Archives available at journals.mriindia.com

International Journal on Advanced Electrical and Computer Engineering

ISSN: 2349-9338

Volume 15 Issue 01s, 2026

Multi Group classification based on Arabic

[1]Suliman Al-Shossi, [2]Akram Alsubari

[1] Faculty of Science, University of Ibb, Yemen

[2] Faculty of Science, University of Ibb, Yemen

Email: [1] suliman.alshossi@ibbuniv.edu.ye, [2]akram.alsubari@ibbuniv.com

Peer Review Information

Submission: 05 Dec 2025

Revision: 25 Dec 2025

Acceptance: 10 Jan 2026

Keywords

Hierarchical Classification of Arabic Texts, Arabic Transformer Models, Multi Group label, Fine-Tuning AraT5 and AraGPT2 Models, Generative Text-to-Text.

Abstract

Hierarchical text classification represents a significant and critical challenge in the field of Arabic natural language processing. This challenge is further complicated by the language's morphological richness and the scarcity of large-scale, structured datasets. This paper presents a comprehensive comparative study of two modern approaches to fine-tuning this task: a generative text-to-text approach using the AraT5 model, and a direct classification approach using the AraGPT-2 model. These models were evaluated on a large, specially collected dataset comprising over 75,000 articles distributed across 600 subcategories, as well as a smaller benchmark dataset compared to the previous literature. Experimental results demonstrated that the generative AraT5 model achieved superior performance and hierarchical consistency on this large and complex dataset. Furthermore, our improved AraGPT-2 model, enhanced with advanced regularization techniques, significantly outperformed the current literature benchmark on the compared dataset, achieving a hierarchical F1 score of 93.54%. The results indicate that, while both approaches are effective, the generative approach demonstrates a clear advantage in dealing with a fuzzy classification space. This work establishes new performance benchmarks and provides critical insights into the impact of fine-tuning strategies and data complexity on the hierarchical classification of Arabic.

Introduction

Language is a complex and dynamic communication system that enables individuals to communicate thoughts, emotions, intentions, and information through an organized set of symbols and rules. The science that studies this complex system is known as linguistics, which seeks to understand its various components, including sounds, words, sentence structures, and meanings [1]. In our digital age, the field of Natural Language Processing (NLP) has emerged, a branch of artificial intelligence that enables machines to understand, interpret, and respond to human language in a meaningful way [2]. Text classification is one of the fundamental and important tasks in the fields of natural language processing and machine learning, as it aims to design algorithms that allow computers to extract features and automatically classify texts [3]. Text classification has numerous applications, spanning broad areas such as information retrieval, topic labeling, sentiment analysis, and news classification [4]. Recent years have witnessed a significant shift in text classification techniques, moving from traditional "shallow learning" methods to "deep learning" approaches. Deep learning models have gained significant traction due to their ability to model complex features without the need for manual engineering [4]. proving to be more powerful than traditional machine learning methods at representing features [3]. At the heart of this development, the transformer architecture has emerged as the dominant architecture in natural language processing, outperforming convolutional neural networks (CNNs) and recurrent neural networks (RNNs). Significant advances in transformer architectures and pre-training have led to the construction of models with higher capacity that can be easily adapted to specific tasks such as text classification and achieve strong performance [5]. Among these models, two major architectures have emerged: the first is "decoder-only" models such as GPT-2, which learn tasks without any explicit supervision [6]. The second architecture is encoder-decoder models such as the T5 model, which rely on a unified framework that transforms all text-based language problems into a text-to-text format[7]. Despite significant progress in this field, Arabic language generation applications remain underrepresented despite the large and diverse speaker base [8]. Hierarchical text classification, which involves classifying texts into a multi-level classification, adds layers of complexity, especially in distinguishing closely related categories within the same superclass [8]. Advanced models have been developed, such as AraGPT2, the first advanced Arabic language generation model based on the GPT-2 architecture, and AraT5, which is based on the T5 framework and is capable of converting all language problems into text-to-text (Text), making it inherently suitable for generation tasks [9]. Accordingly, this paper aims to provide a comparative study of the performance of two.

Related Work

A review of the historical development of text classification techniques reveals an upward trajectory, from traditional statistical methods to the advanced transformer models that dominate the field today. Initially, the classification process relied on successive stages including preprocessing, feature extraction, and then the application of a classification algorithm [4]. "Shallow learning" algorithms such as Support Vector Machines (SVMs), K-Nearest Neighbors (KNNEs), and Naïve Bayes (NBBs) [10]. dominated, requiring manual construction of text features such as "bag-of-words" (BOWs) [4]. For example, the Maximum Entropy algorithm was successfully used to classify Arabic news after applying preprocessing techniques such as word tracing and part-of-speech identification [11]. Later, with the advent of deep learning, neural networks demonstrated superior capabilities that enabled them to surpass traditional methods in many tasks [12]. Convolutional neural networks (CNNs) excelled at extracting local features, while recurrent neural networks (RNNs) and long short-term memory (LSTM) networks excelled at modeling long-term dependencies in text sequences [3]. However, the real paradigm shift came with the emergence of the "Transformer" architecture, which brought about tremendous progress in natural language processing. By relying on a self-attention mechanism, this architecture was able to model the relationships and understand the context between all words in a text, regardless of the distance between them [5]. This architecture has paved the way for the emergence of pre-trained large language models (PLMs) that have redefined the capabilities of the field. These models can be divided into two main categories: decoder-only models such as GPT-2, which excel at generation tasks [6]. and encoder-decoder models such as T5, which address diverse tasks, including classification, within a unified text-to-text framework [7]. When applying these techniques to Arabic, unique challenges arise due to the presence of many different dialects [13]. To address this, dedicated transformer models such as AraT5 have been developed, which has demonstrated state-of-the-art performance on a wide range of tasks thanks to its pre-training on massive Arabic datasets [14]. This becomes more complex when dealing with hierarchical classification, which requires the model to understand the relationships between main and subclasses [8]. In this specific context, the work presented by Bouchiha et al. This is closest to what will be presented in this paper. They fine-tuned the AraGPT2 model for a hierarchical classification task on a limited dataset (12 subclasses), achieving an accuracy of 80.64% [8]. They also explored a hybrid architecture that combines a BERT model with a BiLSTM network for the same task. The BERT model extracts rich contextual representations of words, and the BiLSTM layer then models the entire text sequence before sending the output to the final classification layer [15]. Based on this review, and despite the progress made, it is clear that previous studies, particularly the work of Bouchiha et al. [8], [15], leave a fundamental research gap. Although their approach has proven effective on a small scale, the scalability of these models to handle more complex hierarchical classification problems remains questionable. The field also lacks a direct, large-scale comparison of the performance of different architectures (Decoder-Only versus Encoder-Decoder) on this particular

Methodology

This section reviews the detailed methodology used in the construction and adaptation of deep learning models for the task of hierarchical classification of Arabic texts. The methodology begins with the collection of data from multiple sources using web scraping technology, followed by a critical stage of pre-processing that included standardizing hierarchies and reducing the size of data via sample reduction to address the problem of imbalance between categories. Subsequently, an improved data segmentation methodology (derived from stratified sampling) was applied to ensure a fair representation of all categories in the training and test groups. Finally, the different transfer Learning strategies of both models (AraGPT-2 with direct classification, and AraT5 with generative classification) are explained and each structure is adapted to the task of hierarchical classification.

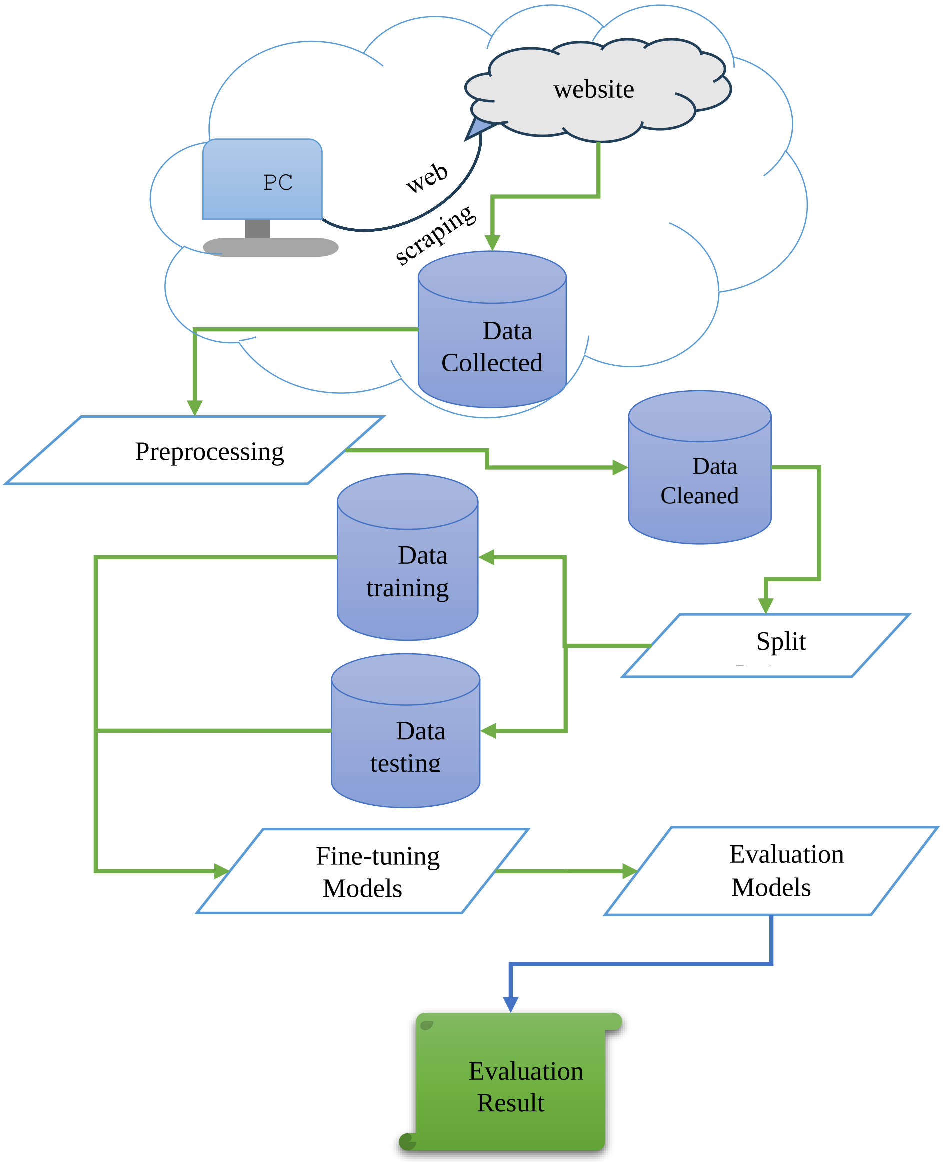

Fig. 1. Steps of the Methodology.

1. Data Collection

Building a robust and representative dataset is the cornerstone of any machine learning study. This section details the methodological steps involved in the data collection process for this study, from identifying and selecting appropriate sources, through the technical process of data extraction, to verifying compliance with ethical considerations.

Data Sources and Characteristics The dataset for this study was constructed by collecting articles from diverse fields to ensure comprehensive coverage. These fields included literature, Islamic topics, general culture, medicine, education, nutrition, culinary arts, stories, travel, commerce, technology, and others. Four leading Arabic websites were selected as data sources: Mqall [16], Mawdoo3 [17], Mhtwyat [18], and Mqalaty [19]. These sources were selected based on their rich content, organized within a clear hierarchical structure consisting of main categories and subcategories, making it easier to identify the areas covered by the articles. The collection process resulted in a final dataset of 94,685 articles, categorized into 72 main categories and 767 subcategories. Table 1 provides a breakdown of the number of categories collected from each site.

Table 1. Statistics for main and subcategories collected from each source.

Websites	Main categories available	Available subcategories	Main categories collected	Subcategories collected
maqall	24	356	22	320
mqalaty	18	99	18	99
mhtwyat	16	107	14	98
mawdoo3	21	328	18	250
Total	76	868	72	767

To visualize the scale and complexity of this hierarchical structure, the relationships between the main and sub-categories were modeled as a network graph, as shown in Figure 2.

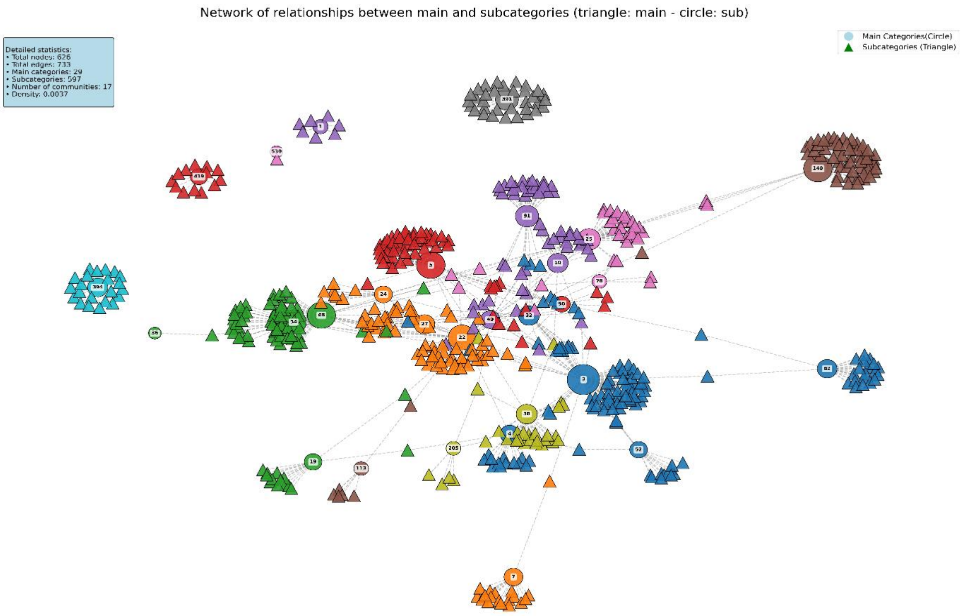

Fig .2. A network visualization of the hierarchical relationships between the main categories (large circles) and sub-categories (triangles).

Node colors represent distinct communities detected within the network, highlighting clusters of thematically related topics. This network visualization reveals key structural characteristics of the dataset that a table alone cannot show. The use of a force-directed layout algorithm groups related topics into distinct "galaxies" or clusters, visually confirming the thematic coherence of the data. The size of each main category node (the large circles) is proportional to the number of sub-categories it contains, immediately highlighting the dominant topics within the corpus. Furthermore, the colors do not represent the main categories themselves, but rather "communities" of nodes that are more densely interconnected. This advanced analysis reveals the inherent semantic overlap between certain high-level topics, which is a key challenge that the machine learning models must overcome. To confirm the accuracy of the number of categories he mentioned, a complete list of all the main and sub-categories and their structure has been included in Appendix A1 of this study.

web scraping. The dataset was created using web scraping, a technique for automatically extracting data from websites using specialized software rather than manually copying and pasting [20] This process required developing custom Python code for each target website, given the differences in their internal programming structures. The extraction process relied on two main libraries:

– *Requests:* A popular HTTP library in Python that facilitates the process of sending HTTP requests to a specified web server and receiving the response content, which is typically the raw HTML code of the page [21]. It was used to send HTTP requests to servers and retrieve the raw, unstructured HTML code of article pages. When retrieved, this code is disorganized and difficult to read.

– *Beautiful Soup:* It was used to parse the retrieved HTML code and convert it into a structured document tree. This tree is an organized hierarchical representation of HTML elements, allowing for accurate text extraction (title, content, and categories) [21]

The extraction process took between two and five hours for each site. After the data was extracted, it was saved in separate CSV files and then combined into a single final data file. Privacy concerns were observed, but all collected data is public content and may be accessed and used in accordance with the ethics of scientific research or the terms of use of the source sites

2. Data Preprocessing

After completing the data collection process, the dataset underwent multiple preprocessing steps to ensure its quality and consistency and prepare it for the training process. This preprocessing was carried out in three main stages: standardization of category labels, addressing category imbalances, and text cleaning and standardization.

Label Normalization. The initial data screening process revealed discrepancies in the labels of the main and subcategories, which was expected given that the data was collected from several different electronic sources. To address this, a manual standardization process was conducted to combine similar names for the main and subcategories. This process resulted in reducing the number of main categories to 29 and the number of subcategories to 600 distinct categories.

Class Imbalance Handling. Analysis of the data distribution revealed significant variation in the number of articles per subcategory, with some categories containing more than 2,000 articles, while others contained fewer than 20. This phenomenon, where some categories (majority categories) are significantly more frequent than others (minority categories), is known as class imbalance, and can bias trained models toward the majority categories. To address this problem, down sampling was used. This is a resampling method that aims to balance the distribution of categories by randomly reducing the number of samples in the majority categories [22]. A Python script was developed that caps the number of articles in each subcategory at 400. Any subcategory containing more than 400 articles was reduced by retaining 400 articles and deleting the rest, while categories containing 400 articles or fewer were left unchanged. As a result of this process, 19,433 articles were removed, bringing the total size of the final dataset to 75,282 articles.

Text Cleaning and Normalization. A comprehensive set of cleaning and normalization processes was applied to the textual content of each article to ensure noise-free and consistent syntax. These processes included:

– Removing common Arabic stopwords.
– Removing all vowel signs (diacritics) and intonation symbols.
– Removing all non-alphabetical symbols, such as punctuation and numbers.
– Normalizing the various forms of the alif letter (آ ,أ ,إ) to the basic form."ا"
– Converting the ta marbuta (ة) to a ha (ه) at the end of words.
– Removing repetitions of letters, newlines, and any text written in languages other than Arabic.

3. Data Splitting Methodology

Given the large number of main classes and subclasses in the dataset, controlling the class balance poses significant challenges, potentially biasing models toward the majority classes in the classification process. To address this challenge, a custom splitting method was designed and implemented to control the distribution of classes between the training and test sets, especially with a data size of 75,752 records and a massive number of subcategory (600). This methodology is a customized and improved version of stratified sampling, and can be referred to as [23]. It is implemented by independently checking the frequency of records in each subclass. If the frequency is greater than 30 records (θ), a fixed number of 26 records (N_large) are selected for testing, while the remaining records in that subclass are allocated to the training set. If the number of records in a subclass is 30 or fewer, 20% (R_small) of these records is taken for testing, and the rest is allocated for training. This method aims to construct a more representative and balanced test set across all classes, rather than taking perfectly equal samples, thus ensuring a fair assessment of model performance. As a result, the test data became 15,184 samples, representing 20% of the total dataset, while the training data became 60,067 samples, representing 80% of the total dataset. This data splitting method is a fundamental step that helped achieve better balance in the test classes and enabled models to train on the data more effectively while reducing the risk of bias in classification processes. To provide a precise description of this method, it is formalized in the following algorithm:

1	**Input:** Dataset D, threshold $\emptyset$, test count for large classes N_{large}, test ratio for small classes R_{small}, random seed S_{seed}
2	**Output:** Training set D_{train}, Test set D_{test}
3	**Initialize** D_{train} as $\emptyset$
	Initialize D_{test} as $\emptyset$
4	Shuffle (D, S_{seed})
	Let C be the set of unique classes in D
5	**For** each class C in C
	Let D_C be the subset of D for class C
	Let n_c = size of D_C
6	**If** n_c> θ $theta$
	K= N_{large}
7	**Else**
	K= [n_c × R_{small}]
8	**End If**
	Append first K items of D_C to D_{test}
	Append remaining $(n_c - K)$ items of D_C to D_{train}
9	Next C
	Shuffle (D_{train}, S_{seed})
	Shuffle (D_{test}, S_{seed})
10	**Return** D_{train},D_{test}

This algorithm is an effective solution to the problem of imbalance in large, multi-class datasets. By applying this methodology, we were able to obtain a test set (D_{test}) that represents all major and minor classes in a balanced manner, ensuring that the model's performance evaluation is fair and accurate. Furthermore, this method provides a rich and diverse training set (D_{train}) that helps the model effectively learn the distinctive features of each class, significantly reducing the likelihood of bias and improving the overall quality of the classification process.

4. Transfer Splitting

The models used in this study are based on the Transformer architecture, a cornerstone of modern natural language processing. This research explores two distinct strategies for implementing transfer learning, using the AraGPT-2 and AraT5 models, which implement the Transformer architecture in fundamentally different ways, resulting in distinct performance and capabilities. This section details the underlying architecture of each model and the fine-tuning methodology applied to adapt them to the hierarchical classification task

Direct Classification Methodology Using AraGPT-2. This approach is based on the AraGPT-2 model, the Arabic version of the GPT-2 (Generative Pre-trained Transformer 2) model developed by OpenAI [6]. AraGPT-2 is a language model based on a unidirectional decoder-only architecture, where each token can only see the tokens preceding it in the sequence, making it ideal for language generation tasks. Each decoder layer consists of a masked multi-head self-attention mechanism and a feed-forward network, with residual correlations and layer normalization to ensure training stability [6]. The attention mechanism in AraGPT-2 is based on the "scaled dot-product self-attention" algorithm, which was first introduced in the transformer architecture [24]. which can be formally described as follows.

Algorithm: Scaled Dot-Product Attention Input: Query matrix Q, Key matrix K, Value matrix V

Compute dot products of the query with all keys: QK^T

Scale the dot products: $\frac{QK^T}{\sqrt{d}}$

Apply a SoftMax function to obtain the weights on the values: $softmax\left(\frac{QK^T}{\sqrt{d}}\right)$

Compute the weighted sum of the values $Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V$

Output: The attention-weighted value matrix.

where Q, K, V are matrices derived from the same input sequence, and dk is the dimension of the key vectors.

The AraGPT-2 model was pre-trained on a massive amount of Arabic text to enable it to understand and generate coherent texts, making it a strong foundation for various Arabic language processing tasks [25]. To leverage and adapt the model for classification, in this study, the aubmindlab/AraGPT-2-base version of the Hugging Face [26]. platform was downloaded. This model was not used for its original generative task, but rather employed as a "feature extractor." The applied architecture consists of passing the input text to the AraGPT-2 model to extract an information-rich numerical representation from its last layer (last_hidden_state). Then, to adapt this understanding to the classification task, a custom class (GPT2ForClassification) was built, adding new layers known as "Classification Heads." Fine-tuning focused on training these new heads, with minor adjustments to the weights of the original model. This process was controlled using the AdamW optimizer and advanced techniques such as using class_weights to address data imbalance, a learning rate regulator with a warm-up period, and early stopping

Generative Classification Methodology Using AraT5. This approach relies on the T5 (Text-to-Text Transfer Transformer) model, a framework developed by Google to unify natural language processing tasks into a text-to-text format. T5 features a full encoder-decoder architecture. The bidirectional encoder reads the entire input text to construct a comprehensive contextual representation, while the unidirectional decoder generates the output sequence symbol by symbol, taking into account the encoder's output. This design makes T5 particularly powerful for tasks that require understanding an entire sequence and then generating a new sequence [7]. This study used the Arabic version of UBC-NLP/AraT5-base, which was specifically trained on Arabic texts [14]. To leverage and adapt the model for the classification task, it was downloaded from the Hugging Face [27]. platform to take advantage of its natural structure. This was achieved by reframing the classification task to suit its generative nature. Instead of adding classification layers, the model was fine-tuned by designing directive inputs (using special tokens) that inform the model of the required task. The model was trained to read the input text and generate a structured string containing category identifiers (e.g., "[CATEGORY] {id} [SUBCAT] {id}"). This was achieved by designing directive inputs and training the model to generate the target string.The two base models, despite their reliance on transformers, differ significantly in their architecture, as summarized in Table 2.

Table 2. Comparison of the infrastructure parameters of the models used in the pre-training phase.

Parameter	**AraGPT2-base**	**AraT5-base**
Architecture	Decoder-Only	Encoder-Decoder
Layers	12	12 for both encoder and decoder
Hidden Size	768	768
Attention Heads	12	12
Total Parameters	~124 million	~220 million

Results And Discussion

This section provides a comprehensive systematic Discussion of fine-tuning experiments and performance evaluation, tracing a multi-stage path starting with the initial tuning and optimization of AraT5 and AraGPT-2 models on different datasets, focusing on the tuning of hyperparameters and the use of regularization techniques to combat overfitting. After the training phase, the transition is made to the analysis of actual performance on large test groups using traditional scales (accuracy, F1-Score) and hierarchical scales that take into account the taxonomic structure. The stage concludes with a comprehensive comparison of the resulting performance and the performance of reference models in the literature, followed by a crucial step is the evaluation of Inference Evaluation on a new, independent and unseen dataset to ensure the generalization power in real-world scenarios, allowing to draw substantial conclusions about the efficiency of each architecture (generative vs. Direct) in the task of hierarchical classification of Arabic texts.

1. Experimental Results

This experimental section is designed for the three experiments tested. The training parameters (hyperparameters) were tuned differently for each model to suit its architecture and the data used, as shown in Table 3. To analyze these results, several fundamental concepts in machine learning were relied upon. The ultimate goal of a model is generalization, which refers to the model's ability to perform well on new and unseen data after being trained on a limited dataset [28]. This ability is estimated using a validation loss. When selection is high, this may indicate underfitting, where the model fails to remove the base model, or overfitting, where the model retains details of the training data instead of learning generalities [28]. To combat overfitting, regularization techniques such as dropout and weight decay are used [29]. These training and optimization processes were performed within a specific computing environment, with all experiments conducted on the Google Colab platform. The differences in computational complexity between the two approaches were reflected in the training time. Initial training of the AraT5 and AraGPT-2 models on the large dataset (75,752 articles) took approximately five and three hours, respectively, using an NVIDIA A100 GPU. Training the optimized AraGPT-2 model on the small dataset (6,000 articles) took only 20 minutes using a T4 GPU.

Training and improvement Optimization. To ensure the validation and continuity of both applications, only different options were applied, selected for the specific details of each model. For the AraGPT-2 model, which includes the addition of new classifier layers, the focus was on stabilizing the training process. The AdamW optimizer was used, a generally improved version of the Adam algorithm that separates the weight update from the weight decay technique, often resulting in better engagement. Additionally, a learning rate scheduler was used, including a warm-up phase (get_linear_schedule_with_warmup) [30]. The diverse training process begins with the warm-up phase and then linearly decreases the learning curve, a new technique among innovations at the beginning of training. As for the AraT5 model, which was trained with a "text-to-text" methodology without major architectural modifications, the AdamW optimizer was also used, but it was different and without a learning rate scheduler, and the model showed good performance even for stable learning.

Table 3. Comparison of fine-tuning parameters for the three models.

Parameter	AraT5 (Seq2Seq)	AraGPT-2 (Initial Classifier)	AraGPT-2(Comparison paper specific)
Learning Rate	3e-5	4e-5	2e-5
Batch Size	16	16	16
Epochs	10	10	15
Max Length	512	512	128
Regularization	AdamW	AdamW, Early Stopping	Dropout=0.4, weight_decay=0.01, EarlyStopping(patience=2)

Model Performance Analysis. To analyze the results, we rely on several fundamental concepts in machine learning. The goal is to achieve generalization, i.e., the model's ability to perform well on new data [28]. This is assessed using the validation loss, with high values indicating either underfitting or overfitting [28]. To combat the latter, regularization techniques are used [29]. Figures 3 and 4 show the learning curves comparing the accuracy and validation loss of the three models. The AraT5 model, represented by the blue curve, shows exceptional performance. In Figure 3, we see that the validation accuracy started very high at 92.42% and steadily increased to peak at 97.21%. Figure 4 supports this stability, with the validation loss steadily decreasing from 0.3579 to 0.1037, indicating excellent generalization ability and a complete absence of overfitting. This strong performance is attributed to the compatibility between the encoder-decoder architecture and the generative methodology, which allowed the model to effectively handle the hierarchical complexity of large-scale datasets using a simple AdamW optimizer. In contrast, the initial performance of the AraGPT-2 model, represented by the orange curve, represents a classic case of underlearning. In Figure 1, the validation accuracy started low at 29.78% and stabilized at a modest level (~60%), while the validation loss remained very high (Figure 4), starting at 1.7801 and ending at 1.0800. Despite the use of advanced optimization strategies such as AdamW and the learning rate scheduler, this indicates that the simple initial parameters listed in Table 6 were insufficient to enable the model to learn complex patterns in the large dataset. Finally, the optimized AraGPT-2 model, represented by the green curve, demonstrates a fine-tuning success story. In Figure 3, we see a sharp and healthy increase in validation accuracy, starting at 65.19% and peaking at 92.35% in the seventh epoch. More importantly, Figure 4 shows that the validation loss reached its lowest point (0.2712) in the same epoch before starting to rise, a clear sign of the onset of overfitting. This outstanding performance is directly attributable to the advanced regularization techniques mentioned in Table 6. The high dropout value (0.4) and weight_decay value (0.01) prevented the model from memorizing the training data, while the low learning rate (2e-5) allowed for more accurate convergence. The Early Stopping mechanism proved effective in capturing the model at its optimal performance point.

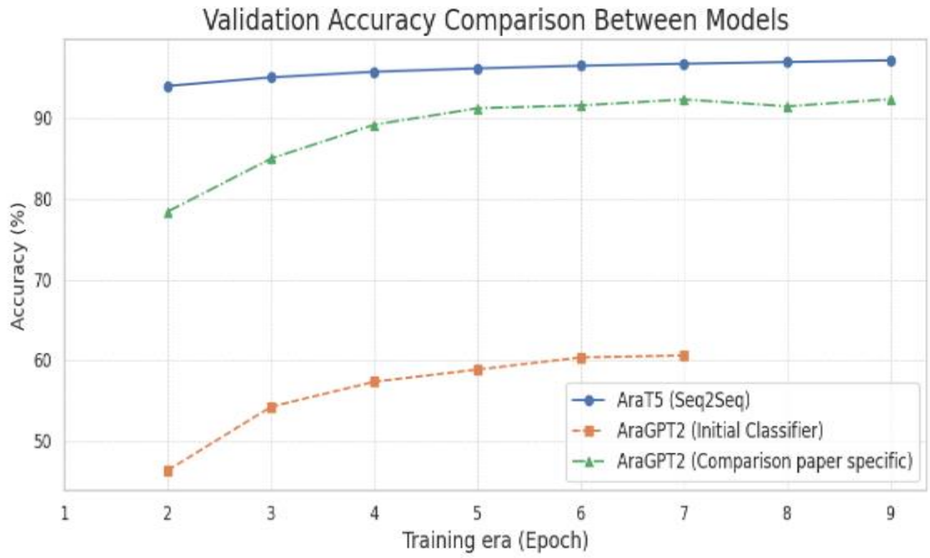

Fig .3. Validation Accuracy Comparison Between Models.

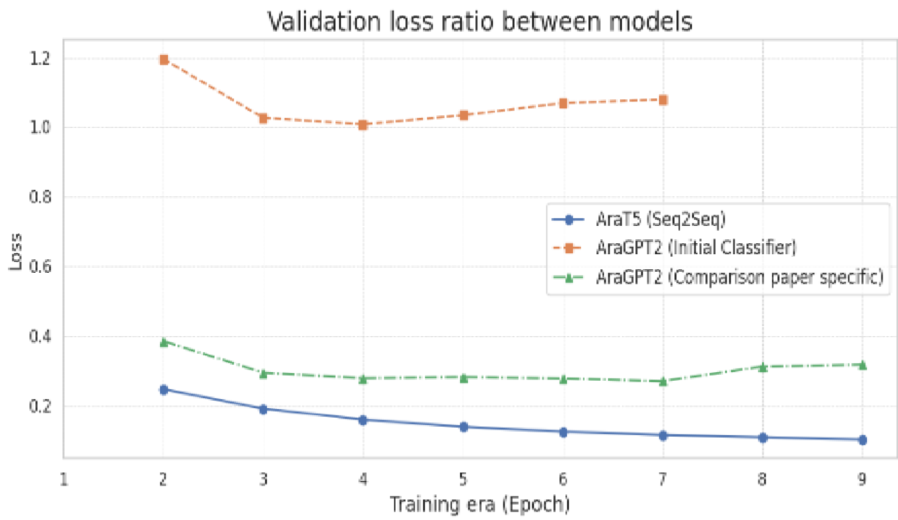

Fig .4. Comparison of validation loss between models.

2. Models Evaluation Results

This section presents a comparative analysis of the performance of the AraT5 and AraGPT-2 models on the hierarchical text classification task. Each model was evaluated using an independent test set consisting of 15,000 articles. The evaluation relied on key quantitative metrics, qualitative error analysis, and visual analysis of the results shown in the appendices. This comparison aims to understand the strengths and challenges of each model and the impact of its architecture (generative versus direct) on overall performance:

AraT5 Model Evaluation Results. The performance analysis of the AraT5 model shows a clear disparity between the two levels of hierarchical classification.

Table 4. Final Performance Metrics for the AraT5 Model on the Test Set

Metric	Main Category	Sub-category
Accuracy	90.45%	66.64%
F1-Score	0.9042	0.6385
Precision	0.9052	0.6565
Recall	0.9045	0.6664

As shown in Table 4, the model achieved exceptional performance in the main categories, with an accuracy of 90.45% and an F1-Score of 0.9042. This is attributed to its Encoder-Decoder architecture, which allows it to comprehend the full context of the text and identify its general domain. Conversely, at the sub-category level, performance was lower but still strong, with an accuracy of 66.64% and an F1-Score of 0.6385. This gap highlights the fundamental challenge of precise classification among 600 potential sub-categories.

Qualitative Error Analysis: The errors made by the model are insightful and not random; they are often "intelligent" and semantically close to the correct answer. As shown in Table 9, these errors can be categorized into two types: hierarchical errors and entirely incorrect classifications. For example, an article on "أسماء سورة التوبة ومعانيها" (The names and meanings of Surah At-Tawbah) was incorrectly classified under "تأملات قرآنية" (Quranic reflections) instead of "قرآن" (Quran), although the main category "إسلام" (Islam) was correct.

AraGPT-2 Model Evaluation Results (Direct Classification Approach). The AraGPT-2 model demonstrated strong performance, comparable to AraT5, confirming the effectiveness of the direct classification approach.

Table 5. Final Performance Metrics for the AraGPT-2 Model on the Test Set

Metric	Main Category	Sub-category
Accuracy	89.42%	65.95%
F1-Score	0.8946	0.6383
Precision	0.9019	0.6636
Recall	0.8942	0.6595

As seen in Table 5, the model achieved an F1-Score of 0.8946 and an accuracy of 89.42% in the main categories, indicating high reliability in the general classification task. Performance decreased as expected at the sub-category level, with an F1-Score of 0.6383 and an accuracy of 65.95%, which is a good result given the complexity of the task.

Qualitative Error Analysis: The errors made by this model were similar to those of AraT5, as detailed in Table 6. They include hierarchical errors, where the correct sub-category is placed under the wrong main category, and semantically close misses, where the model's predictions are logically reasonable but taxonomically incorrect. For instance, an article about "الكركديه والضغط" (Hibiscus and Blood Pressure) was incorrectly classified under "تغذية" (Nutrition) instead of "فن الطهي" (Culinary Arts), an error also made by the AraT5 model, suggesting a potential inherent ambiguity in this data point.

Table 6. Examples of Classification Errors

Model	Title	Correct Main Category	Correct Sub-category	Predicted Main Category	Predicted Sub-category
AraGPT-2	العالم جيلبرت لويس	تعليم	علماء	الحياة والمجتمع	علماء
AraGPT-2	كلام الجمعة الثانية رمضان	الآداب	عبارات وكلمات	الآداب	حكم وأقوال

AraT5 & AraGPT-2	الكركديه والضغط	فن الطهي	مشروبات باردة وساخنة	تغذية	فوائد الأعشاب
AraT5	أسماء سورة التوبة ومعانيها	إسلام	قرآن	إسلام	تأملات قرآنية

The AraT5 and AraGPT-2 models prove to be powerful and effective tools for hierarchical classification, showing convergent performance. The two models excel at identifying high-level topics, and while they face challenges in accurately categorizing among 600 categories, their errors are often linguistically plausible. This suggests that both models have developed a deep contextual understanding from their pre-trained knowledge, regardless of their different structure

3. Comparison and Discussion of Results

The AraT5 and AraGPT-2 models prove to be powerful and effective tools for hierarchical classification, showing convergent performance. The two models excel at identifying high-level topics, and while they face challenges in accurately categorizing among 600 categories, their errors are often linguistically plausible. This suggests that both models have developed a deep contextual understanding from their pre-trained knowledge, regardless of their different structure

Analysis Based on Traditional Metrics. The evaluation using traditional classification metrics confirms the model's high effectiveness. On a test set of 1,764 texts, the model achieved a Category Accuracy of 97.11% and a Sub-category Accuracy of 92.86%. These figures indicate a superior capability in distinguishing between broad categories and maintaining excellent performance even at the more granular sub-category level. To provide a deeper insight beyond accuracy, metrics such as Precision, Recall, and F1-Score were calculated. Precision measures the proportion of correct positive predictions among all positive predictions made, while Recall measures the proportion of actual positives that were correctly identified. The F1-Score, as the harmonic mean of Precision and Recall, offers a balanced assessment of the model's performance, which is especially important in the presence of class imbalance[31]. These are calculated using the following standard equations:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (1)$$

$$\text{Recall: } \frac{\text{TP}}{\text{TP+TN}} \quad (2)$$

$$\text{F1−Score}=2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}} \quad (3)$$

Where TP, FP, and FN represent True Positives, False Positives, and False Negatives, respectively.

Analysis Based on Hierarchical Metrics. In hierarchical classification tasks, traditional metrics can be misleading as they penalize all errors equally, regardless of their proximity within the category tree. Therefore, Hierarchical Metrics are employed, which incorporate the taxonomy's structure into the evaluation[32]. To facilitate a direct comparison with the reference paper, the same hierarchical evaluation approach was adopted. These metrics are calculated based on the hierarchical distance $(d(c, c\text{^}))$ between the true class (c) and the predicted class $(c\text{^})$. This allows for a more nuanced evaluation where, for instance, a misclassification within the same parent category is penalized less severely than a misclassification into a completely different branch of the hierarchy [32]. The core hierarchical metrics are defined as:

$$\text{hP}=\frac{1}{|\text{D}|} \sum_{(\text{X,C})} \in \text{D}\big(1 - \text{d}(\text{c}, \text{c\^{}})\big) \quad (4)$$

$$\text{hR}=\frac{1}{|\text{D}|} \sum_{(\text{X,C})} \in \text{D}\big(1 - \text{d}(\text{c}, \text{c\^{}})\big) \quad (5)$$

$$hF = \frac{2*hp*hR}{hP+hR} \quad (6)$$

Were D is the dataset, x is the text. c is the true class, and $c\text{^}$ is the predicted class.

The results, summarized in Table 7, show that the model achieved excellent performance according to both traditional and hierarchical metrics at both the main and sub-category levels.

Table 7. Comparison of Traditional and Hierarchical Performance Metrics for the Refined Model.

Metric	Main Categories	Sub-Categories
F1-Score	0.9711	0.9286
Precision	0.9714	0.9292
Recall	0.9711	0.9286
HF1-Score	0.9711	0.9354
HPrecision	0.9711	0.9354
HRecall	0.9711	0.9354

Performance analysis shows that the discrepancy between traditional and hierarchical measures confirms that errors in hierarchical classification are not equal. While traditional measures report an error rate of 7.14% across subcategories, hierarchical measures show greater tolerance for semantically close predictions. For example, predicting the category "Math/Analysis" instead of "Math/Algebra" is a complete miss under traditional measures but acceptable under hierarchical measures because it falls within the same parent category. Consequently, the high hierarchical F1 value of 0.9354 demonstrates that the model possesses a deep understanding of the semantic structure of the data, making it suitable for hierarchical classification.

Compare Results with Reference Paper Results. After independently validating the model's effectiveness, a comprehensive comparison of its performance with another leading model in the field of hierarchical classification of Arabic texts was conducted. This comparison aims to evaluate the competitiveness of our refined AraGPT-2 model and measure its superiority over existing state-of-the-art methods. The comparison focuses on key metrics that express accuracy at various hierarchical levels, including main-class accuracy, sub-class accuracy, and the hierarchical F1 score, which is a comprehensive measure of model performance in a hierarchical environment.

Table 8. Performance Comparison of the Refined AraGPT-2 Model with the Reference Model.

Model	Hierarchical F1-Score (hF1)	Main Category Accuracy	Sub-Category Accuracy
Model from Reference	80.64%	83.64%	80.98%
Our Refined AraGPT-2 Model	93.54%	97.11%	92.86%
Improvement	+12.9%	+13.47%	+11.88%

Table 8 shows a clear comparison between the performance of our improved model (AraGPT-2) and the reference model from the research paper. Our model's clear superiority is evident across all three key metrics. At the main-class accuracy level, our model achieved a significant improvement of +13.47%, demonstrating its superior ability to identify general text categories with very high accuracy compared to the reference model. At the sub-class accuracy level, a metric that reflects the model's ability to handle fine-grained details, our model outperformed the reference model by an improvement of +11.88%, confirming the effectiveness of our fine-tuning strategy. Most importantly, at the hierarchical F1 score (hF1), the most comprehensive and accurate metric for assessing hierarchical classification performance, our model achieved a massive improvement of +12.9%, confirming that the model not only excels at correct predictions but also possesses a deeper and more coherent understanding of the hierarchical structure of the data, reducing the penalty for semantically close errors. These results confirm that the improved AraGPT-2 model outperforms the model in the reference paper.

4. Inference Evaluation Models

In this study, an assessment of the generalization ability of trained models was presented, a final inference test was performed on a new, independent dataset. This collection consists of 200 articles collected from the Arabic content site sotor[33]. a source that was not used in any of the previous training or verification stages. This process measures the ability of each model to apply its acquired knowledge to real-world data that comes from a different distribution. The evaluation was conducted using both traditional (Weighted F1-Score) and hierarchical (Hierarchical F1-Score) scales to provide a comprehensive and integrated view of performance .

AraGPT-2 Model Inference. The AraGPT-2 model, following the direct classification approach, showed a strong ability to generalize to new data. At the level of the main class, the model achieved excellent performance with an accuracy of 94.50% and a conventional F1-Score scale of 0.9370. This confirms that the structure of feature extraction with classification headers is able to effectively delimit the general area of the article. At the subcategory level, the performance was very good, with an F1-Score of 0.7874. One of the most important observations on the performance of this model is the difference between its traditional full accuracy (79.00%) and its total hierarchical accuracy scale (87.50%). This significant difference indicates that a significant part of the model's errors were of the type of "hierarchically Near-Misses", that is, even when he made a mistake in determining the exact subcategory, his prediction was often within the correct main category, which reduced the punishment imposed by hierarchical scales.

Arat5 Model Inference. The AraT5 model, which relies on a generative approach, demonstrated exceptional and consistent performance on unseen data, outperforming its counterpart on most metrics. At the main class level, the model achieved an accuracy of 96.97% and an F1-Score of 0.9687. More importantly, on the accurate classification task at the subclass level, it achieved a robust F1-Score of 0.8534. The most impressive result is that the AraT5 model's overall hierarchical accuracy (hF1) reached 91.67%, significantly higher than its traditional full accuracy (86.36%). This significant difference confirms that the model's errors were almost exclusively of the "hierarchically close" type. This behavior demonstrates that the generative approach not only learned how to classify but also learned the structural relationships between main and subclasses, ensuring it adheres closely to the hierarchical structure of the data. Table 9. also summarizes a comprehensive comparison and analysis of the results and performance of both models on the dataset collected for verification, allowing for a thorough analysis of the significant differences in their performance.

Table 9. Detailed Performance Comparison Between AraT5 and AraGPT-2 models.

Models		**AraT5 Model**		**AraGPT2 Model**	
Metric Type	**Metric**	**Main category**	**Subcategory**	**Main category**	**Subcategory**
traditional	Accuracy	96.97%	86.36%	94.50%	80.50%
	F1-Score (weighted)	0.9687	0.8534	0.9370	0.7874
	Precision (weighted)	0.9713	0.8656	0.9344	0.8188
	Recall (weighted)	0.9697	0.8636	0.9450	0.8050
Hierarchical	HF1-Score	0.9697	0.9167	0.9370	0.7874
	HPrecision	0.9697	0.9167	0.9344	0.8188

	HRecall	0.9697	0.9167	0.9450	0.8050

Table 9: provides profound insights when comparing the two models. The results show that both models possess strong generalization capabilities, but their behavior and performance differ, revealing the strengths and weaknesses of each approach. The AraT5 model clearly outperforms all traditional metrics, especially at the subclass level, with a difference of nearly 7 percentage points in the F1-score. This demonstrates that its generative approach was more robust in arriving at completely correct and accurate predictions. Furthermore, AraGPT-2 exhibits interesting unique behavior: Consider the following example:

- *Text*: " القوات السودانية تصد هجوما واسعا للدعم السريع على الفاشر.... "
- *Correct classification:* Worldwide -> Cities and countries
- *AraT5 prediction*: "حول العالم -> مدن وبلدان" (structurally and logically correct)
- *AraGPT-2 prediction:* "حول العالم -> ثورات وحروب".

Therefore, AraGPT-2 made a hierarchical mistake, since the category "revolutions and wars" does not belong to "around the world". Semantically, however, his prediction of the subcategory was more accurate and related to the content of the text than the structurally correct prediction of AraT5. ("Revolutions and wars"), ignoring the fact that it does not belong to the main category predicted by the other head ("around the world"). This suggests that the classification headers work semi-independently, which can lead to very accurate predictions at the hierarchical rule-breaking calculation Sublevel. This is not so much a classification error as a shortcoming in the design of the structure of the categories themselves, since some subcategories may be closely related to more than one main category. In turn, AraT5 shows a strict adherence to the hierarchical structure. The generative methodology "text-to-text" has learned that the generation of a certain subcategory directly depends on the main one that was generated before it in the same sequence. This "structural awareness" is the main reason that its hierarchical scale (91.67%) is much higher than its traditional one, which is what makes it more reliable in applications that require maintaining a hierarchy.

Conclusion

This study provided a comprehensive investigation of the application of advanced transducer models to the complex task of hierarchical classification of Arabic texts. Two distinct fine-tuning methodologies were evaluated and compared: a "text-to-text" generative approach using the AraT5 model, and a direct classification approach using the AraGPT-2 model. The experimental results led to several main conclusions. First, both the generative and direct approaches have proven to be highly effective, demonstrating the deep ability of pre-trained linguistic models to cope with the intricacies of the Arabic language and complex hierarchical structures. The generative AraT5 model showed a slight but consistent outperformance over the large-scale and semantically ambiguous dataset, indicating that its architecture is exceptionally well suited for tasks requiring strong hierarchical cohesion. Secondly, this research has successfully proved that systematic fine - tuning and the application of advanced regulation techniques can lead to significant performance gains. The improved AraGPT-2 model significantly outperformed the reference model from the previous literature, achieving a modern hierarchical F1 score of 93.54% and establishing a new performance standard. Finally, the analysis revealed that the fundamental challenge in this task lies in the complexity and ambiguity inherent in the space of classifications, since both models showed sophisticated and context-aware thinking even when making mistakes

Future Work:

Based on the results of this study, several promising avenues for future research can be identified. First, one of the next crucial steps is to conduct a study focused on improving and enriching the hierarchical classification of a large-scale data set. This may include the use of semi-automated methods for combining semantically similar subcategories, the possibility of adding missing new categories to create a more powerful and less ambiguous classifier space. Evaluating the performance of models on this improved rating can reveal additional insights into the impact of the quality of ratings. Secondly, it would be useful to expand the empirical comparison to other modern architectures. One of the most prominent candidates for this is a model based on AraBERT, which uses only the structure of the encoder (Encoder-Only). Comparing its performance with that of a typical encoder-decoder (AraT5) and decoder-only (AraGPT-2) would complete a comprehensive analysis of the three main transformer architectures for this task. Finally, the models developed in this study can be applied and tested to different areas of Arabic texts, such as social media content or official legal documents, to assess their adaptability and performance across various linguistic styles and dialects.

References

1. Boboqulovna, K.M.: The Linguistic Levels: Classification, Characteristics, And Interrelationships. Int. J. Lit. Lang. 5, 26–30 (2025). https://doi.org/10.37547/ijll/volume05issue04-08.
2. Customer, I.: Artificial Intelligence in the Impact of Natural Language Processing (NLP) on Artificial Intelligence in the Impact of Natural Language Processing (NLP) on International Customer Engagement via Chatbots. (2025).
3. Zulqarnain, M., Ghazali, R., Hassim, Y.M.M., Rehan, M.: A comparative review on deep learning models for text classification. Indones. J. Electr. Eng. Comput. Sci. 19, 325–335 (2020). https://doi.org/10.11591/ijeecs.v19.i1.pp325-335.
4. Gasparetto, A., Marcuzzo, M., Zangari, A., Albarelli, A.: Survey on Text Classification Algorithms: From Text to Predictions. Inf. 13, 1–39 (2022). https://doi.org/10.3390/info13020083.
5. Kumar, L.A., Renuka, D.K.: State-of-the-Art Natural Language Processing. Deep Learn. Approach Nat. Lang. Process. Speech, Comput. Vis. 49–75 (2023). https://doi.org/10.1201/9781003348689-3.
6. Gummadi, V. P. K. (2019). Microservices architecture with APIs: Design, implementation, and MuleSoft integration. Journal of Electrical Systems, 15(4), 130–134.
7. Raffel, C., Roberts, A., Matena, M., Liu, P.J.: Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. 21, 1–67 (2020).
8. Bouchiha, D., Bouziane, A., Doumi, N., Hamzaoui, B.: Fine-tuning AraGPT2 for Hierarchical Arabic Text Classification. 5, 55–65 (2025).
9. Alahmari, S., Atwell, E., Saadany, H.: Sirius _ Translators at OSACT6 2024 Shared Task : Fin-tuning Ara-T5 Models for Translating Arabic Dialectal Text to Modern Standard Arabic. 117–123 (2024).
10. Aliwy, A.H., Ameer, E.H.A.: Comparative Study of Five Text Classification Algorithms with their Improvements. 12, 4309–4319 (2017).
11. El-halees, A.M.: Arabic Text Classification Using Maximum Entropy. 15, 157–167 (2007).
12. Bhavani, A., Santhosh Kumar, B.: A Review of State Art of Text Classification Algorithms. Proc. - 5th Int. Conf. Comput. Methodol. Commun. ICCMC 2021. 1484–1490 (2021). https://doi.org/10.1109/ICCMC51019.2021.9418262.
13. Boughareb, D., Boughareb, R., Hallaci, S., Boukherouba, M.R.E., Seridi, H.: Addressing the Complexity of Dialectal Arabic: An Enhanced Encoder-Decoder Ensemble Approach for Optimized Sentiment Analysis. ACM Trans. Asian Low-Resource Lang. Inf. Process. 24, (2025). https://doi.org/10.1145/3735972.
14. Moatez, E., Nagoudi, B., Abdul-mageed, M.: AraT5 : Text-to-Text Transformers for Arabic Language Generation. 1, 628–647 (2022).
15. Bouchiha, D., Naama, C.U. De, Abdelghani, B., Doumi, N.: BERT-BiLSTM model for hierarchical Arabic text classification BERT-BiLSTM model for hierarchical Arabic text classification. (2024). https://doi.org/10.52549/ijeei.v12i3.5659.
16. Mqall, https://maqall.net/, last accessed 2025/10/02.
17. Mawdoo3, https://mawdoo3.com/, last accessed 2025/10/02.
18. mhtwyat, https://mhtwyat.com/, last accessed 2025/10/02.
19. mqalaty, https://mqalaty.net/, last accessed 2025/10/02.
20. Khder, M.A.: Web Scraping or Web Crawling : State of Art , Techniques , Approaches and Application. 13, (2021). https://doi.org/10.15849/IJASCA.211128.11.
21. Abodayeh, A.: Web Scraping for Data Analytics : A BeautifulSoup Implementation. (2024). https://doi.org/10.1109/WiDS-PSU57071.2023.00025.
22. Henning, S., Beluch, W., Fraser, A., Friedrich, A.: A Survey of Methods for Addressing Class Imbalance in Deep-Learning Based Natural Language Processing. 523–540 (2023).
23. Sharma, G.: Pros and cons of different sampling techniques. 3, 749–752 (2017).
24. Furfari(tony), F.A.: The Transformer. IEEE Ind. Appl. Mag. 8, 8–15 (2002). https://doi.org/10.1109/2943.974352.
25. Antoun, W., Baly, F., Hajj, H.: A RA GPT2 : Pre-Trained Transformer for Arabic Language Generation. 196–207 (2021).
26. Aubmindlab: Arabic GPT2, https://huggingface.co/aubmindlab/aragpt2-base, last accessed 2025/02/01.
27. UBC-NLP: AraT5: Text-to-Text Transformers for Arabic Language Generation, https://huggingface.co/UBC-NLP/AraT5-base, last accessed 2025/02/01.
28. Kim, K.G.: Deep Learning. 22, 351–354 (2016).
29. Hinton, G.: Dropout : A Simple Way to Prevent Neural Networks from Overfitting. 15, 1929–1958 (2014).
30. Schmidt, R.M., Schneider, F., Hennig, P.: Descending through a Crowded Valley —

Benchmarking Deep Learning Optimizers. (2021).
31. Tharwat, A.: Classification assessment methods. 17, 168–192 (2025). https://doi.org/10.1016/j.aci.2018.08.003.
32. Zangari, A., Marcuzzo, M., Rizzo, M., Gasparetto, A., Giudice, L., Albarelli, A.: Hierarchical Text Classification and Its Foundations : A Review of Current Research. (2024).
33. sotor, https://sotor.com/, last accessed 2025/10/02.

Appendix

List view main and sub categories A1

إجمالي المقالات الأصلية: 75252

عدد الأصناف الرئيسية: 29

عدد الأصناف الفرعية: 600

عدد المقالات في كل صنف:

الصنف الرئيسي	عدد المقالات
أبراج	778
أدعية دينية	1060
إسلام	8906
الآداب	3451
الأسرة	2669
الأمومي	390
الأنثوي	1431
الحياة والمجتمع	1942
الخليج العربي	1970
السيارات	51
العناية بالذات	4858
المنوع	2644
تجارة ومشاريع	1602
تسلية وألعاب	507
تعليم	12488
تغذية	3774
تفسير الاحلام والرؤى	1900
تقنية	2534
ثقافة عامة	931
حكم وأقوال	3052
حول العالم	3166
رجيم ورشاقة	325
رياضة	893
سياحة وسفر	324
صحة	4630
طبي	3334
علوم الأرض	1530
فن الطهي	3717
قصص وحكايات	395

الصنف الرئيسي: الخليج العربي .

----

1. السعودية
2. الكويت
3. الإمارات
4. منصات حكومية
5. سلطنة عمان
6. قطر
7. البحرين

الصنف الرئيسي: تعليم .

----

1. ثقافة عامة
2. علوم
3. تاريخ وجغرافيا
4. رياضيات
5. كليات وجامعات
6. كيمياء
7. فيزياء
8. متنوع
9. قواعد اللغة العربية
10. إذاعة مدرسية
11. اللغة العربية
12. تعريفات منوعة
13. أبحاث وتقارير
14. مفاهيم عامة
15. كتب ومؤلفات
16. مواضيع دينية متفرقة
17. فن الكتابة والتعبير
18. موضوعات تعبير
19. أساليب التعليم
20. مكتبة الطالب
21. جسم الإنسان
22. منصات تعليمية
23. تنمية بشرية
24. مهارات إدارية
25. سؤال وجواب
26. ثقافة
27. تخصصات جامعية
28. تاريخ
29. لغة
30. جغرافيا
31. علماء

31. علماء
32. معالم وآثار
33. اختراعات واكتشافات
34. مهارات الحياة
35. تاريخ الدول
36. شخصيات تاريخية
37. أعلام
38. العلوم الإنسانية
39. تنمية الشخصية والقدرات
40. علم النفس
41. فقهاء وأئمة
42. وسائل النقل
43. دول عربية
44. تعريفات وقوانين علمية
45. المجموعة الشمسية
46. سياحة
47. تنمية الذات
48. اسئلة علمية
49. مصطلحات ومعاني
50. تنمية المهارات الشخصية
51. البحث العلمي
52. مهارات دراسية
53. دراسات وأبحاث
54. تنمية الذكاء
55. مبادئ التعليم
56. جبال ووديان
57. التخصصات الجامعية
58. حضارات
59. ملوك وأمراء
60. جامعات ومعاهد
61. مهارات التواصل
62. اسئلة تاريخية
63. تعلم اللغات
64. صناعات
65. علوم الأرض
66. الغطاء النباتي
67. قصص عربية
68. الحروف الابجدية
69. أبحاث علمية
70. مهارات وأساليب
71. موضوعات تعليمية
72. تطوير الذات
73. مهارات فردية
74. جغرافيا و تاريخ
75. ثورات وحروب
76. دول قارة أوروبا
77. أساليب التعلم
78. التعليم الأساسي
79. معاني ومصطلحات

الصنف الرئيسي: إسلام

----

1. معلومات دينية
2. أدعية وأذكار
3. تفاسير قرآنية
4. أحكام شرعية
5. كتب ومؤلفات
6. مواضيع دينية متفرقة
7. قرآن
8. القرآن الكريم
9. شخصيات إسلامية
10. ثقافة إسلامية
11. رمضان
12. أحاديث
13. الصحابة والتابعون
14. معلومات إسلامية
15. أسئلة دينية
16. أحاديث نبوية
17. معتقدات إسلامية
18. رسل و أنبياء
19. أخلاق إسلامية
20. حكم و مواعظ دينية
21. دعاء 1
22. تعريفات إسلامية
23. شهر رمضان
24. السيرة النبوية
25. فقهاء وأئمة
26. عبادات
27. مصطلحات إسلامية
28. التاريخ الإسلامي
29. أحكام الشريعة الاسلامية
30. وضوء و صلاة
31. الأسرة في الإسلام
32. معارك و غزوات
33. قصص إسلامية
34. أذكار
35. مناسبات عربية وعالمية
36. معالم إسلامية
37. النوافل
38. حياة الرسول و الصحابة
39. دعاء 2
40. دعاء 3
41. صحابيات
42. فروض وسنن
43. دعاء 4
44. الحج والعمرة
45. معاملات إسلامية
46. وضوء وطهارة
47. أسئلة دينية
48. معلومات اسلامية
49. وضوء وصلاة
50. الاسرة في الاسلام
51. رسل وأنبياء
52. قصص دينية
53. أركان الإسلام والإيمان
54. الاستخارة
55. حساب الزكاة

56. منوعات اسلامية
57. النبي محمد
58. طرق حفظ القرآن الكريم

الصنف الرئيسي: المنوع

---

1. الرياضة
2. تسلية وترفيه
3. ألغاز مع الحلول
4. ترددات القنوات
5. سؤال وجواب
6. أفضل
7. معلومات ونصائح طبية
8. فوازير وحزازير واختبارات
9. معلومات مفيدة
10. معلومات ثقافية
11. معلومات رياضية
12. معلومات تاريخية وجغرافية
13. أين تقع مدن العالم الأجنبية
14. تأمينات
15. أين تقع مدن العالم العربية
16. عطور
17. أين تقع دول العالم العربي
18. أين تقع دول العالم الأجنبي

الصنف الرئيسي: الأداب

---

1. عبارات وكلمات
2. كتب ومؤلفات
3. فن الكتابة والتعبير
4. أشعار منوعة
5. أدباء وشعراء
6. شعر عربي
7. قصص وشعر
8. فنون أدبية
9. أقوال
10. متفرقات أدبية
11. كلمات معبرة
12. أشعار حب
13. كتب وروايات
14. شعر حزين
15. دول عربية
16. عبارات حزينة
17. حكم المشاهير
18. نكت و رسائل
19. شعر مدح
20. شعر حب وغزل
21. شعر مديح و رثاء
22. شعر غزل
23. حكم وأقوال
24. منوعات أدبية

الصنف الرئيسي: تقنية

---

1. انترنت وكمبيوتر
2. تقنيات منوعة
3. التواصل الاجتماعي
4. جوالات وتطبيقات
5. كمبيوتر
6. موبايل
7. انترنت
8. مواقع التواصل الإجتماعي
9. تطبيقات الكترونية
10. برامج الكترونية
11. ألعاب إلكترونية
12. برمجة وتصميم المواقع
13. اسئلة تقنية
14. مواقع التواصل الاجتماعي

الصنف الرئيسي: العناية بالذات

---

1. العناية بالشعر
2. العناية بالبشرة
3. مواضيع دينية متفرقة
4. العناية بالوجه
5. رجيم وأنظمة غذائية
6. مهارات إدارية
7. العناية بالجسم
8. تفتيح لون البشرة
9. تحاليل طبية
10. صبغات الشعر
11. تمارين رياضية
12. صحة الشعر
13. فوائد الزيوت للشعر
14. الهالات والرؤوس السوداء
15. مهارات الحياة
16. الشاي والقهوة
17. تجميل ومكياج
18. تنظيف وتقشير البشرة
19. طرق زيادة الوزن
20. وزن ورشاقة
21. حب الشباب
22. أزياء وملابس
23. وصفات تطويل الشعر
24. حليّ ومجوهرات
25. مشاكل تساقط الشعر
26. وصفات تكثيف الشعر
27. جمال
28. العناية بالأظافر
29. صحة البشرة
30. حبوب البشرة
31. تجميل الأسنان
32. العناية بالشعر التالف
33. العناية بفروة الرأس
34. وصفات تنعيم الشعر
35. التخلص من التعرق
36. العناية بالبشرة الدهنية
37. فوائد الزيوت للبشرة
38. متفرقات
39. العناية باليدين
40. التخلص من الحشرات
41. جمال البشرة
42. منوعات العناية بالذات
43. العناية بالقدم
44. قصات وتسريحات للشعر
45. حلويات شرقية
46. طرق تنزيل الوزن
47. قصات وتسريحات الشعر
48. العناية بالفم و الأسنان
49. العناية بالفم والأسنان
50. العناية بالشعر المصبوغ
51. العناية بالبشرة الجافة والحساسة

الصنف الرئيسي: الأنثوي

---

1. العناية بالشعر
2. العناية بالبشرة
3. صحة الحامل
4. العناية بالطفل
5. تغذية الحامل
6. العناية الشخصية
7. أمراض الحمل والولادة
8. الحياة الزوجية
9. مراحل نمو الجنين
10. العناية بحديثي الولادة
11. صحة الجنين
12. مراحل الحمل
13. أزياء وموضة
14. المرأة الحامل
15. كيف أهتم بطفلي

الصنف الرئيسي: طبي

---

1. الحمل والولادة
2. أمراض وعلاجات
3. معلومات طبية
4. جسم الإنسان
5. تحاليل طبية
6. دليل الأدوية
7. أمراض جلدية
8. أمراض الجهاز الهضمي
9. أمراض الدم
10. السرطان
11. مرض السكري
12. أمراض الجهاز التنفسي
13. مصطلحات طبية
14. ضغط الدم
15. أمراض الكبد والمرارة
16. أمراض معدية
17. آلام الرأس
18. أمراض الحساسية
19. أمراض القولون
20. جراحة عامة
21. أمراض القدم
22. عادات صحية
23. التخلص من التعرق
24. أمراض الأوعية الدموية
25. آلام الظهر والرقبة
26. المتلازمات
27. أمراض صدرية
28. الحروق
29. الصحة النفسية
30. منوعات طبية
31. اضطرابات النوم وعلاجها
32. صحة الفم والأسنان
33. ارتفاع الحرارة والحمى
34. أمراض الأعصاب وعلاجها
35. عظام وروماتيزم
36. أمراض الأطفال وعلاجها
37. كلى ومسالك بولية
38. أنف وأذن وحنجرة
39. نسائية وتوليد
40. أمراض القلب والشرايين
41. أمراض العيون وعلاجها
42. أمراض الغدد وعلاجها
43. باطني وقناة هضمية

الصنف الرئيسي: الأسرة

---

1. تفسير الأحلام
2. معاني أسماء
3. تنمية بشرية
4. شؤون منزلية
5. العناية بالطفل
6. معاني الأسماء
7. التدبير المنزلي
8. ديكورات
9. العناية بحديثي الولادة
10. صحة الجنين
11. علاقات أسرية
12. مناسبات عربية وعالمية
13. الحمل و الولادة
14. مراحل الحمل
15. صفات الأبراج
16. متفرقات
17. التخلص من الحشرات
18. منوعات أسرة وتسلية
19. طب عام
20. قصص عربية
21. تربية الأطفال
22. الزواج والحب
23. حفظ الأطعمة
24. قصص عالمية

الصنف الرئيسي: الحياة والمجتمع

---

1. معلومات عامة
2. احلام
3. تعريفات منوعة
4. مفاهيم عامة
5. قضايا مجتمعية
6. لغة
7. العناية بالطفل
8. حكم
9. ظواهر اجتماعية
10. أعلام
11. وسائل النقل
12. علاقات أسرية
13. حكم وأقوال في الصداقة
14. مناسبات عربية وعالمية
15. حكم متنوعة
16. مهارات دراسية
17. عادات صحية
18. حكم وأقوال في الحياة
19. مبادئ التعليم
20. حضارات
21. ملوك وأمراء
22. التعامل مع المراهقين
23. عادات وتقاليد
24. أضرار التدخين

الصنف الرئيسي: تغذية

---

1. معلومات غذائية
2. رجيم وأنظمة غذائية
3. فوائد الأعشاب
4. فوائد وأضرار الأطعمة
5. فيتامينات ومعادن
6. منتجات غذائية
7. الشاي والقهوة
8. تغذية الحامل
9. فوائد الفواكه
10. حليب وأجبان
11. حلو عربي
12. العسل
13. فوائد الخضروات
14. تخفيف الوزن
15. فوائد الفيتامينات والمعادن
16. فوائد الزيوت
17. نقص الفيتامينات والمعادن
18. مصادر الفيتامينات والمعادن
19. فوائد البذور
20. طرق حرق الدهون
21. تغذية الطفل
22. التداوي بالأعشاب
23. فوائد الحبوب
24. الفاكهة والخضراوات
25. ماء الورد
26. فوائد البقوليات
27. منوعات طبية
28. فوائد الفواكه والخضروات
29. الزيوت النباتية
30. الحميات الغذائية
31. التخلص من الكرش
32. الاطعمة وفوائدها للجسم
33. التخلص من السمنة

الصنف الرئيسي: حول العالم

---

1. مدن عربية
2. مدن أجنبية
3. أنظمة دولية
4. معالم وآثار
5. جزر العالم
6. مدن وبلدان
7. دول أجنبية
8. مدن ومحافظات
9. الكثافة السكانية
10. دول عربية
11. السياحة في الدول الأجنبية
12. سياحة
13. معالم سياحية
14. السياحة في الدول العربية
15. مناسبات عربية وعالمية
16. عواصم
17. جزر سياحية
18. مساحات الدول
19. وجهات سياحية
20. جزر وقارات
21. دول قارة آسيا
22. دول قارة أفريقيا
23. بماذا تشتهر
24. ممالك وجمهوريات
25. دول قارة أوروبا

الصنف الرئيسي: تسلية وألعاب

---

1. احلام
2. جسم الإنسان
3. أقوال
4. متفرقات أدبية
5. موبايل
6. حكم
7. وسائل النقل
8. السياحة في الدول الأجنبية
9. حكم في الحب
10. تحليل الشخصية
11. جمال
12. عبارات نهائي
13. شخصية و أبراج
14. صفات الأبراج
15. نكت و رسائل
16. منوعات أسرة وتسلية
17. مهارات فردية
18. الألوان ودلالاتها
19. قصص عالمية

الصنف الرئيسي: علوم الأرض

---

1. معلومات عن الدول
2. ظواهر طبيعية
3. ثروات طبيعية
4. زراعة
5. المجموعة الشمسية
6. التلوث البيئي
7. بحار ومحيطات
8. أنهار وبحيرات
9. جبال ووديان
10. تعداد سكان الدول
11. زراعة الخضراوات والفواكه
12. البترول ومشتقاته
13. خرائط الدول

الصنف الرئيسي: صحة

---

1. مواضيع طبية متفرقة
2. معلومات و نصائح طبية
3. العناية بالجسم
4. نسائية و توليد
5. تحاليل طبية
6. صحة الحامل
7. أمراض جلدية
8. حكم و مواعظ دينية
9. الشاي والقهوة
10. طب بديل
11. أمراض الجهاز الهضمي
12. أمراض الدم
13. أمراض الحمل والولادة
14. السرطان
15. مرض السكري
16. كلى و مسالك بولية
17. حب الشباب
18. أطفال
19. أمراض الجهاز التنفسي
20. مراحل نمو الجنين
21. صحة الفم و الاسنان
22. اضطرابات النوم وحلولها
23. مصطلحات طبية
24. صحة الجنين
25. ضغط الدم
26. أمراض الكبد والمرارة
27. أمراض معدية
28. عيون
29. أنف و أذن و حنجرة
30. آلام الرأس
31. الحمل و الولادة
32. أمراض الحساسية
33. اضطرابات نفسية
34. عظام و روماتيزم
35. أمراض القولون
36. جراحة عامة
37. غدد
38. أمراض القدم
39. أمراض الأوعية الدموية
40. آلام الظهر والرقبة
41. التداوي بالأعشاب
42. صحة نفسية
43. المتلازمات
44. أمراض صدرية
45. اضطرابات القناة الهضمية
46. متفرقات
47. الحروق
48. أمراض القلب و الشرايين
49. أعصاب
50. باطني و قناة هضمية
51. الصحة الجنسية
52. منوعات طبية
53. طب عام
54. العناية بالقدم
55. الحمى وإرتفاع الحرارة
56. أمراض الأطفال والمراهقين

الصنف الرئيسي: رجيم ورشاقة

---

1. رجيم وأنظمة غذائية
2. طرق زيادة الوزن
3. طرق حرق الدهون
4. طرق تخفيف الوزن
5. كيف أحافظ على وزني
6. التخلص من الكرش

الصنف الرئيسي: تجارة ومشاريع

---

1. البنوك والقروض
2. التداول والاستثمار
3. نماذج وخطابات
4. التجارة الإلكترونية
5. صناعات
6. أفكار مشاريع مربحة
7. أعمال يدوية
8. بنك الأهلي
9. بنوك
10. أفكار مشاريع إنتاجية
11. مشاريع صغيرة مربحة
12. اقتصاد مالي
13. نصائح مالية
14. بنك مصر
15. مشاريع منزلية
16. عملات
17. اقتصاد العالم
18. قصص نجاح
19. مال وأعمال
20. بنك cib
21. أفكار وجدوى المشاريع

الصنف الرئيسي: فن الطهي

---

1. منتجات غذائية
2. عجائن ومخبوزات
3. حليب وأجبان
4. حلو عربي
5. أكلات خفيفة
6. أطباق رئيسية
7. أطباق جانبية
8. أساسيات فن الطهي
9. حلويات ومعجنات
10. أطباق الدجاج
11. مشروبات باردة و ساخنة
12. تزيين الأطباق
13. أطباق اللحوم
14. كيك
15. أطباق الأرز
16. شوربات
17. مشروبات وعصائر
18. صلصات
19. مأكولات بحرية
20. حلويات القطر
21. حلويات باردة
22. عصائر
23. مقبلات و سلطات
24. طبخ عربي
25. أطباق شرقية
26. حلويات سريعة
27. المخللات
28. أطباق خليجية
29. أكلات متنوعة
30. سلطات
31. حلويات البسكويت والتارت
32. أطباق مصرية
33. أطباق شامية
34. حلو عالمي
35. المتبلات
36. المربى والدبس
37. فطائر
38. حلويات شرقية
39. أطباق مشوية
40. أطباق صحية
41. أطباق عراقية
42. أكلات سريعة
43. أطباق المغرب العربي
44. المحاشي
45. حلويات رمضان
46. المكرونة
47. الكعك والمعمول
48. حلو الكاسات
49. صواني بالفرن
50. البهارات
51. الكنافة

الصنف الرئيسي: حكم وأقوال

---

1. عبارات جميلة
2. كلمات متنوعة
3. أقوال
4. كلمات معبرة
5. حكم
6. حكم في الحب
7. عبارات عن الأسرة
8. حكم وأقوال في الصداقة
9. حكم متنوعة
10. عبارات حزينة
11. خواطر
12. حكم وأقوال في الحياة
13. حكم المشاهير
14. عبارات تهاني
15. رسائل
16. حكم ومواعظ دينية
17. حكم عن الحياة
18. حكم واقوال العظماء والفلاسفة
19. حكم عن الحزن
20. حكم عن الأخلاق والصفات والعادات
21. حكم قصيرة ومفيدة
22. حكم عن الحب
23. عبارات تهنئة
24. بوستات
25. حكم عن الرجال والنساء
26. حكم عن الصداقة
27. عبارات الصباح
28. كلمات الصباح
29. رسائل قصيرة وجديدة
30. كلمات عتاب
31. حكم وأمثال مضحكة
32. عبارات عتاب
33. كلمات المساء

الصنف الرئيسي: رياضة

---

1. كرة القدم
2. تمارين رياضية
3. رياضات منوعة
4. كمال الأجسام
5. رياضات أخرى
6. مهارات الصيد

الصنف الرئيسي: ثقافة عامة

---

1. ظواهر طبيعية
2. اسئلة دينية
3. كائنات حية
4. اسئلة علمية
5. اسئلة تقنية
6. اكتشافات واختراعات
7. وحدات قياس
8. منوعات علمية
9. اسئلة تاريخية
10. مصالح وخدمات مصر
11. مصالح وخدمات السعودية
12. أسماء أولاد
13. أسماء بنات
14. أسماء فيس بوك
15. آثار عربية
16. آثار مصر
17. آثار عالمية
18. قوانين وأنظمة دولية

الصنف الرئيسي: قصص وحكايات

---

1. الصحابة والتابعون
2. أخلاق إسلامية
3. قصص إسلامية
4. صحابيات
5. قصص أطفال
6. قصص وعبر
7. روايات
8. قصص عربية
9. جغرافيا و تاريخ
10. قصص دينية
11. قصص نجاح
12. قصص عالمية

52. لقم صغيرة
53. أكلات رمضانية
54. الكيك الإسفنجي
55. الجلي والكريم كراميل
56. البان كيك والكريب
57. الشوفان
58. سندويشات

الصنف الرئيسي: سياحة وسفر .

---

1. السياحة في الدول الأجنبية
2. السياحة في الدول العربية
3. السياحة في اوروبا
4. سياحة حول العالم
5. السياحة في مصر
6. السياحة في تركيا

الصنف الرئيسي: تفسير الاحلام والرؤى .

---

1. الزواج والطلاق والجماع في المنام
2. الحروب والاسلحة في المنام
3. الحزن والخوف في المنام
4. الموت والمقابر والجنائز في المنام
5. الذهب والفضة والمعادن في المنام
6. أشياء غريبة في المنام
7. الحمل والرضاعة في المنام
8. البحار والأنهار والوديان في المنام
9. الشياطين والسحر في المنام
10. الأرقام في المنام
11. منوعات الأحلام
12. النقود في المنام
13. الطعام والشراب والموائد في المنام
14. العبادات والذبائح في المنام
15. الطبيعة وما فيها في المنام
16. الملابس والأزياء في المنام
17. الرجال والنساء والأطفال في المنام
18. الحشرات والزواحف في المنام
19. الآفات والأمراض والأدوية في المنام
20. الأسماك وصيدها في المنام
21. أثاث المنزل والغرف في المنام
22. الأسنان والأضراس في المنام
23. الجمال والزينة في المنام
24. المدارس والجامعات في المنام
25. الطيور وصيدها في المنام
26. أعضاء الجسم والشعر في المنام
27. الحيوانات وصيدها في المنام
28. وسائل التنقل والسفر في المنام
29. المرحاض والبراز والبول في المنام
30. الهاتف في المنام
31. الألوان في المنام
32. الفرحة والسعادة في المنام
33. قاموس تفسير الأحلام بالحروف

الصنف الرئيسي: أدعية دينية

---

1. أدعية دينية جميلة مكتوبة
2. أدعية الرزق
3. أدعية الزواج
4. أدعية الشفاء
5. أدعية المظلوم
6. أدعية للمتوفي
7. أدعية دينية قصيرة للأحباب
8. أدعية الولادة
9. أدعية الفرج
10. أدعية الصلاة
11. أدعية التحصين
12. أدعية يوم الجمعة
13. أدعية التيسير
14. أدعية الصباح
15. أدعية الحزن
16. أدعية الإمتحانات
17. أدعية المساء
18. أدعية الحمد والشكر
19. أدعية شعبان
20. أدعية التوفيق
21. أدعية الفجر
22. أدعية التوبة

الصنف الرئيسي: أبراج

---

1. برج العذراء
2. برج الجوزاء
3. برج الدلو
4. برج الحوت
5. برج الجدي
6. منوعات الأبراج
7. برج العقرب
8. برج الثور
9. برج القوس
10. برج السرطان
11. برج الحمل
12. برج الميزان
13. برج الأسد

الصنف الرئيسي: السيارات

---

1. أسعار السيارات