Journal of Innovations in Computer Science and Engineering
Shahid Beheshti Unversity

June 2024, Special Issue on AI 4 All- 1

Title Generation for the Qur'anic chapters by summarizing them

Masoume Maleki✉, Code Orcid: 0009-0003-3719-2480
Interdisciplinary Quranic Studies Research Institute, Shahid Beheshti University, Iran, mas.maleki@mail.sbu.ac.ir
Alireza Talebpour, Code Orcid: 0000-0003-2538-3928
Associate Professor of the Faculty of Computer Science and Engineering and the Interdisciplinary Qur'anic Studies Research Institute, Shahid Beheshti University, Iran, Talebpour@sbu.ac.ir
Mostafa Moradi, Code Orcid: 0000-0002-3609-7276
Assistant Professor of Interdisciplinary Qur'anic Studies Research Institute, Shahid Beheshti University, Iran, m_moradi@sbu.ac.ir

Abstract— With the increase in textual data generated on the internet and the limited time individuals have for reading, the need for automatic text summarization is more essential than ever. One application of summarization is title generation. The goal of this study, which falls within the field of digital humanities and interdisciplinary studies, is to provide a framework for title generation through extractive and abstractive summarization methods, focusing specifically on chapters of the Qur'an. For extractive summarization, eleven different methods have been examined, some of which are novel and innovative. For the abstractive part and title generation, several models have been trained to select the most effective one. In this research, the Persian translation of the Qur'an is used as the primary source, and a dataset was created based on the first ten parts (juz) of the Qur'an, including extractive summaries, abstractive summaries, and titles for various sections of the chapters. The results of this study indicate that the titles generated through summarization are close to human-generated titles, based on BERTScore, R-1, R-2, and R-l values of 21.03, 6.85, 20.73, and 52.51, respectively. It is important to note in the evaluation that a single fixed title does not exist for a document; multiple titles may also be valid. In human evaluation, we observed that the average score produced by the proposed approach is 0.59, while for the best results from other approaches, this value is 0.44.

Keywords— extractive summarization, abstractive summarization, title generation for Qur'anic surahs, computational Qur'anic studies

I. Introduction

With the expansion of digital technologies and the increase in textual data, the need for text processing tools has grown. Natural Language Processing (NLP) is a crucial research field that, with the advancement of artificial intelligence, has contributed to data analysis, content generation, and text summarization. One of the main challenges in this field is summarizing long and complex texts.

The Qur'an, as a comprehensive religious text, encompasses a wide range of concepts and teachings on various subjects. This divine book, considered a miracle of the Prophet Muhammad (PBUH) and the only holy book preserved from alteration, has always been introduced by God as a source for contemplation and reflection. However, the diversity and distribution of topics in the Qur'an may make it challenging for some individuals, especially those less familiar with religious concepts, to study and fully understand.

The primary aim of this research is to design and develop an intelligent and efficient framework for summarizing and generating relevant titles for chapters (surahs) of the Qur'an. This framework, using advanced NLP and AI techniques, will be capable of summarizing the Persian translation of the Qur'an both extractively and abstractively and, based on these summaries, generate one or more titles for each surah.

The goal of extractive summarization is to create a summary by selecting a subset of sentences from the input text and connecting them in a way that maximizes coverage of essential content while avoiding redundancy. In contrast, abstractive summarization aims to create a compact representation of the input text and uses natural language generation techniques to produce a summary. Compared to extractive summaries, generating abstractive summaries is more challenging but closer to human-like summaries, as they may include phrases not present in the original text [1].

II. Related Work

Many studies related to these tasks have been conducted, and some of them will be highlighted below. However, the number of studies on summarization and title generation in Persian is less than in languages such as English.

The background of previous research is generally divided into three categories: extractive summarization, abstractive summarization, and title generation. Additionally, some studies have specifically examined these methods in the context of the Qur'an text.

A. Extractive Summarization

The first research on extractive summarization dates back to 1958. In [2], Luhn used keyword repetition and importance to

Submit Date: 2024-10-15
Accept Date: 2025-04-21
✉ Corresponding author

extract relevant sentences from the main text and presented them as a summary. Afterward, most articles adopted the approach of scoring sentences based on predefined features to select summary sentences. Other extractive summarization methods include clustering techniques [3] and graph-based approaches [4]. More modern approaches utilize machine learning, deep learning, and neural networks [5]. Additionally, in article [6], the BERTSUM model is introduced, which is a simplified version of BERT and focuses on extractive summarization.

In Persian, many studies have also been conducted in this field. PARSUMIST is a Persian summarization system that combines classic, statistical, and semantic methods for single- and multi-document texts [7]. Additionally, article [8] introduces a novel application of OIE as an intermediate layer for summarization, where OIE's structured propositions can be used to shorten sentences and create summaries.

B. Abstractive Summarization

In this type of summarization, the goal is to generate new sentences that convey the main idea of the text, rather than simply selecting existing sentences from the text. Advanced models such as recurrent neural networks and attention-based models are used for this task. In [9], a hybrid approach is discussed that combines extractive and abstractive methods. A simple extraction step is performed before generating the summary, which is then used to condition a transformer-based language model on the relevant information before generating the summary. The paper [10] in 2021 introduces a new dataset called pn-summary, which is related to this task in Persian. Then, the models mt5 and a transformer-based encoder-decoder version of ParsBERT are fine-tuned on the introduced dataset.

In [11], a new self-supervised pretraining objective for large transformer-based encoder-decoder models is proposed, which is trained on massive text corpora. In PEGASUS, important sentences from an input document are masked, and the output continuation is generated from the remaining sentences, similar to an extractive summary. Aiming to focus on the semantic similarity of sentences, the ARMAN model, with the same architecture as PEGASUS, is introduced for the Persian language[12]. ARMAN is a transformer-based encoder-decoder model designed with three new objectives. In ARMAN, key sentences from a document are selected using a modified semantic scoring method and masked to create a pseudo-summary.

C. Title Generation

Title generation is considered an abstractive summarization task with a higher level of compression. Title generation models typically use sequence-to-sequence methods, leveraging neural networks such as LSTMs and attention-based models. In article [13], published in 2017, a top-down approach is proposed: it first identifies important sentences using summarization techniques and then utilizes a multi-sentence summarization model with hierarchical attention to generate a title based on these key sentences. The SHEG method in article [14] is a hybrid model that combines extractive and abstractive summaries to produce a concise summary for title generation. To explore how humans can best utilize large language models for writing, article [15] in 2023 compares various types of common human-AI interactions (such as guidance systems, system output selection, and post-editing outputs) in the context of news headline generation using large language models. Research on automatic title generation in Persian is very limited. Elmnet is a dataset for title generation, containing around 400,000 abstract/title pairs from scientific articles. Article [16] evaluates the performance of the most important title generation methods based on the introduced dataset.

D. Summarization and Title Generation in the Qur'an

According to studies, few articles have addressed summarization or title generation in the Qur'an. These articles typically use topic models to generate titles [17][18][19]. Regarding intelligent summarization of Qur'anic chapters, only one thesis [20] was found, aiming to create an automatic extractive summary using RBM (Restricted Boltzmann Machine) on a dataset containing an English interpretation of Surah Al-Fatihah.

None of the studies have used summarization for title generation of Qur'anic chapters, nor have any of them worked on the Persian translation of the Qur'an.

III. Methodology

In the proposed framework of this research, as shown in Figure 1, an extractive summary is first generated, followed by an abstractive summary, and then a title is produced. The summarization steps in this framework establish the main concept of the document, and the title generation step returns the output as a title-like phrase. In each of these stages, various methods were applied and evaluated to ultimately select the best one for title generation.

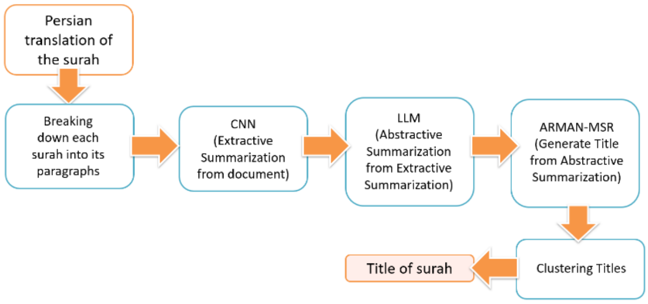

Fig. 1. Proposed framework for title generation

A. Dataset Construction

Due to the lack of a suitable and relevant dataset for this research, a dataset was initially created manually based on the Persian translation of the first ten parts (juz) of the Qur'an specifically using the translation of Makarem Shirazi. To build this dataset, consecutive related verses were grouped together in a single record. The relationships between verses were determined using the meaning and, at times, the interpretation of the verses, as well as expert opinions in this field. After this step, the verses were segmented into sentences based on a pre-existing dataset. Each document thus consists of a paragraph of related verses that have been segmented into sentences. The number of sentences in each document is also specified in the dataset.

For each document in this dataset, an extractive summary, an abstractive summary, and a title were created, each based on the previous one. This means that the extractive summary consists of 20% of the document's sentences, the abstractive summary is based on the extractive summary, and the title is derived from the abstractive summary. This dataset was created for all verses of the Qur'an up to the end of Surah At-Tawbah and will be used in different stages of the research to train and evaluate the models.

B. Extractive Summarization methods

To select the best extractive summarization method, eleven methods were reviewed. In these methods, attempts were made to implement and compare ideas proposed in various studies.

Preprocessing: For each method, the input texts were first preprocessed. The Hazm library was used for normalization and sentence segmentation, and stop words were removed as part of the preprocessing.

The results of these implementations are presented in Table I As can be seen, the CNN model yields the best outcome. This model is a Convolutional Neural Network (CNN). For training, the first 400 samples of the dataset were used, while the last 86 samples were reserved for evaluation. The maximum sentence count per document in this dataset is 58, which determined the input matrix dimensions for this model (58, 768) to accommodate all documents.

For each input document, the representation obtained from ParsBERT for each sentence is placed in a row of the input matrix. If a document has fewer than 58 sentences, the remaining rows are padded with zeros. Each document's labels are represented as a vector of length 58, where each entry is set to zero or one, depending on whether the sentence is included in the summary.

The Mean Squared Error (MSE) loss function is used, representing the mean of the squared differences between actual and predicted values. The model was trained for 18 epochs with a batch size of 32. Finally, sentences corresponding to the elements with the highest values in the output vector are selected.

C. Abstractive Summarization methods

Abstractive summarization in Persian has received far less attention compared to extractive summarization. To perform this task, after reviewing various models, we fine-tuned a pre-trained language model called Arman, introduced in 2021 [12], which was trained on summarization datasets like pn-summary, on our Qur'anic dataset. The input to this model is the extractive summary, and the output is the abstractive summary. This model was also trained on the first 400 samples of the dataset.

D. Ttile Generation methods

For title generation, we also used the Arman model, similar to abstractive summarization. This model was likewise trained on the first 400 samples of the dataset.

E. Large Language Models

The development of large language models and their impressive results in various natural language processing tasks in recent years has attracted significant attention from researchers. Therefore, in this study, we have also utilized these models for generating abstractive summaries. The command-r-plus model is one of the large language models that performs very well on Persian text. In this context, using RAG technology and prompt engineering, efforts were made to generate suitable abstractive summaries from extractive summaries. The results of this model are compared with those of the trained abstractive model.

F. Determining Titles for surahs

The goal of this research is to generate titles for the surahs. As explained so far, we have been able to find a title for each section of the surah. To aggregate these titles and create a title for the surah, the representations obtained from each title were clustered using the k-means algorithm. The number of clusters in this step is defined as 1/20 of the number of titles, and ultimately, the closest sample to the center of each cluster is selected.

Finally, these titles will be evaluated manually.

Results

All extractive, abstractive, and title generation methods have been evaluated based on automatic metrics. Additionally, the impact of using large language models on improving the results will be examined. The titles generated have also been assessed by human evaluators. For automatic evaluation, two metrics, ROUGE and BERTScore, have been used. It is worth mentioning that, as noted in the literature review, no similar work based on artificial intelligence techniques was found for text summarization of the Quran or for generating Surah titles in Persian or even in other languages. Therefore, the proposed approach in this study cannot be directly compared to any existing method. However, the models used in each step have been compared with several other models trained under different settings, and the best-performing model has been selected. For brevity, only the results of the best model are presented.

A. **Extractive Summarization**

In Table 1, the average values of Rouge-1, Rouge-2, Rouge-L, and BERTScore in the F-mode are reported, sorted in descending order based on the average.

Method for calculating the average: In each document, the number of sentences correctly identified by the model is divided by the total number of sentences in the document, and then the average is taken across all documents. This is interpreted as the average number of sentences each method can correctly extract.

As shown in the Table 1, the best method for extractive summarization is the CNN model, which will be 73% successful in creating summaries. Additionally, based on the Rouge metric, which relies on lexical similarity, it scores over 70. The BERTScore metric, which compares summaries based on semantic similarity, shows much better results with scores around 86, indicating that the selected sentences may not exactly match the reference summary sentences but are semantically very close to them. Therefore, this model will be used for extractive summarization.

TABLE I. EXTRACTIVE SUMMARIZATION METHODS

Without Remove stop-words					method
BERT Score	R-L	R-2	R-1	average	
0.86	0.739	0.723	0.757	0.733	CNN
0.881	0.768	0.742	0.786	0.7	Choose the first n sentences
0.827	0.674	0.635	0.696	0.606	n-element combinations with the highest ROUGE score relative to the entire document.
0.825	0.672	0.635	0.698	0.599	Calculate the TF-IDF vector of sentences, sum the values for each sentence, and select the highest ones
0.761	0.516	0.52	0.607	0.517	Calculate the TF-IDF vectors of the sentences, obtain the cosine similarity between vectors, compute the average of the similarities, and select the sentence with the highest similarity to the other sentences
0.762	0.546	0.502	0.593	0.502	TextRank
0.749	0.526	0.479	0.571	0.478	Compare each sentence vector with the average vector excluding the sentence itself
0.739	0.504	0.451	0.553	0.461	Compare each sentence vector with the average vector of the entire document
0.74	0.505	0.453	0.553	0.427	Cluster the sentences and select the centroid of each cluster
0.735	0.495	0.453	0.557	0.374	Cluster the sentences and select the best sentence from each cluster using TF-IDF and summation
0.728	0.489	0.439	0.544	0.373	Cluster the sentences and select the best sentence from each cluster using TF-IDF with the highest similarity

B. *Abstractive Summarization*

To perform abstractive summarization, the ARMAN model has been fine-tuned on the constructed dataset.

The components and settings of the model during the training process are defined as follows: batch size of 4 and 6 epochs are considered, the learning rate is set to 1e-4, and a weight decay of 0.01 is used to prevent overfitting.

The input to the abstractive summarization models is the extractive summaries from the Qur'anic dataset, which need to be converted into titles of the dataset.

The results of the summarization are presented in Table II.

In this approach, the model attempts to convert extractive summaries into abstractive summaries, producing a single sentence. This approach is denoted as "Ext-Abs" in Table II, which stands for "extractive-abstractive." After evaluating various models, the ARMAN-SH model achieved the highest Rouge and BERTScore values compared to other models.

TABLE II. PROPOSED FRAMEWORK EVALUATION

BERT Score	R-L	R-2	R-1	model	approach
0.515	0.23	0.095	0.271	ARMAN-SH-Persian-base-PN-summary	Ext-Abs
0.706	0.548	0.348	0.559	ARMAN-MSR-Persian-base-PN-summary	Abs-Title
0.568	0.257	0.116	0.295	ARMAN-SH	Ext-Abs (framework)
0.512	0.189	0.068	0.201	ARMAN-MSR	Abs-Title (framework)
0.552	0.22	0.073	0.245	cohere	Ext-Abs (LLM)
0.525	0.207	0.068	0.21	cohere	Abs-Title (LLM-framework)

C. *Title Generation*

After testing various models for title generation, the results of the best model can be seen in Table II, which is represented by "abs-title." The ARMAN-MSR model, with R-1, R-2, and R-L, BERTScore values of 0.559, 0.348, 0.548, and 0.706, respectively, was selected as the best model for title generation.

After obtaining the best models for extractive summarization, abstractive summarization, and title generation, we placed them within the proposed framework. In this framework, the output of each model was used as input for the next model.

Initially, the main document is fed into the CNN model, and after obtaining the output, it is used as the input for the abstractive model (ARMAN-SH), and its output becomes the input for the title generation model (ARMAN-MSR).

After applying the models in the proposed framework, the results shown in Table II were obtained.

It is important to note that the errors from each stage inevitably pass on to the next stage, which will undoubtedly affect the quality of the results. In this research, we examine whether the proposed framework can still produce satisfactory results despite these errors.

D. *Large Language Models*

In the next step, we attempted to improve these results using large language models. The results are presented in Table II. After evaluating all the different scenarios, we found that the use of large language models could only slightly improve the automatic evaluation results from the previous section in the phase of converting extractive summaries to abstractive ones. Therefore, in our final framework, we will use LLMs for this phase.

E. *Human Evaluation*

In this section, the titles generated for the last 87 records of the dataset were evaluated by three experts. To perform this, a Qur'anic expert assigned values of 0, 0.5, or 1 to each generated title for each document. The comparison criterion was not the human title for the document but the appropriateness of the generated title for the document. If the title had correct linguistic structure and was appropriate for the document, it received a score of 1. If the title was correct but had minor linguistic issues or was ambiguous, it received a score of 0.5, and in all other cases, it received a score of 0. Finally, the scores were summed and divided by 87.

According to the proposed approach, this score is 0.59. Additionally, only 27 titles out of the 87 titles received a score of 0 in the proposed method. The results of the human evaluation indicate that the proposed method in this research outperformed other methods. Samples of the outputs of this method can be seen in Table III.

TABLE III. RESULTS OF THE PROPOSED FRAMEWORK

title (proposed approach)	document
نادیده گرفتن دلایل روشن پیامبران توسط قوم های پیشین	آیا خبر کسانی که پیش از آنها بودند، به آنان نرسیده است؟! «قوم نوح» و «عاد» و «ثمود» و «قوم ابراهیم» و «اصحاب مدین» [= قوم شعیب] و «شهرهای زیر و رو شده» [= قوم لوط]؛ پیامبرانشان دلایل روشن برای آنان آوردند، (ولی نپذیرفتند؛) خداوند به آنها ستم نکرد، امّا خودشان بر خویشتن ستم می‌کردند!
سوگند خوردن به دروغ	آنها به خدا سوگند می‌خورند که از شما هستند، در حالی که از شما نیستند؛ ولی آنها گروهی هستند که می‌ترسند (و به خاطر ترس از فاش شدن اسرارشان دروغ می‌گویند)! اگر پناهگاه یا غارها یا راهی در زیر زمین بیابند، بسوی آن حرکت می‌کنند، و با سرعت و شتاب فرار می‌کنند.
پذیرش صلح	و اگر تمایل به صلح نشان دهند، تو نیز از در صلح درآی؛ و بر خدا توکّل کن، که او شنوا و داناست! و اگر بخواهند تو را فریب دهند، خدا برای تو کافی است؛ او همان کسی است که تو را، با یاری خود و مؤمنان، تقویت کرد.

F. Title Clustering

Finally, to obtain the title of the Surah, the titles generated for each section of the Surah are clustered, and the title closest to the center of the cluster is selected. The number of clusters is determined by dividing the number of sections by 20, meaning that for every 20 sections, one title is generated. The reason for this is that long Surahs of the Qur'an cover multiple topics, making it difficult to select a single title for them. Therefore, multiple titles are generated for them. A few of the Surah titles can be found in Table IV. The score column represents the average of the ratings assigned to these titles by three experts. For scoring, each expert, based on their knowledge of the Surah and its overall meaning, with a deep understanding of its interpretation and considering the titles proposed by commentators, assigned a score between 0 and 10.

TABLE IV. TITLES OF SURAHS BASED PROPSED FRAMEWORK

Score(10)	title	surah
5	پذیرش صلح نهی از فرار کردن در نبرد با کافران	انفال two titles
6	جنگ با کافران و منافقان نهی از یار شدن با کافران	توبه two titles
9	پاداش بزرگ ترس از پروردگار	ملک one title
7	خلق انسان در بهترین شکل	تین one title
9	خداوند، مقصود نیازمندان	اخلاص one title
8	رخ دادن حادثه ای ویرانگر	قارعه one title

Conclusion

The aim of this research was to investigate whether it is possible to extract the purpose or title of a Surah using summarization techniques. In this research, a framework was proposed to generate titles for Surahs. In this framework, the first step is to create an extractive summary of the document, which plays a key role in identifying important sentences. Then, the extractive summary is transformed into an abstract summary in one sentence to make it coherent. Finally, this sentence is converted into a phrase, which becomes the title.

However, this method is not directly applied to the text of the Surah; instead, each Surah is divided into different sections, and a title is generated for each section. This research was conducted on the Persian translation of the Qur'an.

For training the models, a Qur'anic dataset was manually created with the guidance of a Qur'anic expert, which included the segmentation of the Qur'an into up to ten parts (juz), along with extractive summaries, abstract summaries, and titles for each section.

In this study, eleven methods were evaluated for extractive summarization, and their results were compared. Afterward, the best

method, which was the CNN model, was chosen. For abstract summarization and title generation, various models were fine-tuned on the constructed dataset, and the best-performing ones were selected.
In the next step, after generating titles for the different sections of the Surah, the Surah titles were obtained using clustering and selecting the closest sample to the center of the cluster. According to the Qur'anic expert, these titles effectively convey the overall meaning of the Surah to the audience.

References

[1] Lin H, Ng V. Abstractive summarization: A survey of the state of the art. InProceedings of the AAAI conference on artificial intelligence 2019 Jul 17 (Vol. 33, No. 01, pp. 9815-9822).
[2] Luhn HP. The automatic creation of literature abstracts. IBM Journal of research and development. 1958 Apr;2(2):159-65.
[۳] شفیعی ف، شمس فرد م. سیستم خودکار خلاصه ساز متون فارسی. ۱۳۹۳.
[٤] پورغلامعلی ف. خلاصه‌سازی خودکار متن با استفاده از تکنولوژی داده‌های عظیم. در: سومین کنفرانس ملی در مهندسی کامپیوتر، فناوری اطلاعات و پردازش داده‌ها؛ ۱۳۹۷.
[5] Laugier L, Thompson E, Vlissidis A. Extractive Document Summarization Using Convolutional Neural Networks-Reimplementation. Department of Electrical Engineering and Computer Sciences University of California, Berkeley. 2018.
[6] Liu Y. Fine-tune BERT for extractive summarization. arXiv preprint arXiv:1903.10318. 2019 Mar 25.
[7] Shamsfard M, Akhavan T, Joorabchi ME. Persian document summarization by parsumist. World Applied Sciences Journal. 2009;7:199-205.
[8] Rahat M, Talebpour A. Open information extraction as an intermediate semantic structure for Persian text summarization. International Journal on Digital Libraries. 2018 Nov;19(4):339-52.
[9] Pilault J, Li R, Subramanian S, Pal C. On extractive and abstractive neural document summarization with transformer language models. InProceedings of the 2020 conference on empirical methods in natural language processing (EMNLP) 2020 Nov (pp. 9308-9319).
[10] Farahani M, Gharachorloo M, Manthouri M. Leveraging ParsBERT and pretrained mT5 for Persian abstractive text summarization. In2021 26th International computer conference, computer society of Iran (CSICC) 2021 Mar 3 (pp. 1-6). IEEE.
[11] Zhang J, Zhao Y, Saleh M, Liu P. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. InInternational conference on machine learning 2020 Nov 21 (pp. 11328-11339). PMLR.
[12] Salemi AK, Emad; Neisi Minaei, Ghazal; Shakery, Azadeh, editor ARMAN: Pre-training with Semantically Selecting and Reordering of Sentences for Persian Abstractive Summarization. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing; 2021. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics.
[13] Tan J, Wan X, Xiao J. From Neural Sentence Summarization to Headline Generation: A Coarse-to-Fine Approach. InIJCAI 2017 Aug 19 (Vol. 17, pp. 4109-4115).
[14] Singh RK, Khetarpaul S, Gorantla R, Allada SG. SHEG: summarization and headline generation of news articles using deep learning. Neural Computing and Applications. 2021 Apr;33:3251-65.
[15] Ding Z, Smith-Renner A, Zhang W, Tetreault JR, Jaimes A. Harnessing the power of LLMs: Evaluating human-AI text co-creation through the lens of news headline generation. arXiv preprint arXiv:2310.10706. 2023 Oct 16.
[16] Shenassa ME, Minaei-Bidgoli B. ElmNet: a benchmark dataset for generating headlines from Persian papers. Multimedia Tools and Applications. 2022 Jan 1:1-4.
[17] Siddiqui MA, Faraz SM, Sattar SA. Discovering the thematic structure of the Quran using probabilistic topic model. In2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences 2013 Dec 22 (pp. 234-239). IEEE.
[18] Rolliawati D, Rozas I, Khalid K. Text Mining Approach for Topic Modeling of Corpus Al Qur'an in Indonesian Translation. InProceedings of the 2nd International Conference on Quran and Hadith Studies Information Technology and Media in Conjunction with the 1st International Conference on Islam, Science and Technology, ICONQUHAS & ICONIST, Bandung, October 2-4, 2018, Indonesia 2020 May 25.
[19] Alshammeri M, Atwell E, Alsalka MA. Quranic topic modelling using paragraph vectors. InIntelligent Systems and Applications: Proceedings of the 2020 Intelligent Systems Conference (IntelliSys) Volume 2 2021 (pp. 218-230). Springer International Publishing.
[20] Hummeida NS. Text summarization of Holy Quran interpretation in English using deep learning algorithm (Case study: Surat Al-Fatiha) [Master's thesis]. Sudan University of Science and Technology, College of Graduate Studies, College of Computer Science and Information Technology; 2018

Masoume Maleki
0009-0003-3719-2480
Alireza Talebpour
0000-0003-2538-3928
Mostafa Moradi
0000-0002-3609-7276
Title Generation for the Qur'anic chapters by summarizing them