

Incorporating Deep Median Networks for Arabic Document Retrieval Using Word Embeddings-Based Query Expansion

Yasir Hadi Farhan* 

Department of Medical Physics, College of Applied Sciences, University of Fallujah, Fallujah, Iraq
E-mail: yasir.hadi87@yahoo.com

Mustafa Abd Tareq 

Department of Computer Science, University of Technology-Iraq, Baghdad, Iraq
E-mail: mustafa.t.abd@uotechnology.edu.iq

Mohanaad Shakir 

Department of Management Information System (MIS), College of Business (COB), University of Buraimi (UOB), Buraimi, Oman
E-mail: mohanaad.t@uob.edu.om

Boumedyen Shannaq 

Department of Management Information System (MIS), College of Business (COB), University of Buraimi (UOB), Buraimi, Oman
E-mail: boumedyen@uob.edu.om


ABSTRACT

The information retrieval (IR) process often encounters a challenge known as query-document vocabulary mismatch, where user queries do not align with document content, impacting search effectiveness. Automatic query expansion (AQE) techniques aim to mitigate this issue by augmenting user queries with related terms or synonyms. Word embedding, particularly Word2Vec, has gained prominence for AQE due to its ability to represent words as real-number vectors. However, AQE methods typically expand individual query terms, potentially leading to query drift if not carefully selected. To address this, researchers propose utilizing median vectors derived from deep median networks to capture query similarity comprehensively. Integrating median vectors into candidate term generation and combining them with the BM25 probabilistic model and two IR strategies (EQE1 and V2Q) yields promising results, outperforming baseline methods in experimental settings.

Keywords: automatic query expansion, information retrieval, word embedding, deep median networks, Arabic document retrieval, natural language processing

Received: October 10, 2023
Accepted: May 9, 2024

Revised: April 10, 2024
Published: September 30, 2024

*Corresponding Author: Yasir Hadi Farhan
 <https://orcid.org/0000-0002-2378-9177>
E-mail: yasir.hadi87@yahoo.com



All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

The main challenge for search engines is that user queries are often short and not specific enough to accurately represent their information needs (Azad & Deepak, 2019; Farhan et al., 2020; Nwesri & Alyagoubi, 2015). This is due to a gap in user knowledge, according to the Anomalous State of Knowledge hypothesis (Belkin, 2005). The automatic query expansion (AQE) technique addresses this issue by expanding the primary query with related terms to generate better and more relevant results (Esposito et al., 2020; Farhan et al., 2020; Raza et al., 2019). Commercial search engines such as Yahoo and Google use this technique through linked queries, connected search terms, and auto-completion functionality (Cai & De Rijke, 2016).

1.1. Query Expansion Techniques

Query expansion techniques are classified into global and local methods. Global techniques use a thesaurus such as WordNet to enlarge the initial user queries without relying on retrieval outcomes (Farhan et al., 2020; Pal et al., 2014). Local techniques use relevance feedback from the first retrieval process to select appropriate terms to add to the primary query (Miyaniishi et al., 2013; Takeuchi et al., 2017). The pseudo relevance feedback (PRF) technique is a useful expansion method that automates manual aspects of relevance feedback. It assumes that the top-k results retrieved in the original search contain words that can be used to refine the query further (Farhan et al., 2021b).

The PRF technique is effective, but it has practical issues. Terms generated from resources like WordNet tend to have multiple meanings, requiring a disambiguation strategy before using them for query expansion. Additionally, the PRF approach relies on the highest ranked documents retrieved, which can be impacted by irrelevant terms and multiple topics (Zou et al., 2018), making it challenging to improve both precision and recall metrics simultaneously (Fernández-Reyes et al., 2018). To resolve these issues, recent research in AQE has focused on using word embeddings (WE) as a semantic modelling process in order to have meanings from text (ALMasri et al., 2016; Diaz et al., 2016; Roy et al., 2016).

In natural language processing (NLP), WE use a 1-D mathematical embedding to represent every word as a vector of real numbers in a low-dimensional continuous vector space. The models use the proximity of terms in their corpus for training (Diaz et al., 2016). Word2Vec has two options for training WE models: Continuous Bag-Of-Words (CBOW) and Skip-Gram (SG). CBOW predicts

aim words based on nearby words, while SG predicts surrounding words based on target words. These models can determine terms with syntactic and semantic similarities by using the same context. Most AQE techniques use query term search or neighbourhood strategies to expand the query using proximal N terms (Roy et al., 2016).

Current AQE techniques using WE typically retrieve candidate terms one at a time, without considering the influence of other terms in the query. Researchers argue that AQE can be improved by modelling query semantics as collective vocabulary terms, resulting in higher quality suggested terms that are more semantically relevant. To accomplish this, they suggest employing deep averaging networks (DANs), a neural architecture that computes the average of embedded words for classification and processes them through several linear layers. The researchers suggest that DANs can be used to identify related terms for AQE using the complete query input (Roy et al., 2016), but this approach has not yet been extensively studied.

The use of WE in information retrieval (IR) has been extensively studied, but its usage in Arabic IR has not been properly investigated (Alsmearat et al., 2014; Faqeeh et al., 2014), mainly due to the lack of ontological knowledge bases in Arabic (Mohsen et al., 2018). Previous studies on Arabic IR have mainly focused on assessing or comparing word-stemming methods (Abdelali et al., 2016; Abu El-Khair, 2007; Guirat et al., 2016; Larkey et al., 2002; Mustafa et al., 2008).

1.2. Aims and Objectives

The aim of this research is to address the persistent challenge of query-document vocabulary mismatch within IR systems by enhancing AQE techniques using deep median networks (DMNs). This challenge arises when user queries fail to accurately represent their information needs, leading to suboptimal search results. By leveraging DMNs, the study aims to comprehensively capture query similarity, thus improving the relevance and effectiveness of search engine results.

Furthermore, the objectives of this research encompass several key aspects. Firstly, it seeks to investigate the limitations of current AQE methods, particularly those reliant on WE, in effectively addressing query-document vocabulary mismatch. Secondly, the study aims to explore the potential of DMNs in capturing comprehensive query similarity, thereby overcoming the shortcomings of existing AQE approaches. Thirdly, it endeavours to integrate DMNs into AQE methodologies alongside traditional IR strategies, such as the BM25 probabilistic model and

EQE1 and V2Q models, to assess their combined efficacy in improving retrieval performance. Lastly, the research aims to evaluate the applicability and effectiveness of the proposed approach within the context of Arabic IR, an area that has been relatively underexplored due to the scarcity of ontological knowledge bases.

1.3. Significance of the Study

This study's importance is in its potential to enhance AQE techniques and aid in creating more efficient search engines. By introducing DMNs as a novel approach for capturing query similarity comprehensively, the research aims to overcome the limitations of existing methods and improve retrieval performance. Moreover, the focus on Arabic IR fills a crucial gap in the literature and has implications for improving access to relevant information in Arabic language contexts. Ultimately, findings from this study may inform the development of more efficient and inclusive search engine technologies, particularly in languages with unique linguistic characteristics such as Arabic.

The primary issue addressed in this study is the inherent challenge faced by search engines due to the ambiguity and lack of specificity in user queries. Despite existing techniques like AQE, practical hurdles persist, such as polysemy in ontological knowledge bases and reliance on top-ranked documents, affecting precision and recall. Recent advancements leveraging WE show promise, yet there is a need for improved AQE methodologies, especially in languages like Arabic. This study proposes utilizing DMNs for AQE in Arabic IR, aiming to address term mismatch issues and enhance retrieval performance.

2. RELATED WORKS

IR systems help users locate the information they seek by fetching it from a database in response to their queries (Baeza-Yates & Ribeiro-Neto, 1999). However, one of the main challenges faced by these systems is vocabulary mismatch. To address this issue (Carpineto & Romano, 2012; Farhan et al., 2020), researchers have proposed AQE techniques, which automatically add new terms to the query to improve the accuracy and precision of the IR system (Abbache et al., 2016). One popular approach for AQE is the use of WE, which has gained significant attention (ALMasri et al., 2016; Diaz et al., 2016; El Mahdaouy et al., 2018b; Farhan et al., 2021a; 2021b; Roy et al., 2016).

WE is a technique used in NLP for semantic parsing that helps in extracting meaning from texts. It represents

words as vectors of real numbers in a corpus, which are categorized into local and distributed representations. The distributed representation indicates that all words having similar context display similar vectors within the WE vector space, thus creating a closer distribution. This technique can represent words based on their value vectors and can help in understanding natural language by extracting meaning from texts (Bengio, 2009; Kim et al., 2017; Turney & Pantel, 2010).

Aklouche et al. (2018) developed an AQE technique using Word2Vec toolkit based on WE. The study focused on two models, i.e., SG and CBOW, to learn semantically related words for the main queries from the Text REtrieval Conference (TREC) *Washington Post* Corpus. They reweighted and selected these terms and evaluated the effectiveness of the document retrieval process using the Euclidean distance to compute vector similarity. Two types of candidate vectors were selected, one related to the complete query and the other to individual terms. Their study showed that a query reweighting technique was more effective than other approaches and concluded that assigning the same weight to terms in the expanded query decreases retrieval effectiveness.

Using word vector representations provides an effective foundation for modelling semantic similarity in query expansion (Esposito et al., 2020). However, the context is often neglected during expansion techniques, and formulating a proper context for retrieving useful terms is crucial. Researchers have attempted to tackle this problem, such as Roy et al. (2016), but it has not been completely resolved. Two studies that addressed this problem are those conducted by Fernández-Reyes et al. (2018) and Zamani and Croft (2016). Fernández-Reyes et al. (2018) developed a new Query-Guided AQE strategy known as V2Q, which filtered the primary query and ignored unnecessary terms by considering terms that showed high similarity with those in the primary query.

Zamani and Croft (2016) proposed a new technique called EQE1 for query expansion, which considers the semantic similarity between all terms based on the similarity noted between different WE vectors. It presumes that the query terms are conditionally independent and all expanded query terms must be like the query terms that are selected for adding to the primary query. The experiments conducted in their study used two baseline methods, and the details are discussed in the paper.

Fernández-Reyes et al. (2018) suggested an AQE method that employs Word2Vec (WE) (Mikolov et al., 2013) and took into account the full query to tackle the disam-

biguation problem, enhancing both precision and recall (ALMasri et al., 2016). They proposed two strategies, query-guided association scheme (Q2V) and prospect-guided association scheme (V2Q), which retrieved candidate terms based on query terms and candidate terms that had a high semantic relationship with query terms, respectively. The Q2V approach used general pooling methods to generate Rankq values for every query term and selected top N terms for expansion. However, this approach could lose the context of the primary query. On the other hand, the V2Q approach used candidate expansion terms to vote for query terms and selected the most semantically related terms. Experimental results showed that these techniques improved precision and recall metrics and outperformed conventional IR models.

Zamani and Croft (2016) carried out research aimed at improving the effectiveness of query language models in ad-hoc retrieval tasks by utilizing WE. They proposed a new AQE embedding relevance model based on previous models (Lavrenko & Croft, 2017), and used two estimation processes to incorporate semantic similarity between terms in WE vectors. The techniques that relied on embedding outperformed other baselines in average precision and mean average precision (MAP) on three TREC collections. However, expanding solely on individual query terms can cause query drift, and there is a limitation in the utilization of query terms (Crimp & Trotman, 2018).

Farhan et al. (2021a) proposed a new approach for AQE called DANs. The approach involves using the mean vector of the initial query term vectors to pick potential vectors. They stated that for expanding the whole query sentence, the technique must consider the average vector of the primary query term vectors. DANs was incorporated in the BM25, V2Q, and EQE1 probabilistic models for improving the retrieval performance of Arabic texts. According to the experimental outcomes, the suggested approach enhanced the Arabic text retrieval's performance. It showed a better performance than the standard baseline techniques like BM25, V2Q, and EQE with regards to the precision and MAP values and was seen to be among the top 10 in most of the case studies.

Farhan et al. (2021b) also introduced a new approach to improve the effectiveness of AQE for Arabic text retrieval by utilizing DANs (Farhan et al., 2021b) and the query vectors' average for determining candidate expansion vectors. The study used Word2Vec for training and compared the proposed approach with the Okapi BM25 probabilistic framework, V2Q (Fernández-Reyes et al., 2018), and EQE1 (Zamani & Croft, 2016). The hypoth-

esis was that the DANs-based PRF technique using WE similarity for generating expansion vectors can resolve the issue of term inconsistency and enhance the performance of Arabic text retrieval through AQE. The expansion vectors for potential terms were constructed using the k most relevant documents identified during PRF. Evaluation showed that the proposed approach significantly improved performance compared to baseline PRF frameworks.

Although the DANs technique had advantages, it had limitations such as the average vector not referring to a real vector, being impacted by the position of query term vectors, and showing good performance only with local datasets. To overcome these limitations, researchers proposed a new approach using the median vector instead of the average vector for generating additional candidate vectors. This approach, called DMNs, was not affected by the position of all query term vectors, and can be referred to a real vector.

3. METHOD

This study attempted to determine the probability that DMNs could support AQE for the Arabic IR process. Therefore, the researchers used the DMNs technique for the existing CBOW IR model, i.e., the probabilistic Okapi BM25 model along with the two representative WE techniques for AQE, the V2Q and EQE1 techniques which were initially described by Fernández-Reyes et al. (2018) and Zamani and Croft (2016). The researchers selected the BM25 model, as it showed a true performance in the TREC retrieval experiments and affected ranking algorithms in the commercial search engines (Croft et al., 2010). BM25 was regarded as the main baseline model which could be used for non-AQE results. However, the EQE1, V2Q, and DANs techniques were seen to be the most effective techniques. The proposed technique also compared against the BM25+DANs-PRF, EQE1+DANs-PRF, and V2Q+DANs-PRF techniques introduced by Farhan et al. (2021a; 2021b). The researchers compared the results of the DMNs-based AQE technique with the existing AQE techniques that were presented by these approaches. Thus, they determined whether the DMNs could generate better and increasingly relevant expansion terms and overcome the disadvantages presented by the DANs technique.

3.1. Word2Vec and Deep Median Networks

The Word2Vec deep learning toolkit was introduced

in a previous article (Mikolov et al., 2013) for generating word vectors from a text corpus. It employs CBOW and SG models to generate distributed word representations. DMNs is a new approach that determines the median vector of primary query term vectors to generate candidate expansion vectors, improving automatic IR (AIR) performance and overcoming the limitations of the DANs method. DMNs can be integrated with AQE techniques such as BM25, EQE1, and V2Q models to generate a candidate expansion list. DMNs refers to a sentence embedding process and can be trained faster for data with high syntactic variance, such as Arabic data. The explanations of how Word2Vec and DMNs function are provided below.

Word2Vec and DMNs are both powerful techniques in NLP, but they operate differently and serve distinct purposes.

3.1.1. Word2Vec

- Word2Vec is a popular deep learning toolkit used for generating word vectors from a text corpus.
- It is based on the idea that a word's meaning can be derived from its surrounding context within a vast text corpus.
- Word2Vec provides two primary models: CBOW and SG.
- CBOW: Estimates the target word from its surrounding context words, like completing a missing word based on nearby words.
- SG: Predicts the surrounding context words from a given target word, akin to inferring the words around a specific word.

These models learn to represent words as dense, low-dimensional vectors, where vectors for related words are positioned near each other in the vector space, capturing semantic relationships. Word2Vec's vectors capture syntactic and semantic similarities between words, enabling tasks like word analogy and similarity calculations.

3.1.2. Deep Median Networks (DMNs)

DMNs are a novel approach used to determine the median vector of primary query term vectors to generate candidate expansion vectors. Unlike other techniques that focus on average or summing vectors, DMNs prioritize the median vector, which represents a middle ground among the query terms. This approach is beneficial for capturing the essence of the query while avoiding potential biases introduced by extreme values or outliers. DMNs can be integrated with AQE techniques, such as BM25,

EQE1, and V2Q models, to generate a candidate expansion list that enhances the performance of IR systems. Additionally, DMNs excel in processing data with high syntactic variance, such as Arabic text, making them suitable for diverse linguistic contexts.

In summary, Word2Vec transforms words into dense vectors based on their contextual usage, capturing semantic relationships, while DMNs focus on determining the median vector of query term vectors to improve query expansion and IR performance.

This study explores the effect of using the DMNs approach to expand query sentences in AQE. The approach is applied to three models, namely BM25 IR, EQE1-AQE, and V2Q-AQE. The Word2Vec CBOW technique is trained offline on the Arabic TREC 2001/2002 corpus collection, and the dataset is divided into three groups. The researchers performed a search using the same corpus and evaluated the statistical information regarding the dataset, as shown in Table 1 (Farhan et al., 2021a). The study aims to improve the performance of AQE and overcome the limitations presented by the DANs technique.

Indeed, while the concept of using the median operation in DMNs seems intuitive, it has not been extensively explored in prior research, particularly in the context of query expansion. This gap in existing literature highlights the novelty and contribution of DMNs in introducing this approach to enhance query expansion techniques.

3.2. Basic Query Expansion Based on DMNs (BM25+DMNs)

The Okapi BM25 model is a ranking algorithm employed in IR to assess how relevant documents are to a specific search query (Robertson et al., 1995). It is ground-

Table 1. Statistics for Arabic TREC collections

| Collection | TREC 2001 | TREC 2002 | TREC 2001/2002 |
|------------------------------|------------|-----------|----------------|
| No. of queries | 25 | 50 | 75 |
| Average No. of words/queries | 4.88 | 3.28 | 4.08 |
| No. of documents | 383,872 | | |
| No. of tokens | 76 million | | |
| No. of unique words | 666,094 | | |
| Size (compressed) | 209 MB | | |
| Size (uncompressed) | 869 MB | | |

Data from the article of Farhan et al. (2021a, *Journal of Information Science*, 49(5), 1168-1186). TREC, Text REtrieval Conference.

ed in the probabilistic retrieval model and is an improvement over the earlier Okapi BM model. BM25 takes into account the frequency of query terms in documents, document length, and term frequency within the entire document collection. It adjusts relevance scoring based on these factors, providing more accurate results compared to traditional term frequency-inverse document frequency (TF-IDF) models. BM25 is widely used in search engines and has been shown to perform well across various IR tasks (Lv & Zhai, 2011; Robertson & Zaragoza, 2009). Overall, the Okapi BM25 model offers a robust and effective approach to ranking documents based on their relevance to a given query. Its ability to address the limitations of traditional TF-IDF models makes it a popular choice in modern IR systems (Croft et al., 2010). Equations 1 and 2 present the general BM25 scoring functions.

$$\sum_{i \in Q} \log \frac{\frac{(r_i + 0.5)}{R - r_i + 0.5}}{\frac{(n_i - r_i + 0.5)}{(N - n_i - R + r_i + 0.5)}} \cdot \frac{(k_1 + 1)f_i}{K + f_i} \cdot \frac{(k_2 + 1)qf_i}{k_2 + qf_i} \quad (1)$$

$$K = k_1 \left((1 - b) + b \cdot \frac{dl}{avdl} \right) \quad (2)$$

The formula for calculating relevance in this model involves several parameters, including k_1 , k_2 , and K , which are determined through experimentation. The variable qf represents the frequency of a term in a query, while dl represents the length of a document. Commonly used values for the parameters include $k_1=1.2$, k_2 ranging from 0 to 1,000, and $b=0.75$. Additionally, $avdl$ represents the average document longitude in the set.

It is easy to use AQE-based DMNs technique in the BM25 model, wherein the complete query terms form the input. Using a WE model, vectors were extracted for each query and a mean vector was created using the DMNs method proposed in the study. Thereafter, the vectors that were similar to the median vector were identified after using different cosine similarity techniques, as shown in Equation 3. These similar vectors were regarded as the potential candidate expression vectors, while the candidate expansion terms that were associated with the extracted vectors were also identified with the help of a WE corpus. The highest-ranking 'n' candidates were selected as the terms for expanding the initial query.

$$\text{Cosine}(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}} \quad (3)$$

Equation 4 illustrates the CBOW framework of the Word2Vec model, which is utilized to derive vector representations of terms. In this approach, the model uses the words surrounding a target word to predict its vector representation. The equation incorporates two parameters: $|C|$, which denotes the overall count of words within the corpus, and c , which refers to the dynamic context size of the target word (i.e., the number of words surrounding the target word that are considered in the prediction process).

$$\frac{1}{|C|} \sum_{t=1}^{|C|} \log[\log[P(wt \setminus wt - c, \dots, wt - 1, wt + 1, \dots, wt + c)]] \quad (4)$$

$$m(\vec{X}_k) = \vec{X} \left[\frac{n}{2} \right] + \vec{X} \left[\frac{n+1}{2} \right] \quad (5)$$

Equation 5 is employed to determine the median vector from the initial query term vectors, where X_k is the ordered set of vectors for a certain query. n is the vector length, which is 300.

Here are the steps involved in the proposed approach:

1. Choose a user query Q containing n terms.
2. Use a pre-trained WE model to obtain a vector v_i for each term t_i in Q.
3. Compute the median vector $m(v_k)$ for the associated terms of Q using DMNs.
4. Use cosine similarity to calculate the similarity between $m(v_k)$ and all other vectors in the WE corpus.
5. Select the top k most similar vectors to $m(v_k)$ from the WE corpus, and create a set W containing the corresponding vectors w_1, w_2, \dots, w_k .
6. Obtain the words tw_1, tw_2, \dots, tw_k that match the vectors w_1, w_2, \dots, w_k from the WE corpus using the pre-trained model.
7. Incorporate the terms tw_1, tw_2, \dots, tw_k into the initial query Q to create a revised query Q'.
8. Use the updated query Q' to fetch documents.

3.3. Embedding-Based Query Expansion Method Using DMNs (EQE1+DMNs)

The EQE1 procedure was proposed in 2016 by Zamani and Croft (2016). This technique assumed the presence of conditional independence between all query terms, wherein it was stated that the candidate terms must be like the query terms, which are selected for expansion. Hence, the researchers proposed the incorporation of the DMNs

technique into the EQE1 approach for improving the efficiency and performance of the AIR retrieval.

The proposed EQE1+DMNs approach was carried out in the following manner. Initially, the original query term vectors, v_i , were used for finding set W , wherein set $W = \{w_1, w_2, \dots, w_k\}$ displayed the vectors in the WE corpus which showed the highest similarity to the query term vectors, v_i . Thereafter, based on the DMNs, the researchers calculated the Median Vector $m(V)$. After this step, the researchers compared the vectors in Set W with the $m(V)$. The vectors in Set W , which displayed a similarity to $m(V) \rightarrow 0.7$, were selected. Thus, the vectors that were selected were used for determining their similar and corresponding words in a WE corpus. These new terms would be added to the primary query for acquiring the new query. This newly acquired query would be used for retrieving the new documents.

Here are the steps for the proposed Embedding-Based Query Expansion Method using Dynamic Memory Networks (EQE1+DMNs):

1. Start with a user query Q consisting of n terms.
2. Using the created WE model, obtain the corresponding vector v_i from the WE corpus for each term t_i in Q .
3. Calculate the cosine similarity between the median vector $m(v_k)$ of associated terms (computed using DMNs) and all the other vectors within the WE corpus.
4. The top- k vectors most similar to v_i in the WE corpus and set W were chosen.
5. Compute the median vector $m(v_k)$ of associated terms using DMNs.
6. Calculate the cosine similarity between $m(v_k)$ and vectors in set W .
7. Choose the vectors from set W that have a similarity score of 0.7 or higher for $m(v_k)$.
8. Retrieve the words corresponding to the selected vectors w_1, w_2, \dots, w_k from the corpus based on the created WE model.
9. Incorporate the chosen words into the original query Q to form a revised query Q' .
10. Use the updated query Q' to fetch documents.

3.4. Prospect-Guided Query Expansion Strategy Based on DANs (V2Q+DMNs)

During the application of the proposed techniques, the researchers have suggested using the median vectors of the DMNs for generating the expansion candidate vectors. It was seen that this median vector is not affected by their place of query terms present in a vector space. Further-

more, the median vector is regarded as the real vector in a vector space, not like the average vector in the DANs that was described earlier (Farhan et al., 2021a; 2021b). The average vector was a random number present in the WE vector space. This new method is named as V2Q+DMNs. This method could improve V2Q performance, since the candidate expansion set is developed with the help of a median vector rather than using the direct individual query terms.

This strategy is implemented in the following manner: Step 1 involves finding set $W = \{w_1, w_2, \dots, w_k\}$, wherein it consists of vectors in the WE corpus which showed the highest similarity to the primary query term vectors. Thereafter, the researchers determined the $m(V)$ (median vector) for these primary query term vectors. They also determined Set $V = \{v_1, v_2, \dots, v_k\}$, which included the vectors which showed the highest similarity to the $m(V)$ in this WE corpus. Thereafter, they calculated the similarity (with the help of the cosine similarity) between all vectors in Sets W and V and the $m(V)$, and selected the vectors from both the sets which showed the highest similarity to $m(V)$. Lastly, the vectors that were selected were used for deriving the real words from the WE corpus with the help of the WE model. These new words were then added to the basic query and the documents were retrieved using this new query.

This method hypothesized that the $m(V)$ vector was placed in the centre of the query term vectors present in the vector space. This vector highlighted the actual meaning of the complete query. Thus, it could be seen that the candidate expansion terms that were generated using this proposed vector could be useful expansion terms.

Here is a paraphrased version of the given steps:

1. Choose a user query Q consisting of n terms.
2. For each term t_i in the query Q , use the WE model to derive the associated vector v_i .
3. Calculate the cosine similarity between v_i and every other vector in the WE corpus.
4. Choose the k vectors most similar to v_i from the corpus and compile them into a set W .
5. Compute the median vector $m(v_k)$ for each vector v_i by using DMNs to associate terms with vectors.
6. Compute the cosine similarity between $m(v_k)$ and every other vector in the WE corpus.
7. Identify the k vectors in the corpus that are closest to $m(v_k)$ and assemble them into a set V .
8. Compare the vectors in set W and set V to $m(v_k)$.
9. Select the vectors in the intersection of W and V that

have a cosine similarity to $m(vk) \geq 0.7$.

10. Retrieve the words corresponding to the vectors in set W and set V from the WE corpus.

11. Include the retrieved words in the original query Q to create an updated query Q' .

12. Use the revised query Q' to get new documents.

4. EXPERIMENTS

4.1. Experimental Setup

In the experimental setup, the researchers meticulously selected and compared various techniques, including the primary BM25 model, DANs, EQE1, and V2Q techniques, alongside the proposed DMNs-based expansion approaches. To ensure a comprehensive evaluation, these approaches were benchmarked against each other using standard evaluation metrics. Additionally, the researchers utilized the TREC 2001/2002 Arabic newswire dataset, a widely accepted benchmark for evaluating Arabic text retrieval systems, which comprises news articles from the Middle East published between 1994 and 2000. This dataset has been used recently by many researchers involved in the retrieval of Arabic texts (Abdelali et al., 2016; Darwish & Ali, 2012; El Mahdaouy et al., 2018a). It was seen that the TREC 2001/2002 Arabic newswire consisted of three standard TREC collections, i.e., TREC 2001, containing 25 queries; TREC 2002, containing 50 queries; and TREC 2001/2002, containing 75 queries. Moreover, the creation of the WE corpus using the Word2Vec (Mikolov et al., 2013) process involved processing 383,872 of Arabic newspaper articles from *Agence France Presse*, resulting in a comprehensive dataset exceeding 1 GB in size after encoding and distribution.

Regarding preprocessing, stemming was applied to the dataset using the Farasa stemmer (Abdelali et al., 2016), a recognized tool for Arabic language processing, to address the significant impact of stemming on the performance of Arabic text retrieval systems. The Farasa stemmer was seen to be a very effective stemmer for the Arabic language (El Mahdaouy et al., 2018a). Stemming helps in reducing words to their root form, enhancing the efficiency and effectiveness of retrieval processes. Furthermore, the researchers ensured a consistent experimental setup by retrieving 100 documents for each query and considering nine baselines for evaluation, i.e., (1) Probabilistic Okapi BM25 model (without expansion); (2) BM25+DANs, proposed by Farhan et al. (2021a); (3) BM25+DANs-PRF, proposed by Farhan et al. (2021b); (4) Embedding-based Query Expansion (EQE1) (Zamani & Croft, 2016);

(5) EQE1+DANs, proposed by Farhan et al. (2021a); (6) EQE1+DANs-PRF, proposed by Farhan et al. (2021b); (7) Prospect-Guided QE strategy (V2Q) strategy; (8) V2Q+DANs, proposed by Farhan et al. (2021a); and (9) V2Q+DANs-PRF, proposed by Farhan et al. (2021b), covering a range of AQE techniques and their variants. These baselines included models without expansion, expansion models with DANs, EQE1, and V2Q, as well as their PRF variants proposed in previous studies.

For implementation, the experiments utilized the Whoosh search engine library in Python, a popular tool for constructing search engines and assessing retrieval systems (Mukherjee & Kumar, 2019). The semantic similarity between terms was determined by calculating the cosine similarity of their WE vectors, providing a quantitative measure of similarity between words in the embedding space. Additionally, the assessment metrics employed included MAP and Precision at the top 10 retrieved documents (P@10), provided insights into the effectiveness of the proposed DMNs-based AQE approach compared to existing techniques. Overall, the experimental methodology was carefully designed and executed to ensure robust evaluation and reliable comparisons between different AQE methods.

4.2. Metrics for Assessing Performance

The evaluation of the retrieval performance of the proposed DANs AQE approach and the four baselines was conducted using MAP of the top 100 ranked documents as the primary metric. Additionally, P@10 was also considered. Recall-precision curves were generated to depict the performance of each method across various standard recall levels. The formulas for calculating each performance metric are provided below.

$$Recall = \frac{|Rel| \cap |Ret|}{|Ret|} \quad (6)$$

$$Precision = \frac{|Rel| \cap |Ret|}{|Rel|} \quad (7)$$

$$MAP = \frac{1}{N} \sum_{i=1}^n AP_i \quad (8)$$

In the equation provided, “Ret” refers to the total number of documents that are retrieved by the search engine, while “Rel” refers to the total number of documents that are considered relevant within the dataset being evaluated.

$$AP = \sum_{i=1}^n Precision_i \cdot \Delta Recall_i.$$

The precision at each rank i of the retrieved documents is multiplied by the change in recall from items $i-1$ to i , and the resulting values are summed up to obtain the average precision. Here, the variable n refers to the overall number of retrieved documents, and $Precision_i$ is the proportion of relevant documents among the top i retrieved documents. $\Delta Recall_i$ is the difference in recall between the top $i-1$ and top i retrieved documents.

5. FINDINGS AND ANALYSIS

Table 2 displays precision at 10 and values of MAP for each of the proposed AQE methods and the baseline approaches. The values in mark a) signify that this value has outperformed the baseline approach when utilizing DMNs technique. The experiments have been carried out on three dataset collections: TREC 2001, TREC 2002, and TREC 2001/2002. TREC 2001 has 25 queries, TREC 2002 has 50, and TREC 2001/2002 blends the first two collec-

tions with 75 queries. The collections hit the same dataset corpus. In the experiments, the top 100 retrieved documents were considered.

The AQE methods evaluated include BM25+DMNs, EQE1+DMNs, and V2Q+DMNs, where DMNs are utilized to identify and select potential candidate expansion terms. These methods are compared against baseline approaches such as BM25 (without expansion), BM25+DANs, EQE1, EQE1+DANs, V2Q, and V2Q+DANs methodologies.

The recommended methods were compared against the probabilistic model BM25 (without expansion), and the BM25+DANs, EQE1, EQE1+DANs, V2Q, and V2Q+DANs methodologies. Table 2 depicts the experiment outcomes of all methodologies with regards to MAP and indicates P@10. The hypothesis assumes that DMNs can recommend more pertinent expansion terms, and can overpower the drawbacks of DANs, which would later enhance the AIR systems' retrieval performance.

Generally, the experiment outcomes depicted in Table 2 indicate that the improved BM25+DMNs are outclassing the BM25, BM25+DANs, and BM25+DANs-PRF

Table 2. MAP and P@10

| Technique | Collection | | | | | |
|--------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | TREC 2001 | | TREC 2002 | | TREC 2001/2002 | |
| | MAP | P@10 | MAP | P@10 | MAP | P@10 |
| Technique 1 | | | | | | |
| BM25 | 31.30 | 42.10 | 28.70 | 35.80 | 29.50 | 37.90 |
| BM25+DANs | 19.40 | 26.20 | 28.50 | 39.60 | 25.40 | 35.10 |
| BM25+DANs-PRF | 25.60 | 22.90 | 29.30 | 37.60 | 31.10 | 34.50 |
| BM25+DMNs | 32.30 ^{a)} | 55.50 ^{a)} | 37.20 ^{a)} | 40.10 ^{a)} | 34.90 ^{a)} | 47.80 ^{a)} |
| Technique 2 | | | | | | |
| EQE1 | 30.70 | 42.90 | 25.70 | 33.50 | 26.90 | 30.40 |
| EQE1+DANs | 30.70 | 42.90 | 25.70 | 33.50 | 26.90 | 30.40 |
| EQE1+DANs-PRF | 31.30 | 43.00 | 30.20 | 34.30 | 32.30 | 34.20 |
| EQE1+DMNs | 32.30 ^{a)} | 51.00 ^{a)} | 39.80 ^{a)} | 42.70 ^{a)} | 36.10 ^{a)} | 46.80 ^{a)} |
| Technique 3 | | | | | | |
| V2Q | 27.00 | 35.20 | 26.20 | 33.30 | 26.50 | 33.90 |
| V2Q+DANs | 30.20 | 38.90 | 32.90 | 43.90 | 32.00 | 42.30 |
| V2Q+DANs-PRF | 28.30 | 35.70 | 31.20 | 34.60 | 30.70 | 34.10 ^{a)} |
| V2Q+DMNs | 32.20 ^{a)} | 59.30 ^{a)} | 36.60 ^{a)} | 45.60 ^{a)} | 34.40 ^{a)} | 52.40 ^{a)} |

MAP, mean average precision; P@10, precision at the top 10; TREC, Text REtrieval Conference; DANs, deep averaging networks; PRF, pseudo relevance feedback; EQE1, Embedding-Based Query Expansion Method; V2Q, Prospect-Guided Query Expansion Strategy.

^{a)}The values signify that this value has outperformed the baseline approach when utilizing DMNs technique.

models with regard to MAP and P@10 for all TREC collections. Furthermore, on comparison with the baseline approaches EQE1, EQE1+DANs-PRF, and EQE1+DANs, the EQE1+DMNs method indicates the highest MAP and P@10 values for TREC 2001, TREC 2002, and TREC 2001/2002, and considerably outclasses the baseline approach EQE1 and EQE1+DANs. Moreover, the suggested V2Q+DMNs method significantly outshines their baseline methodologies V2Q, V2Q+DANs, and V2Q+DANs-PRF with respect to Precision at 10 and MAP at 10 for each of the TREC data sets.

In summary, in Table 2, the experimental results were presented in detail to compare the performance of different AQE methods and baseline approaches across three TREC dataset collections. Each method's P@10 and MAP scores are provided, allowing for a comprehensive evaluation of their retrieval effectiveness. The AQE methods considered, such as BM25+DMNs, EQE1+DMNs, and V2Q+DMNs, leverage DMNs to select expansion terms, aiming to address query-document vocabulary mismatch and enhance search results. The comparison against baseline approaches, including BM25, BM25+DANs, EQE1, EQE1+DANs, V2Q, and V2Q+DANs, allows for insights into the relative performance improvements achieved by integrating DMNs. By analysing the results, it can be identified which AQE methods yield the highest precision and MAP values, thereby informing decisions regarding the selection and implementation of query expansion techniques in IR systems.

Finally, our research results diverge from early studies, primarily in the incorporation of DMNs for query expansion within the context of IR. While previous studies may have explored AQE techniques using traditional methods or simpler models, this research introduces a novel approach by leveraging DMNs to generate candidate expansion terms based on the median vector of query term vectors derived from Word2Vec WE. This departure from conventional methods results in significant improvements in retrieval performance, as demonstrated through empirical evaluation against established techniques, including Okapi BM25, V2Q, EQE1, and their variations with DMNs.

6. CONCLUSIONS, FUTURE WORKS, AND LIMITATIONS

In conclusion, this study proposed the utilization of DMNs in improving the performance of AIR. By leveraging DMNs to generate candidate expansion terms based

on the median vector of query term vectors derived from Word2Vec WE, the study aimed to address limitations in existing techniques. Comparative analysis with various methods including Okapi BM25, V2Q, V2Q+DANs, BM25+DANs, EQE1, and EQE1+DANs revealed significant improvements in retrieval performance with the incorporation of DMNs into WE models and the BM25 model. Specifically, BM25+DMNs utilized the median vector of original query term vectors for expansion, while EQE1+DMNs and V2Q+DMNs combined original query term vectors with the median vector of DMNs. This approach yielded more relevant expansion terms compared to the DANs method. Key findings underscored the effectiveness of DMNs in enhancing query expansion and ultimately improving search results.

In looking toward future research endeavours, future research should delve into the semantic aspects of query terms to further refine retrieval techniques and ensure the retrieval of even more relevant terms. By delving deeper into the semantic understanding of queries, researchers can continue to refine and advance IR methodologies. This study's contribution lies in not only proposing a novel approach but also demonstrating its efficacy through empirical evaluation against established techniques. The findings underscore the potential of DMNs to enhance query expansion processes, offering valuable insights for the ongoing evolution of IR systems.

While the study addresses several limitations, further considerations could encompass potential challenges in implementing the proposed model and its generalizability. Challenges in implementation may arise due to computational complexity and resource requirements, particularly in real-world applications where computational resources may be limited. Additionally, integrating the DMNs-based AQE approach into existing IR systems may necessitate significant modifications to infrastructure and workflows. Moreover, the study's findings may have limited generalizability beyond the specific context and dataset used. Variations in dataset characteristics, such as document lengths and domain-specific terminology, could impact the model's performance. Additionally, factors such as the quality and size of the training corpus for WE, as well as parameter choices, may influence the effectiveness of the model across different languages, domains, and datasets. Thus, while the study provides valuable insights, further research is needed to validate its applicability in diverse contexts.

Future research endeavours should emphasize the divergence of our results from prior studies, highlighting

the efficacy of DMNs in query expansion. Furthermore, exploring the integration of DMNs into related libraries and databases could enhance IR methodologies, fostering advancements in library and information science research.

7. AUTHOR CONTRIBUTIONS

The author contributions can be summarized as follows:

1. Proposal of DMNs Integration: The study proposes the integration of DMNs into the AIR process to improve performance.

2. Utilization of DMNs for Candidate Expansion Terms: By leveraging DMNs, the study aims to generate candidate expansion terms based on the median vector of query term vectors derived from Word2Vec WE, addressing limitations in existing techniques.

3. Comparative Analysis and Evaluation: The study conducts a comparative analysis with various methods, including Okapi BM25, V2Q, V2Q+DANs, BM25+DANs, EQE1, and EQE1+DANs, demonstrating significant improvements in retrieval performance with the incorporation of DMNs into WE models and the BM25 model.

4. Identification of Future Research Directions: The study suggests future research directions, focusing on delving into the semantic aspects of query terms to further refine retrieval techniques and ensure the retrieval of even more relevant terms, offering valuable insights for the ongoing evolution of IR systems.

Overall, the contributions include proposing a novel approach, demonstrating its effectiveness through empirical evaluation, and identifying potential future research directions for advancing IR methodologies.

ACKNOWLEDGMENTS

We express our gratitude to the LDC for granting us the LDC2001T55 Arabic Newswire Part 1 without any charges and for presenting us with the LDC Data Scholarship in the autumn of 2012. This study is partially supported by the Universiti Kebangsaan Malaysia grant: DCP-2017-007/4.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

- Abbache, A., Meziane, F., Belalem, G., & Belkredim, F. Z. (2016). Arabic query expansion using WordNet and association rules. In *Information Retrieval and Management: Concepts, Methodologies, Tools, And Applications* (pp. 1239-1254). IGI Global.
- Abdelali, A., Darwish, K., Durrani, N., & Mubarak, H. (2016, June 12-17). *Farasa: A fast and furious segmenter for Arabic*. Paper presented at Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, San Diego, California.
- Abu El-Khair, I. (2007). Arabic information retrieval. *Annual Review of Information Science and Technology*, 41(1), 505-533.
- Aklouche, B., Bounhas, I., & Slimani, Y. (2018, November 13-16). *Query expansion based on NLP and word embeddings*. Paper presented at Text Retrieval Conference (TREC), Gaithersburg, Maryland, USA.
- ALMasri, M., Berrut, C., & Chevallet, J. P. (2016, March 20-23). *A comparison of deep learning based query expansion with pseudo-relevance feedback and mutual information*. Paper presented at Advances in Information Retrieval: 38th European Conference on IR Research, ECIR 2016, Padua, Italy.
- Alsmearat, K., Al-Ayyoub, M., & Al-Shalabi, R. (2014, May 19-22). *An extensive study of the Bag-Of-Words approach for gender identification of Arabic articles*. Paper presented at 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA), Aqaba, Jordan.
- Azad, H. K., & Deepak, A. (2019). Query expansion techniques for information retrieval: A survey. *Information Processing & Management*, 56(5), 1698-1735. <https://doi.org/10.1016/j.ipm.2019.05.009>
- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retrieval* (Vol. 463). ACM press New York.
- Belkin, N. J. (2005). Anomalous state of knowledge. In K. E. F. S. E. Robertson, & E. F. McKechnie (Eds.), *Information Research: Theory and Practice* (pp. 1-12). American Society for Information Science and Technology.
- Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends® in Machine Learning*, 2(1), 1-127. <http://dx.doi.org/10.1561/22000000006>
- Cai, F., & De Rijke, M. (2016). A survey of query auto completion in information retrieval. *Foundations and Trends® in Information Retrieval*, 10(4), 273-363. <http://dx.doi.org/10.1561/15000000055>
- Carpineto, C., & Romano, G. (2012). A survey of automatic query expansion in information retrieval. *Acm*

- Computing Surveys (CSUR)*, 44(1), 1-50. <https://doi.org/10.1145/2071389.2071390>
- Crimp, R., & Trotman, A. (2018, December 10-12). *Refining query expansion terms using query context*. Paper presented at Proceedings of the 23rd Australasian Document Computing Symposium, Melbourne, Australia.
- Croft, W. B., Metzler, D., & Strohman, T. (2010). *Search engines: Information retrieval in practice* (Vol. 520). Addison-Wesley Reading.
- Darwish, K., & Ali, A. M. (2012, July 8-14). *Arabic retrieval revisited: Morphological hole filling*. Paper presented at Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, Jeju, Korea.
- Diaz, F., Mitra, B., & Craswell, N. (2016, August 7-12). *Query expansion with locally-trained word embeddings*. Paper presented at Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany.
- El Mahdaouy, A., El Alaoui, S. O., & Gaussier, E. (2018a). Improving Arabic information retrieval using word embedding similarities. *International Journal of Speech Technology*, 21(1), 121-136. <https://doi.org/10.1007/s10772-018-9492-y>
- El Mahdaouy, A., El Alaoui, S. O., & Gaussier, E. (2018b). Word-Embedding-based pseudo-relevance feedback for Arabic information retrieval. *Journal of Information Science*, 45(4), 429-442. <https://doi.org/10.1177/0165551518792210>
- Esposito, M., Damiano, E., Minutolo, A., De Pietro, G., & Fujita, H. (2020). Hybrid query expansion using lexical resources and word embeddings for sentence retrieval in question answering. *Information Sciences*, 514, 88-105. <https://doi.org/10.1016/j.ins.2019.12.002>
- Faqeeh, M., Abdulla, N., Al-Ayyoub, M., Jararweh, Y., & Quwaider, M. (2014, August 27-29). *Cross-Lingual short-text document classification for Facebook comments*. Paper presented at 2014 International Conference on Future Internet of Things and Cloud, Vienna, Austria.
- Farhan, Y. H., Mohd, M., & Noah, S. A. M. (2020). Survey of automatic query expansion for Arabic text retrieval. *Journal of Information Science Theory and Practice*, 8(4), 67-86. <https://doi.org/10.1633/JISTaP.2020.8.4.6>
- Farhan, Y. H., Mohd Noah, S. A., Mohd, M., & Atwan, J. (2021a). Word-Embedding-based query expansion: Incorporating deep averaging networks in Arabic document retrieval. *Journal of Information Science*, 49(5), 1168-1186. <https://doi.org/10.1177/01655515211040659>
- Farhan, Y. H., Noah, S. A. M., Mohd, M., & Atwan, J. (2021b). Word embeddings-based pseudo relevance feedback using deep averaging networks for Arabic document retrieval. *Journal of Information Science Theory and Practice*, 9(2), 1-17. <https://doi.org/10.1633/JISTaP.2021.9.2.1>
- Fernández-Reyes, F. C., Hermosillo-Valadez, J., & Montes-y-Gómez, M. (2018). A prospect-guided global query expansion strategy using word embeddings. *Information Processing & Management*, 54(1), 1-13. <https://doi.org/10.1016/j.ipm.2017.09.001>
- Guirat, S. B., Bounhas, I., & Slimani, Y. (2016). Combining indexing units for Arabic information retrieval. *International Journal of Software Innovation (IJSI)*, 4(4), 1-14. <https://doi.org/10.4018/IJSI.2016100101>
- Kim, H. K., Kim, H., & Cho, S. (2017). Bag-Of-Concepts: Comprehending document representation through clustering words in distributed representation. *Neurocomputing*, 266, 336-352. <https://doi.org/10.1016/j.neucom.2017.05.046>
- Larkey, L. S., Ballesteros, L., & Connell, M. E. (2002, August 11-15). *Improving stemming for Arabic information retrieval: Light stemming and co-occurrence analysis*. Paper presented at Proceedings of the 25th annual international ACM SIGIR conference on Research and Development in Information Retrieval, Tampere, Finland.
- Lavrenko, V., & Croft, W. B. (2017, August 2). *Relevance-based language models*. Paper presented at ACM SIGIR Forum, New York, USA.
- Lv, Y., & Zhai, C. (2011, October 24-28). *Lower-bounding term frequency normalization*. Paper presented at Proceedings of the 20th ACM International Conference on Information and Knowledge Management, Glasgow, UK.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013, December 5-10). *Distributed representations of words and phrases and their compositionality*. Paper presented at Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA.
- Miyaniishi, T., Seki, K., & Uehara, K. (2013, October 27-November 1). *Improving pseudo-relevance feedback via tweet selection*. Paper presented at Proceedings of The 22nd ACM International Conference on Information & Knowledge Management, San Francisco, CA, USA.
- Mohsen, G., Al-Ayyoub, M., Hmeidi, I., & Al-Aiad, A. (2018, April 9-11). *On the automatic construction of an Arabic thesaurus*. Paper presented at 2018 9th International Conference on Information and Communication Systems (ICICS), Amman, Jordan.
- Mukherjee, S., & Kumar, N. (2019, December 12-14). *Duplicate question management and answer verification system*. Paper presented at 2019 IEEE Tenth International Conference on Technology for Education (T4E), Bhubaneswar, India.
- Mustafa, M., AbdAlla, H., & Suleman, H. (2008, December

- 2-5). *Current approaches in Arabic IR: A survey*. Paper presented at International Conference on Asian Digital Libraries, Hyderabad, India.
- Nwesri, A. F. A., & Alyagoubi, H. A. (2015, August 31-September 4). *Applying Arabic stemming using query expansion*. Paper presented at 2015 26th International Workshop on Database and Expert Systems Applications (DEXA), Vienna, Austria.
- Pal, D., Mitra, M., & Datta, K. (2014). Improving query expansion using WordNet. *Journal of the Association for Information Science and Technology*, 65(12), 2469-2478. <https://doi.org/10.1002/asi.23143>
- Raza, M. A., Mokhtar, R., Ahmad, N., Pasha, M., & Pasha, U. (2019). A taxonomy and survey of semantic approaches for query expansion. *IEEE Access*, 7, 17823-17833. <https://doi.org/10.1109/ACCESS.2019.2894679>
- Robertson, S., & Zaragoza, H. J. F. (2009). The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends® in Information Retrieval*, 3(4), 333-389. <https://doi.org/10.1561/1500000019>
- Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M. M., & Gatford, M. (1995). *Okapi at TREC-3*. Paper presented at Third Text REtrieval Conference (TREC-3), Gaithersburg, MD, USA.
- Roy, D., Paul, D., Mitra, M., & Garain, U. (2016). *Using word embeddings for automatic query expansion*. Paper presented at CoRR, Austin, TX, USA.
- Takeuchi, S. I., Sugiura, K., Akahoshi, Y., & Zettsu, K. (2017). Spatio-temporal pseudo relevance feedback for scientific data retrieval. *IEEJ Transactions on Electrical and Electronic Engineering*, 12(1), 124-131. <https://doi.org/10.1002/tee.22352>
- Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37, 141-188. <https://doi.org/10.1613/jair.2934>
- Zamani, H., & Croft, W. B. (2016, September 12-16). *Embedding-based query language models*. Paper presented at Proceedings of The 2016 ACM International Conference on the Theory of Information Retrieval, Delaware, Newark, USA.
- Zou, S., Tao, G., Wang, J., Zhang, W., & Zhang, D. (2018, July 8-12). *On the equilibrium of query reformulation and document retrieval*. Paper presented at Proceedings of the 2018 ACM SIGIR International Conference on Theory of Information Retrieval, Ann Arbor, MI, USA.