts using traditional and deep learning models. *Social Network Analysis and Mining 11*, 1 (2021), 72.

[15] FULLER, J. *News values: Ideas for an information age*, vol. 10. University of Chicago Press, 1996.

[16] GANS, H. J. Deciding what's news: A study of cbs evening news, nbc nightly news. *Newsweek, and Time. New York: Pantheon 42* (1979), 48.

[17] GENUER, R., POGGI, J.-M., TULEAU-MALOT, C., AND VILLA-VIALANEIX, N. Random forests for big data. *Big Data Research 9* (2017), 28–46.

[18] GUO, B., ZHANG, C., LIU, J., AND MA, X. Improving text classification with weighted word embeddings via a multi-channel textcnn model. *Neurocomputing 363* (2019), 366–374.

[19] HASSANI, H., BENEKI, C., UNGER, S., MAZINANI, M. T., AND YEGANEGI, M. R. Text mining in big data analytics. *Big Data and Cognitive Computing 4*, 1 (2020), 1.

[20] HELWE, C., ELBASSUONI, S., AL ZAATARI, A., AND EL-HAJJ, W. Assessing arabic weblog credibility via deep co-learning. In *Proceedings of the Fourth Arabic Natural Language Processing Workshop* (2019), pp. 130–136.

[21] IKEDA, D., TAKAMURA, H., AND OKUMURA, M. Semi-supervised learning for blog classification. In *AAAI* (2008), pp. 1156–1161.

[22] ISLAM, T., PRINCE, A. I., KHAN, M. M. Z., JABIULLAH, M. I., AND HABIB, M. T. An in-depth exploration of bangla blog post classification. *Bulletin of Electrical Engineering and Informatics 10*, 2 (2021), 742–749.

[23] JANNATI, R., MAHENDRA, R., WARDHANA, C. W., AND ADRIANI, M. Stance classification towards political figures on blog writing. In *2018 International Conference on Asian Language Processing (IALP)* (2018), IEEE, pp. 96–101.

[24] KEŠELJ, V. Automated authorship attribution using cng distance on blog posts in the serbian language. In *2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH)* (2023), IEEE, pp. 1–8.

[25] LERNER, K. M. Journalists believe news and opinion are separate, but readers can't tell the difference. https://theconversation.com/journalists-believe-news-and-opinion-are-separate-but-readers-cant\-tell-the-difference-140901, June 22 2020. Accessed on December 22, 2023.

[26] MANJOO, F. I was wrong about facebook. https://www.nytimes.com/2022/07/21/opinion/farhad-manjoo-facebook.html, July 21 2022. Accessed on December 22, 2023.

[27] MICROSOFT. Is it an article, a column, or an editorial (and why does it matter)? https://www.microsoft.com/en-us/microsoft-365-life-hacks/writing/article-column-or-editorial, March 20 2023. Accessed on December 22, 2023.

[28] MUKHERJEE, A., AND LIU, B. Improving gender classification of blog authors. In *Proceedings of the 2010 conference on Empirical Methods in natural Language Processing* (2010), pp. 207–217.

[29] PRAMANIK, M., PRADHAN, R., NANDY, P., QAISAR, S. M., AND BHOI, A. K. Assessment of acoustic features and machine learning for parkinson's detection. *Journal of healthcare engineering 2021* (2021).

[30] PSA RESEARCH CENT. Top 100 u.s. magazines by circulation. http://www1.psaresearch.com/images/TOPMAGAZINES.pdf, November 9 2014. Accessed on December 22, 2023.

[31] SCHERLEN, A. Part i: Columns and blogs: Making sense of merging worlds. *The Serials Librarian 54*, 1-2 (2008), 79–92.

[32] SHERIF, S. M., ALAMOODI, A., ALBAHRI, O., GARFAN, S., ALBAHRI, A., DEVECI, M., BAKER, M. R., AND KOU, G. Lexicon annotation in sentiment analysis for dialectal arabic: Systematic review of current trends and future directions. *Information Processing & Management 60*, 5 (2023), 103449.

[33] SHIRSAT, V. S., JAGDALE, R. S., AND DESHMUKH, S. N. Sentence level sentiment identification and calculation from news articles using machine learning techniques. In *Computing, Communication and Signal Processing: Proceedings of ICCASP 2018* (2019), Springer, pp. 371–376.

[34] SINGH, V. K., PIRYANI, R., UDDIN, A., AND WAILA, P. Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification. In *2013 International mutli-conference on automation, computing, communication, control and compressed sensing (imac4s)* (2013), IEEE, pp. 712–717.

[35] SULLIVAN, M. An uneasy mix of news and opinion. https://www.nytimes.com/2015/01/11/public-editor/an-uneasy-mix-of-news-and-opinion.html, January 10 2015. Accessed on December 22, 2023.

[36] SUN, A., SURYANTO, M. A., AND LIU, Y. Blog classification using tags: An empirical study. In *Asian Digital Libraries. Looking Back 10 Years and Forging New Frontiers: 10th International Conference on Asian Digital Libraries, ICADL 2007, Hanoi, Vietnam, December 10-13, 2007. Proceedings 10* (2007), Springer, pp. 307–316.

[37] VIJAYAN, V. K., BINDU, K., AND PARAMESWARAN, L. A comprehensive study of text classification algorithms. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (2017), IEEE, pp. 1109–1113.

[38] WEBSTER, N. Merriam-webster online dictionary. https://www.merriam-webster.com/, 1828. Accessed on December 22, 2023.

[39] ZAMZAMI, N., HIMDI, H., AND SABBEH, S. F. Arabic news classification based on the country of origin using machine learning and deep learning techniques. *Applied Sciences 13*, 12 (2023), 7074.