*World of Computer Science and Information Technology Journal (WCSIT)*
*ISSN: 2221-0741*
*Vol. 9, No. 3, 12-16, 2019*

# Speculation and Negation Detection for Arabic Biomedical Texts

Fatima T. Al-Khawaldeh
Department of Computer Science
University of York
York, United Kingdom

Abstract—There are many reasons behind research on speculation and negation: there is a lot of irrelevant (nonfactual) information, and a huge changing with new discovering information may strengthen or weaken previous knowledge. Speculation and negation values are considered as one of the main factors which play an essential role to predict the factuality of event or sentence. Negation reverses the truth of a statement to give the opposition and speculation increase or decreases the uncertainty of statement. Recently, Deep Neural Networks (DNN) have proven better performance to distinguish factual from nonfactual information. Most previous approaches have been dedicated to the English language. To our knowledge, there is no previous developed research to identify the negative or speculative expression for biomedical texts in the Arabic language. This research will develop DNN-based Speculation and negation detection models that able to check claims (negated or speculated sentences) by considering syntactic paths between speculation or negation cues and the remaining words (candidates) in biomedical texts, using Stanford dependency parser. In this paper, the implemented models are evaluated based on the BIOARABIC corpus. Experiments on BIOARABIC corpus show that DNN models achieve a competitive performance and the Attention based Bidirectional Long Short-Term Memory model achieves the best F-scores of 73.55.

Keywords- Arabic NLP; negation; speculation; biomedical (medical and biological); factuality.

## I. Introduction

The huge amount of biomedical data and its growth recently increases the need to analyze by depth. Deep learning is based on the artificial neural network where input is mapped to output according to mathematical manipulation and training to find the probability of assigning a certain result. Deep Neural Networks proved its effectiveness in different fields such as speech recognition[29], computer vision [30] and natural language processing [31]. Deep learning-based algorithms show remarkable performance to extract shallow and deep features, also, to learn the structure from complex biomedical data to make sense [32]. Negation and speculation detection plays a significant role to decide the factuality of biological and medical texts (biomedical) and their confidence. Deep learning is applied for negation and speculation detection and proved its effectiveness and accuracy in different language such as English [14]. Due to the scarcity of datasets for training and testing in the Arabic language, less effort has been made.

***Negation***

In negation, the statement is turned into its opposite by negation cue. Two main types of negation are distinguished: clausal, in which an entire statement is negated and local where a specific part of the statement is negated [1]. According to [2], negation has four categories: denials, imperatives, rejections and questions. Negation has its major effect on several NLP takes such as Text entailment wherein existing negation cue turn the statement to its opposite. For example, S1 does not entail S1:

S1: The weather is very hot

S2: The weather is not hot

In sentiment analysis, negation revere the polarity; for example, S3 has a negative polarity while S4 has a positive polarity:

S3: The iPhone is very expensive

S4: The iPhone is not very expensive, it has many services

For the Arabic language, The most used negative cues:

لا (la), ليس (lys), لا شيء (la shy), لا أحد (la ahd), لا هذا ولا ذاك (la hdha wla dhak),لا مكان (la makan),أبد (abd),الكاد (alkad), نادرا

(naderan), بالكاد (belkad), لم يكن (lam yakon), لا ينبغي (la yanbghey), لن (lan), لم أستطع (lam astata), متعود (mtwad),لا يمكن( la ymkn),

### *Speculation*

Speculation is the degree of uncertainty where there is a lack of information, and a reader could not decide the truth of information [3]. There are four levels that correspond to the degree of uncertainty ranging from low to high [4]:

- ❑ Speculation: e.g. She was probably right.
- ❑ Hedging: e.g. 3% of students failed the exam
- ❑ Investigation: e.g. they examined the contextual representation of the cues.
- ❑ Weaseling: e.g. it is known as a health problem

Speculation is the most important type of uncertainty; the first definition of speculation was presented by [5], as using words (speculation cues) hold fuzzy. Speculation cues are words that tell the author's judgement [6]. In [7], The authors studied speculation language and investigated levels of belief (i.e., hypotheses, tentative conclusions, hedges, and speculations). A list of hedges words and its phenomena words according to the degree of category membership are mentioned like (particularly, somewhat, basically, actually ... etc.) and called them predicate modifiers [5]. There are at least four types of criteria for category membership: Definitional (technically), primary (strictly speaking+), secondary (loosely speaking) and characteristic though incidental (regular) [5]. The authors classified the speculation based on the speculation cues to six categories [8]:

• Auxiliaries - may, might, can, would, should, could …etc.

• Epistemic verbs - suggest, presume, suppose, seem, appear, indicate …etc.

• Epistemic adjectives - probable, possible, likely, unlikely, unsure, not sure … etc.

• Epistemic adverbs - probably, possibly, presumably, perhaps, potentially… etc.

• Epistemic nouns - possibility, probability, hypothesis, suggestion… etc.

• Conjunctions - or, and/or, either ... or, whether ... or, whether… etc.

In Arabic language, the most used speculation cues are:

قد (qd) , ربما(rbma) , يمكن(ymkn) , سوف(swf) , ينبغي(ynbghy) , يمكن (ymkn), توحي(twhy) , افترض(aftrd) , افتراض(aftrad) , احتمل (ahtml) , احتمال (ahtmal) , تشير ( tshyr), تظهر( tzhr) , يبدو(ybdw ), فرضية(frdyt) , اقتراح ( aqtrah), , إما ... أو( ama ... aw) , سواء ... أو( swa ... aw) , سواء ... محتمل( swa ...Mhtml) , ممكن (mmkn), محتمل(mhtml) , غير مرجح(ghyr mrjh) , غير متأكد (ghyr mtakd), غير مؤكد (ghyr moakd) ربما (rbma) , يحتمل (yhtml)

## II. RELATED WORK

Negation: In [9], the authors developed DEEPEN system can solve the problem of Incorrect negation assignment in composite sentences NegEx's system by incorporating the dependency relationship. Dependency relationship is the grammatical relationship between the governor (head) word and dependent word which attached to governor. The results Performance showed reduced false positives in NegEx's system. In [10] the authors analysed the negation in dialogue and proposed sequence labelling model using CRFs to detect the negated fragments not only in the same sentence but also in previous dialogue which has not been done before. They annotated tutorial dialogues to use as a dataset (called DT-Neg corpus). It is available for research purposes. The F-1 score obtained (0.826). In [11] they proposed a system that able to affirm or negate suicide mental health records using probabilistic context free grammars. The algorithm presented better results in information retrieval. In [12] the authors created a new corpus of negation and speculation in the veterinary clinical note domain called VetCompass and defined it's the annotation and trained the CRF model over VetCompass training data. The obtained results showed that in domain training data gives better result than on training on bioscope corpus. In [13] the authors realised that there are drawbacks in previous systems like dedicated to a specific domain and most of them are related to the English language rather than they are highly engineered and depend on a reliable representation for a sentence of parser result. They tried to tackle such problems by developing neural network and word embedding's system. Due to the important role of backward processing (negated token may come before cue), it is shown that using both feed-forward neural with Bidirectional Long Short-Term Memory is better than using feed-forward neural alone. Best results of performance are obtained compared by previously scope systems on the same domain but not a different domain. They outperformed the best result of *Sem shared task 2012. In [14] A Convolutional Neural Network (CNN)-based model with probabilistic weighted average pooling has been recently proposed to handle negation and speculation texts automatically. Each candidate token labelled by A, B or O after before and outside respectively, to describe the location relationship among the cue tokens and other (distance of the cue to the candidate token.). Syntactic paths between the cues and the candidate tokens in both constituency and dependency parse trees are extracted (Path Feature can offer effective features to determine whether a token belongs to the scope.). Position Feature and Path Feature are concatenated to form the one feature vector, which is then fed into a SoftMax layer to compute the confidence scores then labelled it. The evaluation showed that the system achieves the best results on the abstract in BioScope corpus. In [15] In 2018, the authors proposed the first system that has the ability to distinguish the negated part in tutorial dialogues using deep learning methods (Long Short-Term Memory (LSTM). Two types of answer negation are: explicit where the negation cue is present and implicit where the text is negated by context without negation cues. The annotated tutorial dialogues data prepared in [10] used to train the models. The performance results are f-score (0.839) are better than [10] of (0.826). DT-Neg corpus. In [16] the authors proposed NegBio whose code focused on negation. Their system is based on rules defined from dependency graphs (universal dependency graph (UDG)). It gives directed graph (universal dependency graph (UDG)) which is a sketch of grammatical relationships in a sentence could be understood even by non-linguists. For the Arabic language, the importance of contradiction classification in RTE systems is studied in [17].

Speculation: In [18] they used a memory-based system that relies on information from lexical and parse tree dependencies, Depending on a heuristic approximation of nearest neighbour, get the highest performance of the task detection the scope of speculation cues. In [19], A Combination of lexical and syntactic patterns as rule-based system has been used to detect speculation and negation scopes and proved its effectiveness compared to other machine learning based systems on BioScope corpus. In [8] they applied a Conditional Random Fields (CRFs) algorithm to learn models whether a word is speculation cue or not based on provided with linguistic features. In the speculation detection stage, a subtree of phrase parse tree generated by Stanford Parser is selected to be the scope of a cue, depending on a set of linguistic heuristics determined previously. The performance results to detect the sentences that contain cues were better than detecting the scope task. In [20] they used syntactic parse information tree for tree kernel-based negation and speculation scope detection. Compared with recently scope detection systems, the system obtained high an F-score of 76.90% on the BioScope corpus reports. In [ 21] they proposed the first n speculation detection system for official monetary policy statements. They built two data set OF Debates and decisions of U.S. central bank and labelled them for speculation by professional annotators. On the two datasets, numerous rule-based and machine learning (ML) approaches were used to test their value. The classifiers obtained 0.70 F-score on the speculative class). In [ 22] They studied cross-domain automatic speculation detection and showed good performance by training SVM on general domain and applying on new specific-domains specific dictionaries (small in-domain list of speculation triggers) like the monetary policy domain in their research. For the Arabic language, a machine learning model is implemented to get uncertainty cues in addition to their scope and holder, 75,9 % f-score is obtained.

## III. Arabic based Deep learning models for speculation and negation detection

### *Input Representation*

In our training models, each model has an input of constituency and dependency Shortest syntactic path. The shortest syntactic path between the cues and the candidate tokens in constituency and dependency parse are extracted as its effectiveness is proven in [14]. An Arabic example with its constituency and dependency parsed result.

Sentence: قَدْ qadيَكونُ yakwnu العِلاجَ alilaja جِراحيًّا jirahyaa عِنْدَما indama يَصْعُبُ yasubu العِلاجُ alilajuالدَّوائيُّ aldawayyu

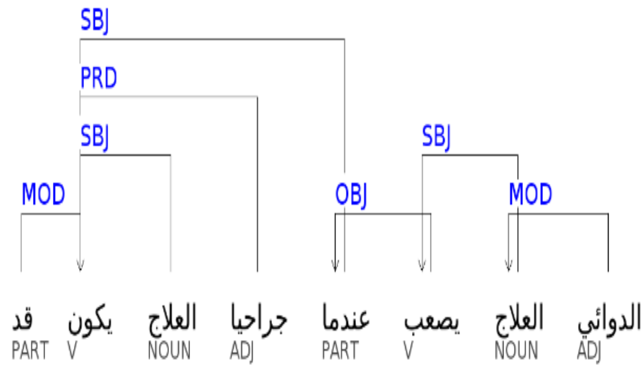


Figure 1: Example of Arabic constituent parsed sentence

```
(ROOT
 (S
  (VP
   (PRT (RP قد))
   (VBP يكون)
   (NP (DTNN العلاج))
   (NP
    (NP (NN جراحيا) (JJ عندما))
    (SBAR
     (S
      (VP (VBP يصعب)
       (NP (DTNN العلاج) (DTJJ الدوائي)))))))))
```

Figure 2: Example of Arabic dependent parsed sentence

### *DNN models in this research*

A Convolutional Neural Network (CNN): CNN-based model has the benefit of fast training to extract a rich representation of syntactic features of speculated and negated elements in a text and deals with sentence by part(windows) [23]. Long Short-Term Memory (LSTM) had memory helps to predict current output based on previous result [24]. Bidirectional Long Short-Term Memory (BLSTM) -based model has the role of extraction semantic features from an entire sequence and learning linguistic relationships over long sequences to obtain a rich representation of statement. Pooling is implemented to get the highest contextual representation of the two statements. The output of the previous stage is fed to BiLSTM to extract sequences of representations [25]. In [26] Attention Based Bidirectional Long Short-Term Memory Networks (AttBLSTM) proposed to detect Relation classification and the necessary information for this detection, without depending on extracted features from lexical resources This approach has the benefit of focusing on relevant words related to the task like speculation and negation cues in our subject, so more able to extract relevant relation to uncertainty cues with related words. The hidden vectors which comes from

LSTM is used in attention mechanism to represent a sentence, following these equations:

$$M= \tanh(H) \quad (1)$$

$$a = \mathrm{softmax}(w^T M) \quad (2)$$

$$r = H*a^T \quad (3)$$

***BioArabic Corpus***

In previous research, a corpus has been built to handle negation and speculative in Arabic. BioArabic corpus consists of 10165 sentences, 26.2% of these sentences have negation and speculative cues, annotated by computational linguistics and biologists. Due to much sentences have much information in English, lead to incorrect produced parsing, they are removed. A filtered corpus of 3000 is used in our training of 31.84% negated and speculated sentences [27].

## IV. RESULTS

The results obtained by our different models over Bio-Arabic corpus are presented in Table 1. All the result reported after applying word embedding in addition to the constituency and dependency parsing to each model. From obtained results, it is observed that CNN has less performance than Recurrent Neural Network based models since it could not capture more information from the previous context of negation or speculation cue. It could be noticed that Bi-LSTM has more ability to get contextual representation by forward and backward learning and has the benefit of learning long sentences. Attention combination with Bi-LSTM has the best performance as an attention mechanism focuses more on relevant information.

TABLE 1: RESULTS OF SPECULATION NEGATION SCOPE DETECTION USING DIFFERENT DEEP LEARNING MODEL

| Model | Precision | Recall | F-score |
|---|---|---|---|
| CNN | 66.07 | 68.93 | 67.46 |
| LSTM | 68.37 | 71.93 | 70.10 |
| CNN+LSTM | 69.07 | 73.53 | 71.23 |
| CNN+pooling+LSTM | 69.97 | 73.87 | 71.86 |
| CNN+pooling+BiLSTM | 70.18 | 75.42 | 72.7 |
| Attention- Bi LSTM | 71.63 | 75.57 | 73.55 |

### ERROR ANALYSIS

After analytical operations, it is noticed that there are two main problems, first: the used dataset has a lot of English medical or biological vocabularies, produced imprecision parsed sentences, second there are a lot of long sentences separated by comma which may be divided into different sentences, to reduce the problem of very long dependencies. The third problem is the main role of applied parsers to get correct labels (where the produced parsed sentences are the input of each model), which mean incorrect parsed input leads to reduce the accuracy of the models.

## V. CONCLUSION

This paper applied different deep learning-based models for speculation and negation scope detection for Arabic models. To my best knowledge, this is the only research focused on speculation and negation detection for biomedical texts to the Arabic language. Due to process more, long dependences and attention mechanism focused on relevant cues, Attention- Bi LSTM model performs best compared to other models, 73.55% f-scor. For our future work, we tend to build different datasets for specific and general domains. We are going to study speculation and negation on event level not only sentence level and enrich more semantic relations between candidate scope words like gloss relation, holonym relation (is a part of another thing), and meronym relation (opposite to holonym) and studied the effect of this detection of speculation and negation on more NLP tasks. The implementation of this paper is available on [28].

## REFERENCES


[1] Klima, Edward S. 1964. Negation in English. In Jerry A.Foder and Jerrold J. Katz, eds., The Structure of language (Englewood Cliffs, New Jersey: Prentice-Hall).

[2] Tottie, Gunnel. "Negation in English speech and writing: A study in variation." Language 69.3 (1993): 590-593.

[3] Szarvas, György, et al. "Cross-genre and cross-domain detection of semantic uncertainty." Computational Linguistics38.2 (2012): 335-367.

[4] Zerva, Chrysoula, et al. "Using uncertainty to link and rank evidence from biomedical literature for model curation." Bioinformatics 33.23 (2017): 3784-3792.

[5] Lakoff, George. "Hedges: A study in meaning criteria and the logic of fuzzy concepts." Contemporary research in philosophical logic and linguistic semantics. Springer, Dordrecht, 1975. 221-271.

[6] Verbeke, Mathias, et al. "Kernel-based logical and relational learning with kLog for hedge cue detection." International Conference on Inductive Logic Programming. Springer, Berlin, Heidelberg, 2011.

[7] Light, Marc, Xin Ying Qiu, and Padmini Srinivasan. "The language of bioscience: Facts, speculations, and statements in between." HLT-NAACL 2004 Workshop: Linking Biological Literature, Ontologies and Databases. 2004.

[8] Yang, Hui, et al. "Speculative requirements: Automatic detection of uncertainty in natural language requirements." 2012 20th IEEE International Requirements Engineering Conference (RE). IEEE, 2012.

[9] Mehrabi, Saeed, et al. "DEEPEN: A negation detection system for clinical text incorporating dependency relation into NegEx." Journal of biomedical informatics 54 (2015): 213-219.

[10] Banjade, Rajendra, Nobal B. Niraula, and Vasile Rus. "Towards Detecting Intra-and Inter-Sentential Negation Scope and Focus in Dialogue." The Twenty-Ninth International Flairs Conference. 2016.

[11] Gkotsis, George, et al. "Don't Let Notes Be Misunderstood: A Negation Detection Method for Assessing Risk of Suicide in Mental Health Records." Proceedings of the Third Workshop on Computational Lingusitics and Clinical Psychology. 2016.

[12] Cheng, Katherine, Timothy Baldwin, and Karin Verspoor. "Automatic Negation and Speculation Detection in Veterinary Clinical Text." Proceedings of the Australasian Language Technology Association Workshop 2017. 2017.

[13] Fancellu, Federico, Adam Lopez, and Bonnie Webber. "Neural networks for negation scope detection." Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vol. 1. 2016.

[14] Qian, Zhong, et al. "Speculation and negation scope detection via convolutional neural networks." Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016.

[15] Gautam, Dipesh, et al. "Long Short Term Memory Based Models for Negation Handling in Tutorial Dialogues." The Thirty-First International Flairs Conference. 2018.

[16] Peng, Yifan, et al. "NegBio: a high-performance tool for negation and uncertainty detection in radiology reports." AMIA Summits on Translational Science Proceedings 2017 (2018): 188.

[17] AL-Khawaldeh, Fatima T. "A Study of the Effect of Resolving Negation and Sentiment Analysis in Recognizing Text Entailment for

Arabic." World of Computer Science & Information Technology Journal 5.7 (2015).

[18] Morante, Roser, Vincent Van Asch, and Walter Daelemans. "Memory-based resolution of in-sentence scopes of hedge cues." Proceedings of the Fourteenth Conference on Computational Natural Language Learning---Shared Task. Association for Computational Linguistics, 2010.

[19] Apostolova, Emilia, Noriko Tomuro, and Dina Demner-Fushman. "Automatic extraction of lexico-syntactic patterns for detection of negation and speculation scopes." Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2. Association for Computational Linguistics, 2011.

[20] Zou, Bowei, Guodong Zhou, and Qiaoming Zhu. "Tree kernel-based negation and speculation scope detection with structured syntactic parse features." Proceedings of the 2013 conference on empirical methods in natural language processing. 2013.

[21] Štajner, Sanja, et al. "Automatic detection of speculation in policy statements." Workshops on Natural Language Processing and Computational Social Science. Association for Computing Machinery. 2016.

[22] Štajner, Sanja, et al. "Domain adaptation for automatic detection of speculative sentences." 2017 IEEE 11th International Conference on Semantic Computing (ICSC). IEEE, 2017.

[23] Kalchbrenner, Nal, Edward Grefenstette, and Phil Blunsom. "A convolutional neural network for modelling sentences." arXiv preprint arXiv:1404.2188 (2014).

[24] Zhang, Dongxu, and Dong Wang. "Relation classification via recurrent neural network." arXiv preprint arXiv:1508.01006(2015).

[25] Wang, Peilu, et al. "A unified tagging solution: Bidirectional lstm recurrent neural network with word embedding." arXiv preprint arXiv:1511.00215 (2015).

[26] Zhou, Peng, et al. "Attention-based bidirectional long short-term memory networks for relation classification." Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Vol. 2. 2016.

[27] AL-Khawaldeh, Fatima T. "Speculation and Negation Annotation for Arabic Biomedical Texts: BioArabic Corpus." World of Computer Science & Information Technology Journal , (2016).

[28] Code : https://github.com/fatimakhawaldeh1/bio-ADNN-

[29] Graves, Alex, Abdel-rahman Mohamed, and Geoffrey Hinton. "Speech recognition with deep recurrent neural networks." 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, 2013.

[30] Shin, Hoo-Chang, et al. "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning." IEEE transactions on medical imaging 35.5 (2016): 1285-1298.

[31] Costa-Jussà, Marta R. "From feature to paradigm: deep learning in machine translation." Journal of Artificial Intelligence Research 61 (2018): 947-974.

[32] Cao, Chensi, et al. "Deep learning and its applications in biomedicine." Genomics, proteomics & bioinformatics 16.1 (2018): 17-32.