



## An Efficient Hybrid Bi-LSTM Attention Model for Claims Extraction from Research Articles through Deep Learning

Salma Asif<sup>1\*</sup> and Aiman Khan<sup>1</sup>

<sup>1</sup> SLOSH AI SOLUTIONS, Islamabad, Pakistan

\*Corresponding author

### Abstract

The research literature is growing rapidly. A research article contains massive amounts of textual information. Claims are the most significant information in a research article that needs to be retrieved to understand the gist of the research work. A research article contains a number of claims in different sections (abstract, introduction, results, and discussion) of an article. A literature review shows that a few studies of claim extraction have been conducted and they are limited to extracting claims from the abstract section of the article only. In existing studies claims are classified either on the basis of the keywords or all the words. In existing works semantics and context are ignored, and Bag of Words (BoW) representation is used. Deep learning architectures such as Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) have the potential to produce better results through the use of deep learning. Attention mechanism and Bidirectional Long Short-Term Memory (Bi-LSTM) have been used for multiple tasks in Natural Language Processing (NLP) and give effective results. In this work, we propose a hybrid Bi-LSTM attention model. Bi-LSTM model captures the long-term sequences of words and the attention mechanism highlights the important words or keywords in text. A number of experiments have been performed on research articles claim and standard IBM datasets. We verified our proposed model on two datasets for claim/non-claim classification and our model gives 96.7% and 95% on both datasets. The results show that our proposed model through deep learning outperforms existing models.

**Keywords:** Claim Extraction, Research Articles, Attention Model, Information Retrieval, Natural Language Processing.

### 1. Introduction

The volume of research literature is growing exponentially rendering it almost impossible to manually sift through it and identify claims in individual works in a reasonable amount of time. It is especially cumbersome and time consuming for funding agencies that have to decide on the fate of the applications for funding in a short span of time. Claims made in research papers is the crux of the work being written about and any future funding decisions are dependent upon claims made in new papers. Claims, backed by evidence, embody the true contribution made by the scientists working on a specific problem [1]. However, depending on the writing style, claims made by the authors are spread all over the place within a research article. Even a diligent reader at times fails to comprehensively identify all claims made within a research article. Blake [2] found that claims are not distributed equally in all sections. They discovered that abstract, introduction,

results, and discussion sections contained 7.84%, 28.56%, 23.44% and 43.15% of claims respectively. It makes the task of claims extraction a bit tricky and necessitates their automatic extraction to save on time [3].

Artificial Intelligence (AI) is a broad field of Natural Language Processing (NLP). Claims extraction can be broadly identified as information retrieval or extraction, which falls under the domain of NLP. Text mining plays a vital role to retrieve relevant information from text documents [4]. Text mining techniques and methodologies can parse a huge amount of text in relatively little time, extracting useful information. To construct a system for textual information retrieval, there are a lot of approaches. One method is to manually define rules or patterns using regular expressions to retrieve information. Blaschke and Valencia [5] designed a *suisegi* system by manually designing patterns that extract information on proteins from biomedical documents or text. Another approach is to automatically learn pattern-based extraction rules to identify the relation or entity type. Machine learning techniques are another method for information extraction or classification. Classifiers have been used to predict the label of a sentence based on a token and its context.

In literature rule-based approaches have also been used for claims extraction or classification [6], [7], [8], [9]. De Ribaupierre[7] annotated each sentence of a document and used syntactic rules to identify the discourse type of every sentence. Risk of annotation noise can increase when there are a large number of rules. Jansen and Kuhn [6] proposed a rule-based approach to help researchers know about recent developments by extracting a core claim from the abstract. They used term frequency (*tf*) for keyword extraction. Sateli and Witte [10] used a rule-based approach to extract claim or contribution sentences from full research articles and they used predefined keywords that point the claims in a sentence. Keywords play an important role in claim extraction.

Biomedicine related articles contain a large number of claims. Claims in the domain are not consistently reliable and might contradict each other. Thus, identification and rectification of contradictory claims play a big role in the improvement of the system. Alamri and Stevenson [11] extracted the contradictory claims from the abstract and proposed systematic reviews related to four cardiovascular topics. Using RNN they extracted different types of sentences from the lawsuit documents and every sentence was annotated with one of the five labels [12]. Every sentence in a document must have discourse type (like definition, hypothesis, method, result etc.). They extracted these types of sentences using a syntactic approach.

In previous work, not all claims were extracted from the full articles and Bag-of-Words (*BoW*) representations were used to assign the weights to words. However, to the best of our knowledge semantic relations between words in the domain of claims extraction has not been explored. Semantics is hidden within the context of words as is popularly known in the NLP community that “A word is known by the company it keeps.” In existing work claims are classified either on the basis of the keywords or all the words, regardless of the importance of the individual words or keywords.

To resolve the issues identified above, we have proposed a hybrid Bi-LSTM attention model and its comparison is done with existing studies [12], [22], [24],[29], [30]. We implemented our model on a standard IBM dataset and on claim sentences of research articles. On both datasets our model performed well.

Main contributions of this study are given below:

- We propose a novel hybrid Bi-LSTM attention model to classify claims from the research articles. Our novel hybrid model improves the performance effectively for claim classification in research articles as compared to previous state-of-the-art methods. To capture the semantics of the word, Word2Vec model is used. We implemented Bi-LSTM to capture long-distance sequences in backward and forward directions.
- Keywords are the most important aspect in text classification, so we introduce the attention mechanism for claims/non-claims classification. The attention mechanism gives high weights to keywords and attention is more suitable to learn the weights of keywords. To give the attention to keywords we embed attention mechanism in neural networks.
- Proposed a rule-based approach with different keyword extraction techniques to build a dataset of research articles claim and non-claim sentences and compared with the existing study. In our study we used these claim and non-claim sentences as dataset and performed experiments.

The rest of the paper is organized as follows. Section 2 presents related work. About corpus the detail is discussed in section 3. The proposed model for claim extraction is presented in section 4. Section 5 presents the experimental setup. The experimental results, comparison with existing studies are presented in Section 6. At the end this study is concluded in section 7.

## 2. Related work

Much effort has gone in the recent past to extract information from research articles. Researchers used machine learning and deep learning approaches for information extraction. In existing studies, authors extracted entities and relation from biomedical research papers using deep learning [13], [14],[15], [16], [17]. Researchers present in [18], [19] a hierarchical neural network model to classify sentences from abstract of biomedical. Wackerbauer [20] used a binary classifier to identify the key sentences or summary of research articles. The comparative sentences identification systems are presented in [21], [22]. Authors identified hypothesis of research from the abstract of biomedical research literature [9].

In a research paper claims are the most important information. However, little work has done for claim extraction from research papers. To retrieve claim from text rule-based and machine learning approaches are used. Blake introduced a Claim Framework, which is an annotation scheme that shows how scientists communicate claims and findings of the empirical study of biomedicine in full-text articles [2] and they defined different types of claims and identified explicit claims. Ahmed et al, identified location and number of claims [23] by Claim Framework introduced by Blake in the domain of social sciences. In studies[13], [24], [25] authors identified different types of claims sentences from text using classifiers. In [9], the authors proposed a method which extracts claims from research articles of the biomedicine field. They are extracting claims from abstract only. For claim extraction authors used the rule-based method and used  $tf$  for keyword extraction. They assigned scores to sentences and extract a single sentence as a claim which has the highest rank. Researchers are identifying Salient factual claims using a neural network model in the study [26]. An idea is introduced by authors [27] of discovering two types of scientific claims: dominant and dominated to annotate them. They introduced the set of features to focus on the claim and its content.

Researchers automate the process of identifying claim sentences from the literature review. Rule-based or machine learning techniques are used for claim extraction. In rule-based approaches, a large number of rules increase the annotation noise that affects performance [7]. The scientific literature contains multiple types of many claims, however, existing studies not extracted all types of claims. In discussed literature review *BoW* representations were used to assign the weights to words, however, they are unable to capture the semantics between words. So, there is a need of representation which captures the semantics. Keywords play an important role in text classification. In existing studies, in rule-based authors manually defined the keywords and in machine learning they used keyword extraction techniques to classify the claims. Later neural networks assign weights to all words, regardless of the importance of the words. There is a need for a model which resolve shortcoming of related work. Table 1 represents the summary of some related work.

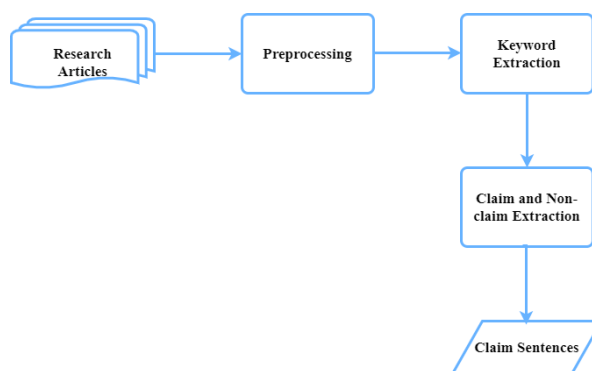
**Table 1** Literature Review Summary

Reference	Paper Name	Proposed Method	Technique	Feature Selection	Dataset	Limitation
[2]	Beyond genes, proteins, and abstracts: Identifying scientific claims from full-text biomedical articles	Introduced a claim framework, which reflects how authors communicate claims in research articles and identify explicit claims	claim extraction by semantics and syntax	N/A	29 full-text articles of biomedicine	Define the different type of claims but identify only explicit claims from biomedicine domain
[6]	Extracting Core Claims from Scientific Articles	Extracting a claim from scientific articles	Rule-based	Term frequency(tf)	125 articles for training and 125 for testing of biomedicine	Extracting single claim from abstract only
[21]	Identifying comparative claim sentences in full-text scientific articles	Identify comparison sentences form full-text articles	SVM, Naïve Bayes, Bayesian Networks	syntactic and semantic features	122 full-text articles of toxicology	Identified only comparison claims
[23]	Identifying claims in social science literature	Used Blake's (2010) claim framework to identify the number of claims and location of the claim	semantics and syntax		8 full-text articles of social science	Only identify location and number of claims

[24]	A Method to Automatically Identify the Results from Journal Articles	Proposed a model which identifying result and non-result sentences from research articles	SVM and Naïve Bayes	MI, CHI, IG strategies used for feature selection	A small corpus of 17 articles	Limited to a single domain and identified results from journals articles only
------	--	---	---------------------	---	-------------------------------	---

### 3. Corpus Formation

For the purposes of this research, we have used two datasets. One is the IBM dataset of claims<sup>1</sup> and the other is our own. Extraction of our own dataset the classification in Fig. 1 as a block diagram, in which the steps performed for sentence extraction has been shown. As a first step, scientific research articles were collected, and preprocessing was performed. The research articles were then given as input to model and keywords were extracted from the text data. Both claim and non-claim sentences were extracted. The classification was done using neural networks and the claims were identified.



**Fig. 1** Block Diagram of Corpus Building

#### Extraction of sentences

Algorithm 1 shows the process claim extraction from research articles to build the dataset. Claims form the gist of what a research article is about. In other words, they summarize the topic. Our work revolves around claim extraction. In our work, has content we extracted the claim sentences. We extracted the sentences on the basis of the following factors:

- match sentence with defined patterns
- keywords extracted by RAKE

**Algorithm 1** Shows the Process of Dataset Preparation from Research Articles

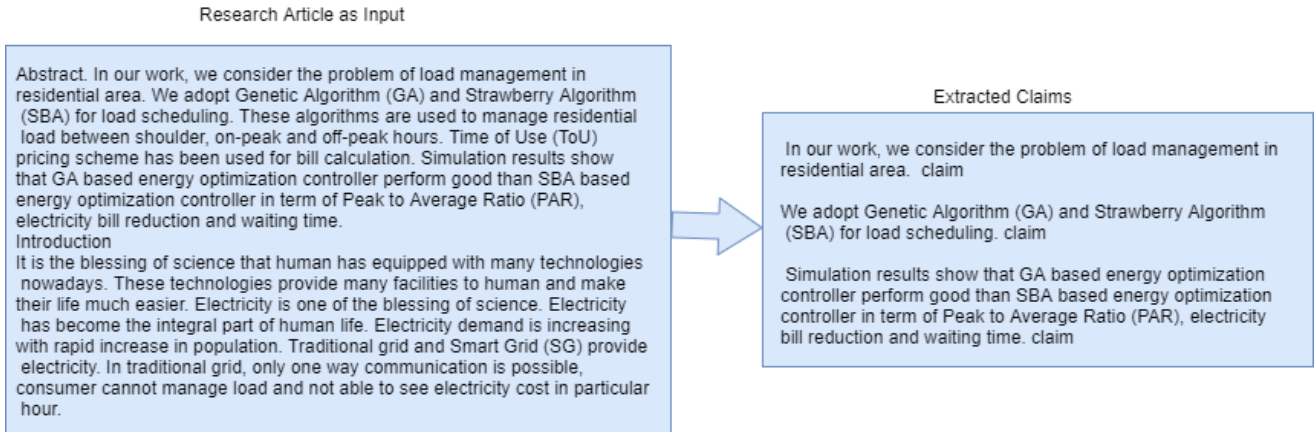
<sup>1</sup>[https://www.research.ibm.com/haifa/dept/vst/debating\\_data.shtml](https://www.research.ibm.com/haifa/dept/vst/debating_data.shtml)

---

**Input:** text documents ( $D$ )  
**Output:** sentences

1. repeat step 2 to 4
2. for each  $d \in D$ 
  - $keywords \leftarrow keyword\_extraction(lines)$
  - for each  $line \in lines$   $\parallel$  reading document line by line
    - if ( $acknowledge, body, reference$  in  $line$ )
      - $start \leftarrow FALSE$
      - if ( $abstract$  in  $line$ )
        - $start \leftarrow TRUE$
    - End if
  - End for
  - $scored\_sentences \leftarrow call\ ranking\ (lines, keywords)$   $\parallel$  call a function which assigns score to sentences
  - $core\_sentence \leftarrow call\ select\_sentence\ (scored\_sentences)$   $\parallel$  call a function which extract a sentence which have highest score
3. End for
4. Return sentences
5. End repeat

---



**Fig. 2** Extracted claims from a research article

Claim sentences usually start with phrases such as "*in this paper, we proposed*", "*our results show*", "*this study reveals*", "*in conclusion*", "*these findings*" etc. A sentence with these types of structure is most probably a claim sentence. To discover these types of sentences we defined the patterns. Fig. 2 shows the extracted claims through our technique when a research article is given as input. We also extracted keywords from the full research article using different keyword extraction techniques. These extracted keywords are also used for sentence extraction.

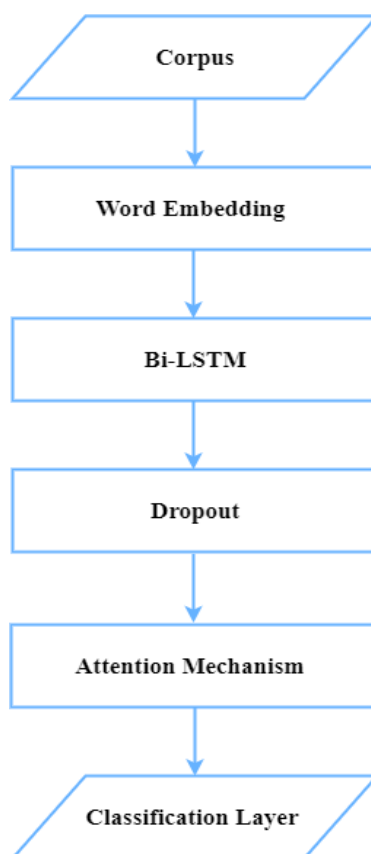
**Table 2** Corpus Building Results Comparison

	Existing ( <i>tf</i> ) [6]	TextRank	RAKE
Extracted claims	69.6%	87.2%	88%

The research articles are given as input and different rules are applied using a regular expression. For keyword extraction, we used RAKE [28]. In a scientific article multiple claims are defined in different sections. In our work, we extracted claim sentences from the full articles. Table 2 shows the results of corpus building experiments, here RAKE outperformed TextRank and *tf* used in existing work [6]. We used these extracted claim and non-claim sentences as a corpus in further work.

#### 4. Proposed method

Our aim is to automatically find claim sentences from research articles, which can be framed as a classification problem. To identify claims, we performed classification using well-known machine learning and deep learning techniques and approaches such as Support Vector Machine (SVM), Naïve Bayes, Logistic Regression, Random Forest, CNN, Bi-LSTM and the attention-based model. Fig. 3 shows the overall flow of our proposed model. It takes the text corpus as an input and performs the pre-processing. In the next step, the model applies word embedding which takes contextual information into account. After that, it captures the long-term dependencies between the words using the Bi-LSTM model. The attention mechanism is used to give attention to important words in a sentence. In the end, the *SoftMax* function is used to perform the classification.



**Fig. 3** Flow diagram of claim classification

#### 4.1 Background of deep learning

##### 4.1.1 Word to vector

Word2Vec is an unsupervised deep learning model. It creates continuous or numeric vector values of words in a sentence. Word2Vec model measures the distance between words on the basis of word meaning. It embeds each word to the high dimensional vector space. It is able to capture the semantics of words. To obtain equal size vectors padding is applied.

### 4.1.2 Long short-term memory (LSTM)

LSTM is a type of Recurrent Neural Network (RNN). RNNs lose some information due to the vanishing gradient problem. LSTM resolves it and performs well for long term dependencies. It captures the information of past time steps only. However, it cannot capture dependencies if they are too long and does not perform well when the sequence length increases beyond 30 [31]. In our model, we used the Bidirectional version of the LSTM which captures the information of past and future time steps. Its context understanding is better than the simple LSTM and has the ability to capture deeper semantics of the words. It reads the input sequence of words  $X = \{x_1, x_2, \dots, x_i\}$  and calculate the forward  $\vec{h}_i$  and  $h_i$  backward sequences of hidden statements. The forward and backward hidden states are concatenated to obtain the representation  $h_t$ . Equation 1 shows the RNN model, where  $f$  is the function,  $h_{t-1}$  is the previous state,  $h_t$  is the new state and  $x_t$  is the input at time  $t$ . Equation 2 represents the concatenation of the forward and backward sequences. Equations 3-6 represent the cell and gates of LSTM. In equations 3-6,  $x_t$  is input at time  $t$ ,  $b$  is the bias,  $W$  is the weight vector of hidden states,  $h_{t-1}$  is previous state and  $I_t, F_t, M_t$  and  $T_o$  denote input gate, forget gate, memory cell and output gate.

$$h_t = f(h_{t-1}, x_t) \quad (1)$$

$$h_t = [\vec{h}_i; h_i] \quad (2)$$

$$I_t = \sigma(w_i \cdot (h_{t-1}), x_t + b_i) \quad (3)$$

$$F_t = \sigma(w_f \cdot (h_{t-1}), x_t + b_f) \quad (4)$$

$$M_t = \tanh(w_m \cdot (h_{t-1}), x_t + b_m) \quad (5)$$

$$T_o = \sigma(w_o \cdot (h_{t-1}), x_t + b_o) \quad (6)$$

### 4.1.3 Attention mechanism

Attention mechanism has been successfully used [32] in many NLP tasks such as sentiment analysis, machine translation, document classification and question answering system. The intuition behind the attention mechanism is to give attention to important words in the text. Fig. 4 shows the process of how to give attention to important words or keywords. First, we calculate the attention weights on the basis of input vectors and sentence level feature vectors. Then we multiply attention weights ( $A$ ) with context vectors ( $c$ ), which are then combined with sentence-level feature vectors to generate the attention vectors ( $h_a$ ). At last, the attention vector is used for output. In equation 8  $h_i$  is word representation of hidden states of Bi-LSTM. We take  $h_i$  hidden vectors as input. The equation 8 is used to calculate the scores of input and target vectors. Equation 9 is used to normalize the scores by using SoftMax activation function. The equation 10 is used to derive the context vector  $c$  that captures the input information to predict the  $y_t$ . To produce the attention vector  $h_a$ , we concatenate both the vectors  $c$  and  $h_t$ . After extracting the attention vector, the vector is fed to SoftMax for classification that produces the output as shown in equation 12.

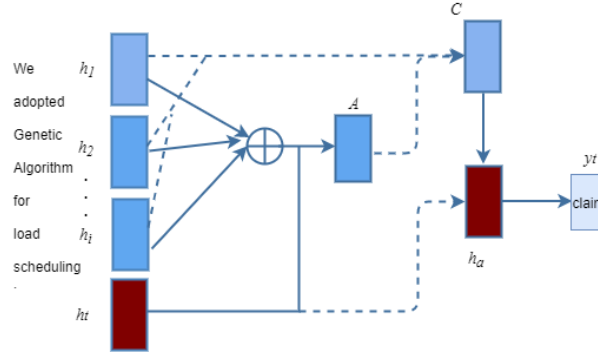
$$a_t = (h | t^T \cdot h_i) \quad (8)$$

$$A = \text{softMax}(a_t) \quad (9)$$

$$c = \sum_{i=1}^i (A_i \cdot h_i) \quad (10)$$

$$h_a = [c; h_t] \quad (11)$$

$$y_t = \text{softMax}(h_a) \quad (12)$$



**Fig. 4** Attention mechanism

#### 4.1.4 Convolutional neural network (CNN)

The CNN model is a type of neural networks. It is adopted from the image processing field, however, it performed well in NLP field. A CNN model is presented in Fig. 5. It is a combination of convolutional and pooling layers followed by a fully connected layer. CNN has performed well on different NLP tasks. It extracts high-level features from the text. The convolutional layer is used to capture the dependencies. Pooling layers are applied to extract important features or information and it reduces the computational power. The fully connected layer is applied to perform classification. Equation 13 is used to calculate the number of output features for each dimension in CNN. CNN model does not capture the information of both directions. To capture the backward and forward sequences we implemented Bi-LSTM with attention.

$$N_0 = \frac{N_i + 2P - k}{s} + 1 \quad (13)$$

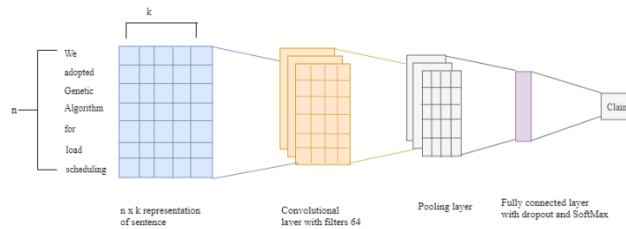
$N_0 = \text{output features}$

$N_i = \text{input feature}$

$P = \text{padding}$

$k = \text{kernel}$

$s = \text{stride}$



**Fig. 5** The architecture of CNN

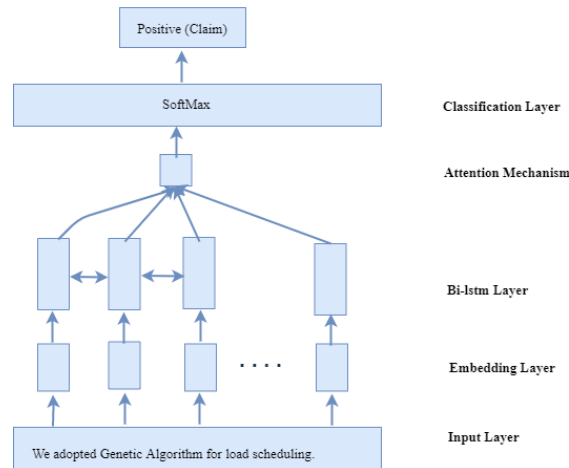
#### 4.2 Proposed hybrid Bi-LSTM-attention model

The whole formation of architecture is represented in equations 1-12 in the above section. Our proposed model resolved above mentioned issues. Firstly, our model captures the semantics of words using Word2Vec method to resolve the issue of *BoW*. For better understanding of context and capture the information in both directions we implemented Bi-LSTM. All words cannot be considered equal, to give high weights to important words used attention mechanism.

The architecture of the proposed hybrid Bi-LSTM attention model with all layers of the model is presented in Fig. 6.



1. **Input Layer:** feeds the sentences to the model.
2. **Embedding Layer:** this layer maps each word of a sentence into a high dimensional vector that captures approximate semantics of a word
3. **Bi-LSTM:** this layer gets high-level features from the embedding layer. We used dropout to prevent our model from overfitting. It discards the unnecessary information which does not enhance the performance of the model. Bi-LSTM generates a sentence level feature vector.
4. **Attention Mechanism:** finds the attention weight vector and merges it with the context vector and produces the output vector. For claim classification, existing methods either classify on the basis of keywords or treat all words equally, regardless of their importance. Attention resolves this issue. Attention mechanism assigns weights to important words on the basis of input vectors and sentence level feature vectors generated by Bi-LSTM. Attention mechanism outputs an attention vector, which is fed to the next layer for classification.
5. **Output Layer:** attention vectors are used for claim classification. *SoftMax* is used in the last layer, which gives the results in the form of 0,1.



**Fig. 6** Architecture of proposed hybrid Bi-LSTM attention model

## 5. Experimental setup

In this section, we provide the experimental setup details for this study. We performed the experiments on the windows 7, processor core m3-7Y30 1.61 GHz, RAM 8.00GB and hard disk 1024 GB. Experiments were performed in Spyder (python 3.6) in this work. Experiments were performed on our own research articles dataset and IBM claim dataset. We split the dataset in 70:30, 70% for training and 30% for testing. Several experiments were performed using deep learning and machine learning techniques such as SVM, Naïve Bayes, Logistic Regression, Random forest, Ensemble learning, CNN and RNN. Our proposed hybrid Bi-LSTM attention model achieve better results with high precision, accuracy, and recall as compared to traditional existing techniques or methods. Table 3 shows the tuning parameters used in our proposed hybrid Bi-LSTM attention model. When training a neural network model, it is required to take a lot of decisions about these tuning parameters. These parameters have a lot of impact on model working.

**Table 3** Proposed hybrid Bi-LSTM attention model parameters

Tuning Parameters	Bi-LSTM Attention
Batch size	100
Filters	64
Epochs	70
Embedding size	32
Learning rate	0.01
Dropout	0.3

### 5.1 Evaluation model

Here, we discuss the evaluation method which is used to evaluate the results of experiments in this study. In our model we used Adam optimizer to measure accuracy and used standard evaluation metrics (precision, recall) to measure the performance. Recall and precision formulas are presented in equations 14 and 15.

$$Recall = \frac{Truepositive(TP)}{Truepositive(TP)+Falsenegative(FN)} \quad (14)$$

$$Precision = \frac{Truepositive(TP)}{Truepositive(TP)+Falsepositive(FP)} \quad (15)$$

## 6. Result and discussion

In this section, we discuss the results of experiments that we performed by implementing approaches of existing papers and our proposed model. We perform a comparison of the results produced by all approaches.

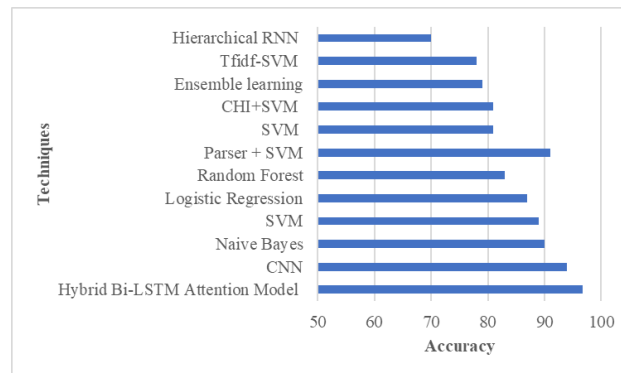
### 6.1 Claim classification

In this study, we proposed a hybrid Bi-LSTM attention model for claim extraction by implementing deep learning Bi-LSTM and attention mechanism. In the proposed model a word2vec technique is used to convert words into vector form and apply Bi-LSTM on vectors to extract the features. Bi-LSTM captures the sequences in both forward and backward directions. Bi-LSTM captures the long-term dependencies in a better way as compared to traditional techniques. The attention mechanism is implemented to give high weights to important words. Due to these mentioned advantages our proposed hybrid Bi-LSTM attention model perform well than other implemented approaches. We implemented the existing techniques and provided a comparison. A number of experiments are performed by applying many machine learning and deep learning techniques. For comparison we applied deep learning CNN architecture. We applied our proposed hybrid model on both datasets that comprised of many claim and non-claim sentences and our proposed model performed well on both datasets. Our proposed hybrid Bi-LSTM attention model outperformed other techniques. We compared our proposed model with existing approaches [12], [22], [24], [29], [30] in Table 4. These approaches are used to extract claims sentences from research articles or law documents. We used two datasets to evaluate our proposed hybrid Bi-LSTM attention model. In this study, CNN performance is comparable, however, multiple convolutional layers are used to capture the long-term dependencies. So, the model becomes very deep and complex.

**Table 4** Claim classification results comparison

	Accuracy	Precision	Recall
<b>Proposed Hybrid Bi-LSTM attention Model</b>	<b>96.7</b>	<b>91</b>	<b>88</b>
CNN	94	85	89
Naive Bayes	90	86	88
SVM	89	88	65
Logistic Regression	87	90	60
Random Forest	83	89	50
SVM [22]	81	77	80
CHI+SVM [24]	81	79	86
Ensemble learning [29]	79	73	81
Tfidf-SVM [31]	78	81	84
Hierarchical RNN [12]	70	48	91

Results show that in comparison with CNN and other machine learning approaches the proposed hybrid Bi-LSTM attention model outperformed. Fig. 7 presents the accuracy comparison of implemented approaches in graphical representation. Our model accuracy is high because our model has the ability to capture sequence in both direction and attention give a high score to the important words produced by the hidden states of Bi-LSTM. Keywords play an important role in text classification. Attention model gives the concept of the keywords in deep learning. Traditional models assign weights to all words and perform classification on the bases of these weights. However, attention gives high weights to keywords and other words have fewer weights. This increase the accuracy of the model. Traditional machine learning approaches used *BoW* technique for word representation. *BoW* technique does not care about the semantics of words. It also does not respect the order of the words in a sentence. Machine learning techniques do not have the ability to capture long-term dependencies in both directions. All traditional machine learning classifiers performance is average in term of accuracy.

**Fig. 7** All implemented techniques accuracy comparison

The CNN have produced 94% accuracy, 85% precision and 89% recall, however the proposed Bi-LSTM with attention model have produced 96.7%, 91% and 88% accuracy, precision and recall respectively. We verified our proposed hybrid Bi-LSTM attention model on IBM and research articles claims dataset. Table 5 shows the result of our model on both datasets in terms of accuracy, recall and precision. On both datasets our model performance is good. Our model gives 96.7% accuracy on the dataset of research articles built by us and 95% accuracy on IBM dataset.

**Table 5** The proposed hybrid Bi-LSTM attention model performance on two datasets

Dataset	Accuracy	Precision	Recall
Research Articles	96.7	91	88
IBM	95	89	90

## 7. Conclusion and Future Work

Automatic extraction and classification of required information such as claims from research articles is an important task. In this study, for claim classification, we proposed a hybrid Bi-LSTM attention model. In our model to captures semantics deeply we implemented Word2Vec and to capture long distance-sequences we applied Bi-LSTM. To learn the high weights of keywords in neural networks we applied attention. We evaluated our proposed model on two data sets such as our own data set of research article claim sentences and IBM dataset of claim sentences. On both datasets our model provided 96.7% and 95% accuracy. We compared our results with existing approaches used for claims extraction. Our model achieves high accuracy as compared to existing techniques used in the literature. All results show that our proposed hybrid Bi-LSTM attention model outperforms other state-of-the-art claim classification techniques in term of accuracy.

In this study, we performed claim extraction in the research articles written in the English language. In future, the articles are written in other languages such as Chinese, Urdu, French can be explored for claim extraction. More deep learning models can be investigated.

## References

- [1] K. J. Mayberry, *Everyday arguments: A guide to writing and reading effective arguments*, 3rd ed. Boston, MA: Houghton Mifflin, 2008.
- [2] C. Blake, "Beyond genes, proteins, and abstracts: Identifying scientific claims from full-text biomedical articles," *J. Biomed. Inform.*, vol. 43, no. 2, pp. 173–189, 2010.
- [3] E. S. Tellez, D. Moctezuma, S. Miranda-Jiménez, and M. Graff, "An automated text categorization framework based on hyperparameter optimization," *Knowl. Based Syst.*, vol. 149, pp. 110–123, 2018.
- [4] C. C. Aggarwal and C. Zhai, "An introduction to text mining," in *Mining Text Data*, Boston, MA: Springer US, 2012, pp. 1–10.
- [5] C. Blaschke and A. Valencia, "The frame-based module of the SUISEKI information extraction system," *IEEE Intell. Syst.*, vol. 17, no. 2, pp. 14–20, 2002.
- [6] T. Jansen and T. Kuhn, "Extracting core claims from scientific articles," in *Communications in Computer and Information Science*, Cham: Springer International Publishing, 2017, pp. 32–46.
- [7] H. de Ribaupierre and G. Falquet, "Extracting discourse elements and annotating scientific documents using the SciAnnotDoc model: a use case in gender documents," *Int. j. digit. libr.*, vol. 19, no. 2–3, pp. 271–286, 2018.
- [8] S. Ribeiro, J. Yao, and D. A. Rezende, "Discovering IMRaD structure with different classifiers," in *2018 IEEE International Conference on Big Knowledge (ICBK)*, 2018.
- [9] M. Shardlow, R. Batista-Navarro, P. Thompson, R. Nawaz, J. McNaught, and S. Ananiadou, "Identification of research hypotheses and new knowledge from scientific literature," *BMC Med. Inform. Decis. Mak.*, vol. 18, no. 1, p. 46, 2018.
- [10] B. Sateli and R. Witte, "Semantic representation of scientific literature: bringing claims, contributions and named entities onto the Linked Open Data cloud," *PeerJ Comput. Sci.*, vol. 1, no. e37, p. e37, 2015.
- [11] A. Alamri and M. Stevenson, "A corpus of potentially contradictory research claims from cardiovascular research abstracts," *J. Biomed. Semantics*, vol. 7, no. 1, p. 36, 2016.
- [12] X. Rao and Z. Ke, "Hierarchical RNN for information extraction from lawsuit documents," *arXiv [cs.CL]*, 2018.
- [13] Z. Yin *et al.*, "Relation classification in scientific papers based on convolutional neural network," in *Natural Language Processing and Chinese Computing*, Cham: Springer International Publishing, 2019, pp. 242–253.
- [14] D. Sousa, A. Lamurias, and F. M. Couto, "Using neural networks for relation extraction from biomedical literature," *Methods Mol. Biol.*, vol. 2190, pp. 289–305, 2021.
- [15] F. Li, M. Zhang, G. Fu, and D. Ji, "A neural joint model for entity and relation extraction from biomedical text," *BMC Bioinformatics*, vol. 18, no. 1, p. 198, 2017.
- [16] Y. Zhang *et al.*, "Neural network-based approaches for biomedical relation classification: A review," *J. Biomed. Inform.*, vol. 99, no. 103294, p. 103294, 2019.
- [17] Y. Zhang *et al.*, "A hybrid model based on neural networks for biomedical relation extraction," *J. Biomed. Inform.*, vol. 81, pp. 83–92, 2018.

- [18] X. Jiang, B. Zhang, Y. Ye, and Z. Liu, "A Hierarchical Model with Recurrent Convolutional Neural Networks for Sequential Sentence Classification," in *CCF International Conference on Natural Language Processing and Chinese Computing*, Springer, 2019, pp. 78–89.
- [19] D. Jin and P. Szolovits, "Hierarchical neural networks for sequential sentence classification in medical scientific abstracts," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018.
- [20] M. Wackerbauer, "Automatically Identifying Key Sentences in Biomedical Abstracts Using Semi-Supervised Learning," Humboldt-Universität zu Berlin Mathematisch-Naturwissenschaftliche Fakultät II Institut für Informatik, 2017.
- [21] D.H. Park and C. Blake, "Identifying comparative claim sentences in full-text scientific articles," in *Proceedings of the workshop on detecting structure in scholarly discourse*, Association for Computational Linguistics, 2012, pp. 1–9.
- [22] S. Gupta, A. S. M. A. Mahmood, K. Ross, C. Wu, and K. Vijay-Shanker, "Identifying Comparative Structures in Biomedical Text," in *BioNLP 2017*, 2017.
- [23] S. Ahmed, C. Blake, K. Williams, N. Lenstra and Q. Liu, *Identifying claims in social science literature*. In Proceedings of the iConference, Fort Worth, USA: iSchools, 2013.
- [24] H. A. Gabb, A. Lucic and C. Blake, *A method to automatically identify the results from journal articles*. iConference 2015 Proceedings: iSchools, 2015.
- [25] W. Ma, W. Chao, Z. Luo, and X. Jiang, "Claim Retrieval in Twitter," in *Web Information Systems Engineering – WISE 2018*, Cham: Springer International Publishing, 2018, pp. 297–307.
- [26] D. Jimenez and C. Li, "An empirical study on identifying sentences with salient factual statements," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018.
- [27] J. M. G. Pinto and W. Balke, "Scientific Claims Characterization for Claim-Based Analysis in Digital Libraries," in *International Conference on Theory and Practice of Digital Libraries*, Springer, 2018, pp. 257–269.
- [28] S. Siddiqi and A. Sharan, "Keyword and keyphrase extraction techniques: A literature review," *Int. J. Comput. Appl.*, vol. 109, no. 2, pp. 18–23, 2015.
- [29] M. Hussain and S. Lee, "Information extraction from clinical practice guidelines: A step towards guidelines adherence," in *Advances in Intelligent Systems and Computing*, Cham: Springer International Publishing, 2019, pp. 1029–1036.
- [30] S. Priyanta, S. Hartati, A. Harjoko and R. Wardoyo, "Comparison of sentence subjectivity classification methods in Indonesian News," *International Journal of Computer Science and Information Security*, vol. 14, no. 5, p. 407, 2016.
- [31] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv [cs.CL]*, 2014.
- [32] D. Hu, "An introductory survey on attention mechanisms in NLP problems," in *Advances in Intelligent Systems and Computing*, Cham: Springer International Publishing, 2020, pp. 432–448.