Knowledge Based Word Sense Disambiguation with Distributional Semantic Expansion for the Persian Language

Hossein Rouhizadeh
Shahid Beheshti University
Tehran, IRAN
Email: h.rouhizadeh@mail.sbu.ac.ir

Mehrnoush Shamsfard
Shahid Beheshti University
Tehran, IRAN
Email: m-shams@sbu.ac.ir

Masoud Rouhizadeh
Johns Hopkins University
Baltimore, Maryland,USA
Email: mrouhizadeh@gmail.com

***Abstract*—Word Sense Disambiguation (WSD) can be the key component of downstream NLP applications. Existing WSD methods and systems are mostly developed and evaluated on English and low-resource languages such as Persian have not been well studied. In this paper, we propose a new knowledge-based method for Persian WSD. Using a pre-trained LDA model, we retrieve the topics of each document and assign each ambiguous content word to one of the topics. For each possible sense s of a given word w, we compute the similarity between the FarsNet (the Persian WordNet) gloss of s and the words of the assigned topic of w. We then choose the sense with the highest score as the most probable one. We evaluated our method on a Persian all-words WSD dataset and show that, compared to other knowledge-based methods, we could achieve state-of-the-art performance.**

1. Introduction

Word Sense Disambiguation (WSD) is an open problem in Natural Language Processing (NLP), aims to identify the most relevant meaning of ambiguous words in context. WSD has applications in other NLP tasks such as Machine Translation [1], Information Retrieval and Extraction [2], [3] and Question Answering [4]. WSD methods can be classified into two major classes: supervised and knowledge-based. According to a unified evaluation of different WSD methods by Raganato et al. [5], supervised methods outperform the knowledge-based methods in most cases. This is mainly due to the fact that such systems use sense-annotated corpora such as SemCor for training. Annotated corpora are difficult and expensive to create and they are not available in many languages such as Persian. On the other hand, knowledge-based WSD systems are mostly dependent on lexical resources such as WordNet [6] that are available in a larger number of languages, making these types of methods applicable on a wider variety of languages.

It is worth noting that many knowledge-based WSD systems use sense annotated corpora to extract statistical information such as sense frequency (the probability of any given sense s for a particular word). Therefore, they may not be considered *purely* knowledge-based as they require the sense distribution from such corpora. It also shows the necessity of developing *purely* knowledge-based systems (systems with no dependency on sense frequency information from a sense annotated corpora), since such information is not available for low-resource languages like Persian.

Persian is an Indo-European (IE) language that is currently spoken with more than 200 million people in Iran, Afghanistan, and Tajikistan. Due to the lack of resources and tools, the Persian language can be considered as a low resource language. As a result, many NLP tasks including WSD has not been well studied and developed for Persian. Among all the works on Persian NLP, we found Reakbsaz et al. [7]and Sarafzadeh et al. [8] as the most recent developed WSD systems. They both developed Cross-Lingual WSD systems. Persian language also suffers from a lack of standard evaluation framework for WSD. To the best of our knowledge, the only available evaluation platform has been developed by Rekabsaz et al. [9] which is a standard test set for English to Persian Cross-Lingual Word Sense Disambiguation. In this paper, we present a novel knowledge-based method for Persian WSD. The method uses latent Dirichlet allocation (LDA) [10] for semantic expansion. Our method is based on the assumption that lexical units of a given topic are semantically related, and as a result, we expect clustering words in document topics would help with identifying document word senses. Considering this, we first extract the topics of each document (using a pre-trained LDA model) and then assign a topic to each ambiguous content word w. Next, we calculate the score for each sense s of w which is the cosine similarity of the assigned LDA topic of w and the FarsNet gloss of s. The sense with the highest score will be chosen. We evaluate our method on a Persian All-Words WSD dataset to show that our proposed method can outperform other knowledge-based Persian WSD systems. The main contributions of the proposed method are as follows:

1) No dependency on sense annotated corpora for training: unlike many knowledge-based WSD

978-1-7281-8566-8/20/$31.00 ©2020 IEEE

systems, we only require lexical information from the Persian WordNet (FarsNet).

2) Low dependency on lexical graph depth and density: unlike the English WordNet, FarsNet semantic graphs are not very dense and deep. As a result, our system is designed in such a way that it only relies on hypernymy and hyponymy relations in the FarsNet semantic graphs.
3) Training feasibility: our algorithm uses LDA topics and word embeddings as two key components, both requiring only raw text corpora such as Wikipedia.
4) Persian all-words WSD test set creation: To be able to test our WSD system in terms of disambiguating Persian words, we created a standard Persian all-words WSD test set. In creating the test set we follow the instruction of Semeval-2015 multilingual WSD task [11].

We suggest that our method could be an applicable system for other low-resource languages with no sense annotated corpora and rich lexical resources.

2. Related Work

Over recent decades many knowledge-based WSD systems have been developed for a variety of languages. However, due to lack of resources and tools, only limited number of systems have been developed for Persian WSD. The most recent works in Persian WSD are propsed by Mahmoodvand and Hourali [12] and Reakbsaz et al. [7] which are not knowledge-based and as a result, can not be compared to our method.

Compared to the other languages, English has the most number of developed WSD systems and also most amount of progress in this area of study. So, in the following, we briefly review the proposed knowledge-based WSD systems for English and then review two recent works in Persian WSD.

Lesk [13] is one the most traditional WSD algorithms which choose the proper meaning of each context word based on the overlap between definition of different senses and the context words. The algorithm chooses the sense with the highest number of mutual words as the most probable one. Banerjee et al. [14] enhanced Lesk algorithm by taking advantage of term frequency-inverse document frequency scores [15] which weights different words in the context based on their importance. Another extension of Lesk algorithm has been developed by Basile et al. [16] [1]. Instead of counting the mutual words between context and sense definitions, they used word embedding to calculate cosine similarity between the words in the context and gloss of different senses. The system also makes use of SemCor information to give more chance to more predominant senses. Taking advantage of the structure of lexical resources such as WordNet, many graph-based WSD systems has been developed. Mihalcea et al. ([17], [18]) proposed a graph-based method building a graph on the basis of the target word and words in the window of size six on the left and on the right of that. The system takes advantages of combination of several similarity metrics and graph centrality algorithms to compute the importance of different nodes of the built graph. Another development of graph-based WSD systems has been done by Navigli et al. ([19], [20]). Given a sentence γ, the system obtains a subgraph of the entire lexicon and determine the importance of each node by local or global connectivity measures. For each word, the most important node is the most probable sense in the context. Agirre et al. ([21], [22], [23]) also proposed another graph-based WSD method (UKB) which makes use of WordNet to create a graph based on the content words and scores each node by Personalized PageRank. As a result, the node with the highest score will be considered as the best sense of the target word in the context.[2]

In contrast to the other graph-based WSD systems which make use of WordNet graph, Moro et al. [24] proposed a system (Babelfy) which uses BabelNet [25] (a semantic network integrating WordNet with knowledge-based resources such as Wikipedia or Wiktionary) to build the target graph. Babelfy is a unified graph-based approach to Entity linking and WSD which makes use of random walks with restart over BabelNet. It must be noted that most of the recent knowledge-based WSD systems ([16], [24], [22], [26])use sense frequency information in their pipelines. As a result, due to lack of sense distribution information for Persian word senses, they are not applicable for Persian WSD.

As mentioned, only a limited number of studies have been done on Persian WSD. Our survey includes two most recent works on Persian WSD (Reakbsaz et al. [7] and by Mahmoodvand and Hourali [12]) which are cross-lingual and semi-supervised, respectively. Rekabsaz et al. [7] developed a cross-lingual system (with application of English to Persian) which makes use of word embedding to calculate the semantic similarity between context words. The system uses word embedding of the candidate translations of target word and translated context terms to compute similarity between them. Finally, the most similar translation to the context will be chosen as the most suitable one. Another recent system developed for Persian WSD is proposed by Mahmoodvand and Hourali [12] . The system is developed in a semi-supervised manner and aims to disambiguate three frequent Persian words (i.e, Shir (شیر), Rast (راست), Tar (تار)). The system first crawls several Persian documents and then labels the target words in the documents. The system then applies a collaborative learning method on the labeled documents to disambiguate the target words.

1. We will refer to the system as 'Basile14' throughout the paper

2. In the rest of the paper, the systems will be refered as 'UKB'

FarsNet components	Number of occurrence
Noun Synsets	28196
Verb Synsets	5410
Adjective Synsets	6861
Adverb Synsets	1243
All Synsets	41170
Sense per Word	>1.2
Words per Synset	>2.5
Synsets Relation	>3000

TABLE 1: General statistics of FarsNet 3.0. The symbol > means 'greater than'.

		Test Set	Tuning Set	All
# Docs		13	13	16
# Tokens		5045	847	5892
Number of Instances per PoS	Nouns	1764	307	2071
	Verbs	494	70	564
	Adjectives	515	95	610
	Adverbs	111	11	122
Mean Sense per PoS	Nouns	4.0	3.9	4.0
	Verbs	3.4	2.9	3.3
	Adjectives	1.6	1.7	1.6
	Adverbs	1.2	1.3	1.2

TABLE 2: General statistics of SBU-WSD-Corpus

3. Data and Resources

3.1. FarsNet (the Persian WordNet)

WordNets are the most predominant lexical ontology developed for a variety of languages. It was initially developed for English at the University of Princeton [6]. The main components of WordNet are synsets each of which groups a set of synonyms that share a common meaning and POS. Additionally, each synset includes a gloss (a short definition of the synsets words) and a usage example of the synset words. WordNet can be viewed as a lexical graph in which the synsets are connected by lexical or semantic relations (e.g, *Hypernymy*, *Hyponymy*, *Antonymy*, etc.). The latest version of WordNet, i.e, WordNet 3.0, includes 155278 unique English words (117798 nouns, 11592 verbs, 21497 adjectives and 4481 adverbs) organized in 82115 noun synsets, 13767 verb synsets, 18156 adjective synsets and 3621 adverb synsets.
Due to the wide usage of WordNet in many NLP tasks, in recent decades, many other WordNets have been developed for different languages including Persian. The Persian WordNet (FarsNet) is the first Persian lexical ontology [27] for the Persian language which has been developed in the Natural Language Processing lab of Shahid Beheshti University. The first version of FarsNet (FarsNet 1.0), included 17000 Persian words and phrases organized in more than 10000 synsets. Over the last decade, a variety of developments (including [28], [29], [30], [31], [32]) have been done on FarsNtet. FarsNet 3.0, the latest version of FarsNet, includes 10000 Persian words in more than 40000 synsets. Figure 1 shows the synsets of the word people (مردم) in the FarsNet 3.0, obtained by online FarsNet search [3] .General statistics of FarsNet 3.0 are shown in Table 1.
In this research, we used FarsNet 3.0 as the sense inventory for Persian WSD[4].

3.2. Persian all-words WSD dataset creation

Over recent decades, a variety of All-words WSD corpora have been developed for multiple languages. However, to the best of our knowledge, no standard All-words WSD data set is available for Persian. To deal with this, we followed Senseval-2 guidelines to create a standard All-words WSD data set for Persian. The guidelines of Senseval-2 suggest that a standard All-words WSD must have two important features:

1) The test set must include at least 5000 words of running text.
2) All content words of the corpus should be tagged

Following the mentioned guidelines, we selected 19 documents from our in-house news corpus and then manually (5892 words of Persian running text) annotated the context words with sense from FarsNet 3.0 senses. The documents include a variety of domains (e.g Sports, Medical, Science, Technology, etc) and are consist of 5892 words of Persian running text. We first manually tokenized, PoS tagged, and lemmatized all the documents. As the next step, we asked two native Persian speakers to manually annotate the context words of the corpus with senses from FarsNet 3.0 sense inventory. Then an expert linguist revised all tags and re-annotated all the words with different tags. The corpus includes 3371 sense annotated words (2073 nouns, 566 verbs, 610 adjectives, and 122 adverbs). We also divided the corpus into two tuning set and test set parts. The tuning set and test set include 3 and 16 documents, respectively. In Table 2, we show the statistics of the corpus. The statistics include the number of documents, number of tokens, number of sense annotated words per PoS (noted as Number of Instances per PoS), and ambiguity level of words per PoS (noted as mean sense per PoS) for both tuning and test set of the corpus. It worth noting that the developed corpus is not large enough to be used as a training set and can be only used as a testbed of Persian WSD systems.

3.3. Persian-News Corpus

One of the key components of many NLP tasks is a large corpus. Typically, many NLP researchers use Wikipedia dumps[5] which is freely available in several languages and is large enough. However, Persian Wikipedia articles are often short and do not seem a reliable source

3. http://farsnet.nlp.sbu.ac.ir/Site3/Modules/Browser/Default.jsp
4. FarsNet web service is freely available at farsnet.nlp.sbu.ac.ir
5. www.dumps.wikimedia.org

Figure 1: FarsNet synsets of the word *people* (مردم)

especially for NLP tasks such as building word embeddings that need a large corpus to train. In order to deal with this, we crawled $1,000,000$ Persian news from different Persian news websites to train our models on that.

4. Method

Our WSD system takes a document as input and identifies the most relevant meaning of polysemous content words. For each ambiguous word w in the document, the system retrieves all possible senses from FarsNet. Next, for each sense s it computes a score and selects the sense with the highest score as the most probable one. The overall architecture of the system is shown in Figure 2. A key assumption in our method is that words in any given LDA topic are semantically inter-related. We, therefore, anticipate that using lexical clusters within document LDA topics will help in identifying document word senses, and consequently, disambiguating content words of a document. We first trained an LDA model on the Persian News corpus (see section 3.3) using Gensim software package [33], and then we utilized the existing function of Gensim to retrieve topics of a document and assign each ambiguous word w to one of these topics. In addition, we obtain the word vectors from word2vec word embeddings. The model is trained by Gensim software package on Persian News corpus.
Next, for each sense s of word w, we identify score of s as the cosine similarity between the FarsNet sense vector of s (the average of word vectors in the FarsNet gloss of s, its hyponym, and hypernyms) and a lexical cluster of topic words. The presence of unrelated words in topics will be problematic. In order to deal with this, we create a lexical cluster by only the words which are highly similar (above a certain threshold) to the target word. As a result, we calculate the score of each s based on the similarity of s sense vector with only the words in the created cluster. For instance, consider document D in Figure 2 shares a topic including the words *people*(مردم), $Human$(انسان), $father$(پدر), *Population*(جمعیت), etc. The algorithm assigns the target word *people*(مردم) to this topic and creates a lexical cluster which consists of words such as *population*(جمعیت), $Human$(انسان), $father$(پدر), etc. (The cosine similarity between the word *people*(مردم) and each of these words satisfies a certain threshold). For each sense, the algorithm computes the similarity between the built sense vector and the words of the cluster as its first score. More formally, if cluster units are w_1,...,w_n and sv_s be s sense vector, $score_1$ of s will be computed as follows:

$$Score(s) = \frac{1}{n}\sum_{i=1}^{n} Cos(sv_s, w_i) \tag{1}$$

Another challenge of using LDA in a WSD pipeline is choosing the optimal number of k for topics number. Although some methods have been developed to determine an efficient k, none of them can determine the optimal k exactly. As observed in [34] choosing a high number for k will results in over-clustering in topics; however, choosing a low value for k topics will be so broad. Since there is no difference between our evaluations with

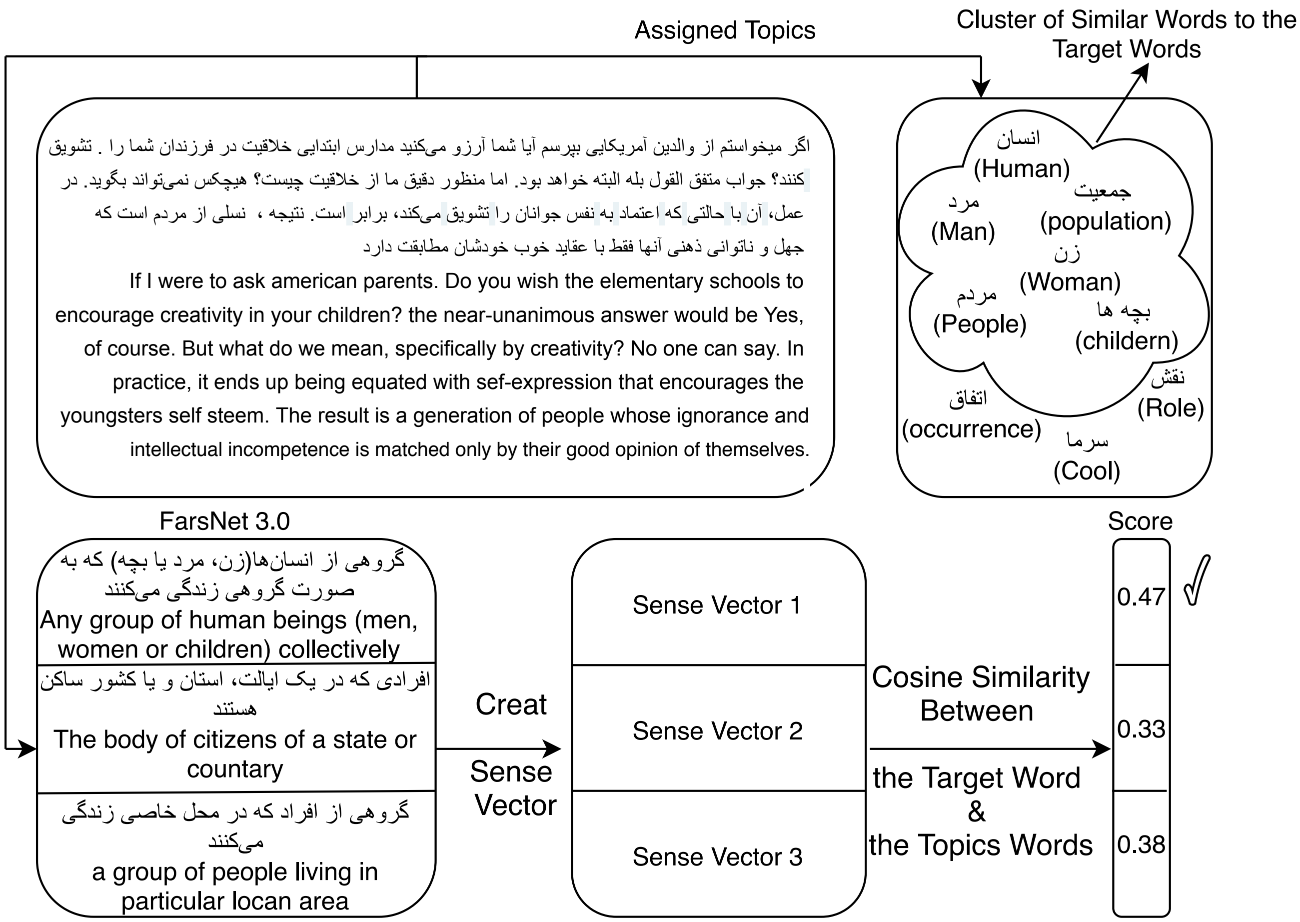

Figure 2: Overall architecture of the propsed dismabiguation pipeline, disambiguating the word *people* (مردم).

training LDA on 90, 100, and 110 topics; it seems that our approach to cluster lexical units of topics is helpful to make the model resistance about choosing the optimal number for k. However; when k is too low (30 in our experiments), this approach cannot find a good lexical cluster. it seems that it is because of the broadness of topics.

When no topic is assigned to the word (for instance in compound words), our system chooses the FarsNet first sense as the most common sense in the Persian language to disambiguate the word meaning.

5. Experimental Setup

5.1. Comparision systems

Since no sense tagged corpora is available for Persian, we can just compare our method with the English WSD systems which do not use sense frequency. Among these systems, we compare our method with Extended Lesk, UKB, and Basile14 systems as the best performing English knowledge-based systems. Our evaluation also includes FarsNet 1st sense as the baseline approach. The approach is context-independent and always chooses the first sense of each FarsNet entry as the best one. [6] Although both UKB and Basile14 systems make use of sense frequency, it is optional and can be enabled. Enabling sense frequency, we applied the algorithms on the FarsNet graph. As mentioned in section 2, the pipeline of the Basile14 system includes a word embedding model. To have a fair comparison, we used the same word2vec model for Basil14 and our proposed method.

5.2. Parameters

We fixed the parameters of our proposed method and baselines by optimizing $F1$ on the tuning set of the developed All-words WSD data set. Note that the only baseline which includes a parameter to tune is Basile14. The pipeline of the Basile14 system only includes one parameter to tune, i.e. context size. We used the available implementation to evaluate systems with context sizes

6. The codes of Extended Lesk, Basile14 and UKB are available at https://github.com/pippokill/lesk-wsd-dsm and https://github.com/alvations/pywsd and https://github.com/asoroa/ukb , respectievly

	SBU-WSD-Corpus
FarsNet 1st sense	48.4
Extended Lesk	48.9
Basile14	62.7
Agirre18(UKB)	58.4
WSD-DSE	**63.3**

TABLE 3: The F1 performance of the proposed Persian WSD system on the noun subset of the SBU-WSD-Corpus. The best result is in bold.

3, 5, 10, 20, and the whole text. We only report the best result obtained by context size 3. During tuning the parameters of our proposed method, we found 100 as the best number for LDA topic number and 0.25 as a threshold for creating a lexical cluster in topics.

6. Results

The results of our experiments can be found in Table 3. As can be seen in Table 3, our proposed method outperforms Basile14 and UKB on disambiguating Persian nouns. It worth nothing that the former Persian WSD approaches are not knowledge-based and as a result are not comparable with our system.
As it shown in Table 3, the FarsNet 1st sense approach can not be considered as a hard to beat baseline for knowledge-based approaches. As mentioned in section 2, the pipeline of Extended lesk and Basile14 are highly similar. A comparison between the results obtained by these systems indicates that the use of word embedding can have a significant impact on the performance of the system. As the performance of Basile14 improved by a large margin (12 percent), compared to the Extended Lesk. As shown in Table 3, Basile14 can outperform UKB. However, as reported in (Raganato et al. [35]) the UKB system can easily outperform Baile14 in disambiguating English words. The difference between results in English and Persian can be possibly explained by the dependency of the UKB pipeline to the lexical graph. Since FarsNet is not as rich as WordNet (note that WordNet includes 117000 synsets, compared to FarsNet covers 40000 ones). A comparison between the performance of our proposed system with Basil14 shows that using LDA topics can be helpful in a disambiguation pipeline. Both systems are inspired by the Lesk algorithm. However, unlike the Basile14 system, our developed model is not limited to the context words of the document. In other words, our method uses LDA topics of a document as a source of semantically related words to expand the semantic space of the document. In the next section, we discuss the benefit of using LDA topics in our WSD pipeline [36].

7. Discussion

In this section, we analyze the benefit of using LDA topics in our WSD pipeline. A part of SensEval-2 documents is shown in Figure 2. Consider that we want to disambiguate the word (*people*)(مردم) in the document. As can be seen, Due to the lack of related words to this word, it is hard to recognize its correct meaning by using just the context words. Using LDA topics, our system makes use of a number of related words in the assigned topic of *people*(مردم) to disambiguate it. As shown in Figure 2 the similar words to the word *people*(مردم) in the assinged topic are *Human*(انسان), *father*(پدر), *Population*(جمعیت), etc. Taking advantage of such words, our method assigns the highest score to the first sense of *people*(مردم) in FarsNet which is the true meaning of that in the context.

8. Conclusion

In this paper, we presented a WSD system for Persian language that uses LDA topics for semantic expansion of document words. Our evluations shows that using the semantically related words, extracted form document topics would be effective in a dismabiguating context words. One of the problems in LDA topics is the presence of unrelated words. We removed the topics words which their similarity to the WordNet glosses of the topic words is below a certain threshold. A possible extension to this work will include obtaining sense ranking from other resources than sense-annotated corpora. To do this, we plan to use statistical and graph-based sense ranking methods such as PageRank.

References

[1] M. Carpuat and D. Wu, "Improving statistical machine translation using word sense disambiguation," in *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, 2007.

[2] Z. Zhong and H. T. Ng, "Word sense disambiguation improves information retrieval," in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*. Association for Computational Linguistics, 2012, pp. 273–282.

[3] C. D. Bovi, L. Telesca, and R. Navigli, "Large-scale information extraction from textual definitions through deep syntactic and semantic analysis," *Transactions of the Association for Computational Linguistics*, vol. 3, pp. 529–543, 2015.

[4] G. Ramakrishnan, A. Jadhav, A. Joshi, S. Chakrabarti, and P. Bhattacharyya, "Question answering via bayesian inference on lexical relations," in *Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering-Volume 12*. Association for Computational Linguistics, 2003, pp. 1–10.

[5] A. Raganato, J. Camacho-Collados, and R. Navigli, "Word sense disambiguation: A unified evaluation framework and empirical comparison," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, vol. 1, 2017, pp. 99–110.

[6] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to wordnet: An on-line lexical database," *International journal of lexicography*, vol. 3, no. 4, pp. 235–244, 1990.

[7] N. Rekabsaz, M. Lupu, A. Hanbury, and A. Duque, “Addressing cross-lingual word sense disambiguation on low-density languages: Application to persian,” *arXiv preprint arXiv:1711.06196*, 2017.

[8] B. Sarrafzadeh, N. Yakovets, N. Cercone, and A. An, “Cross-lingual word sense disambiguation for languages with scarce resources,” in *Canadian Conference on Artificial Intelligence*. Springer, 2011, pp. 347–358.

[9] N. Rekabsaz, S. Sabetghadam, M. Lupu, L. Andersson, and A. Hanbury, “Standard test collection for english-persian cross-lingual word sense disambiguation,” in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*, 2016, pp. 4176–4179.

[10] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent dirichlet allocation,” *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[11] A. Moro and R. Navigli, “Semeval-2015 task 13: Multilingual all-words sense disambiguation and entity linking,” in *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*, 2015, pp. 288–297.

[12] M. Mahmoodvand and M. Hourali, “Semi-supervised approach for persian word sense disambiguation,” in *2017 7th International Conference on Computer and Knowledge Engineering (ICCKE)*. IEEE, 2017, pp. 104–110.

[13] M. Lesk, “Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone,” in *Proceedings of the 5th annual international conference on Systems documentation*. ACM, 1986, pp. 24–26.

[14] S. Banerjee and T. Pedersen, “Extended gloss overlaps as a measure of semantic relatedness,” in *Ijcai*, vol. 3, 2003, pp. 805–810.

[15] K. Sparck Jones, “A statistical interpretation of term specificity and its application in retrieval,” *Journal of documentation*, vol. 28, no. 1, pp. 11–21, 1972.

[16] P. Basile, A. Caputo, and G. Semeraro, “An enhanced lesk word sense disambiguation algorithm through a distributional semantic model,” in *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 2014, pp. 1591–1600.

[17] R. Mihalcea, “Unsupervised large-vocabulary word sense disambiguation with graph-based algorithms for sequence data labeling,” in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2005, pp. 411–418.

[18] R. Sinha and R. Mihalcea, “Unsupervised graph-basedword sense disambiguation using measures of word semantic similarity,” in *Semantic Computing, 2007. ICSC 2007. International Conference on*. IEEE, 2007, pp. 363–369.

[19] R. Navigli and M. Lapata, “Graph connectivity measures for unsupervised word sense disambiguation.” in *IJCAI*, 2007, pp. 1683–1688.

[20] ——, “An experimental study of graph connectivity for unsupervised word sense disambiguation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 4, pp. 678–692, 2010.

[21] E. Agirre and A. Soroa, “Personalizing pagerank for word sense disambiguation,” in *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 2009, pp. 33–41.

[22] E. Agirre, O. López de Lacalle, and A. Soroa, “Random walks for knowledge-based word sense disambiguation,” *Computational Linguistics*, vol. 40, no. 1, pp. 57–84, 2014.

[23] E. Agirre, O. L. de Lacalle, and A. Soroa, “The risk of sub-optimal use of open source nlp software: Ukb is inadvertently state-of-the-art in knowledge-based wsd,” *arXiv preprint arXiv:1805.04277*, 2018.

[24] A. Moro, A. Raganato, and R. Navigli, “Entity linking meets word sense disambiguation: a unified approach,” *Transactions of the Association for Computational Linguistics*, vol. 2, pp. 231–244, 2014.

[25] R. Navigli and S. P. Ponzetto, “Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network,” *Artificial Intelligence*, vol. 193, pp. 217–250, 2012.

[26] A. Popov, K. Simov, and P. Osenova, “Know your graph. state-of-the-art knowledge-based wsd,” in *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*, 2019, pp. 949–958.

[27] M. Shamsfard, A. Hesabi, H. Fadaei, N. Mansoory, A. Famian, S. Bagherbeigi, E. Fekri, M. Monshizadeh, and S. M. Assi, “Semi automatic development of farsnet; the persian wordnet,” in *Proceedings of 5th global WordNet conference, Mumbai, India*, vol. 29, 2010.

[28] M. Rouhizadeh, M. Shamsfard, and M. A. Yarmohammadi, “Building a wordnet for persian verbs,” *GWC 2008*, p. 406, 2007.

[29] M. Rouhizadeh, A. Yarmohammadi, and M. Shamsfard, “Developing the persian wordnet of verbs: Issues of compound verbs and building the editor,” in *Proceedings of 5th Global WordNet Conference*, 2010.

[30] N. Mansoory, M. Shamsfard, and M. Rouhizadeh, “Compound verbs in persian wordnet,” *International Journal of Lexicography*, vol. 25, no. 1, pp. 50–67, 2012.

[31] F. Khalghani and M. Shamsfard, “Extraction of verbal synsets and relations for farsnet,” in *Proceedings of the 9th Global WordNet Conference (GWC 2018)*, 2018, p. 424.

[32] M. A. Yarmohammadi, M. Shamsfard, M. A. Yarmohammadi, and M. Rouhizadeh, “Sbuqa question answering system,” in *Computer Society of Iran Computer Conference*. Springer, 2008, pp. 316–323.

[33] R. Rehurek and P. Sojka, “Software framework for topic modelling with large corpora,” in *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Citeseer, 2010.

[34] D. Greene, D. O’Callaghan, and P. Cunningham, “How many topics? stability analysis for topic models,” in *Machine Learning and Knowledge Discovery in Databases*, T. Calders, F. Esposito, E. Hüllermeier, and R. Meo, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 498–513.

[35] A. Raganato, J. Camacho-Collados, and R. Navigli, “Word sense disambiguation: A unified evaluation framework and empirical comparison,” in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*. Valencia, Spain: Association for Computational Linguistics, Apr. 2017, pp. 99–110. [Online]. Available: https://www.aclweb.org/anthology/E17-1010

[36] H. Rouhizadeh, M. Shamsfard, and M. Rouhizadeh, “Knowledge-based word sense disambiguation with distributional semantic expansion,” in *Proceedings of the 2019 Workshop on Widening NLP*. Association for Computational Linguistics, 2019.