An Investigate of the Efficacy Of CNN and RNN Models for Extraction of Aspect-Based Sentiments

Usha G R[1] and Dr. J V Gorabal[2]

[1]Department of Information Science and Engineering, Sri Dharmasthala Manjunatheshwara Institute of Technology, Ujire – 574 240 and affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India
ushagr2@gmail.com

[2]Professor, Department of Computer Science and Engineering, ATME College of Engineering, Mysuru, Karnataka, India, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India
jvgorabal@gmail.com

Abstract. Several architectures and approaches have been used to investigate sentiment analysis in recent years. In this communication, The aspect extraction experiment is carried out using different sequence labeling and in combination with the LSTMs. In this research, different pre-trained word embedding techniques used are Glove, Word2vec, FastText. Every individual and combination of different techniques were simulated till we get steady results. Data used in this experiment is Semeval-2014. The added CRF layer over the neural network architecture significantly improved the performance of all models. Finally, an exhaustive analysis is carried out by combining sequence models and different word embedding methods in association with the LSTM architecture to understand the effectiveness of the particular model. In all the models least F1 score is realized using CRF, and LSTM word2vec as 75.45%, and 77.91%, consequently LSTM CRF, BiLSTM CRF with Glove.840B exhibit 84.23%, 85.06% and BilSTM CRF FastText show 85.01%. The experimental findings reveal that a combination of CNN and RNN models with proper word embedding techniques provides optimum results for sentiment analysis.

Keywords: Aspect extraction, CNN, RNN, LSTM, sequence models, word embedding

1 Introduction

The identification and analysis of sentiments, judgements, and attitudes on individual perspectives, feelings, and qualities regarding goods, services, politics, a person, etc. have become increasingly essential in recent years thanks in large part to natural language processing. this process is termed sentiment analysis. The inception and rapid growth of technology in natural language processing make life easier, for example, today if one would like to buy consumer goods, one may not take opinions of dear ones because there have been numerous user reviews and debates on the products in the website. Similarly, for an institute's opinion polls and focus group, interviews and data collection is no more essential because this information is open to the public in abundance on the web. In the last decade, it is also observed that development in business is based on conceited public opinion in social media. Therefore, our social and political structures were significantly impacted by the emotions and feelings expressed on online forums. Such sentiments have rallied crowds worldwide for political changes. Thus, it is very necessary to gather and study the opinions [1].

However, due to the abundance of websites, it still takes a great deal of work to find, assess, and summarise the information they provide. Normally each website has a great number of opinions, which is not always easy to uncover in big blogs and forums. It is difficult for common people to retrieve and categorize opinions. Therefore, the aforementioned scenarios demand automated sentiment analysis. The sentiment analysis can be carried out on numerous levels and complexities. More generally, sentiment analysis is categorized as document level, sentence level, entity, and aspect levels as shown in Figure. 1 [2-3].

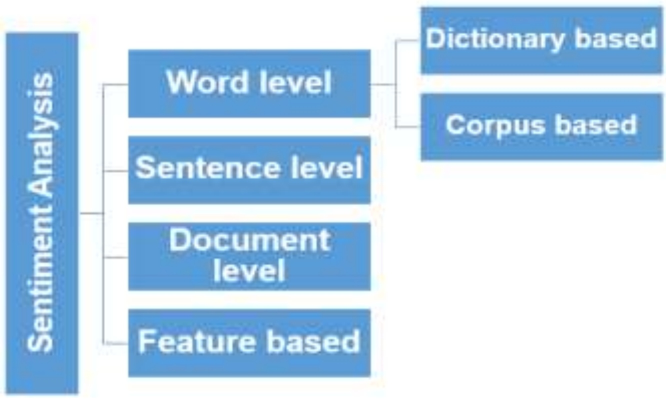

Figure 1. Classification of sentiment analysis [2-3].

The sentiment analysis at the document level is used to assess whether the entire document communicates a favorable or unfavorable view. The most basic method of sentiment classification is document-level sentiment examination. The objective implies that every document conveys a single point of view about a single entity. This assertion, unfortunately, doesn't really hold true when several entities are assessed in a single document. As a result, a more thorough examination is necessary. The aim of sentiment analysis at the sentence level is to assess whether a statement exhibits a good, unfavorable, or neutral attitude toward an entity. This level of sentiment analysis has a neutral opinion, which does not exist in the previous approach (document level). A neutral perspective is just one where there is no opinion stated in a phrase. The assertion used in this approach is one entity is conveyed throughout the sentence, similar to the one at the document level.

The word-level Sentiment Analysis is the challenge of giving a sentiment score to each phrase in an assessment while taking into account the contextual effect of the other terms. For example, consider the word “achievable”, which defines positivity as something that can do or can be accomplished. Nevertheless, “not achievable” define negativity by appending the phrase not. The word level is further classified into a dictionary and corpus-based sentiment evaluation. When a complete text or sentence points to a single entity, the first two approaches (document and sentence) are pretty successful. People, on the other hand, frequently talk about entities having a variety of characteristics. People have differing viewpoints on each element. Therefore, feature or aspect-based sentiment assessment is a fine grain approach. The assessment of sentiments involves four components as Sentiment, the target entity against the feeling is addressed, the person communicating the opinions, and the moment when the emotion is conveyed [2]. Many approaches concentrate on the combination of sentiment and target entity. So identifying the elements and specific aspects categorised as neutral, positive, and negative is necessary for understanding the accurate thoughts offered by the person.

The literature in this field exhibits various techniques to perform sentiment analysis such as machine learning approaches, lexicon-based approaches, and a combination of different approaches as hybrid approaches. Each approach is further classified into different techniques like supervised learning, unsupervised learning, dictionary-based, and corpus-based techniques [4]. In Aspect Term Extraction, A growing emphasis has been placed on neural network architectures and sequential labeling frameworks. These techniques have been demonstrated to be effective for tasks such as named entity recognition (NER), part-of-speech tagging (POS), and character segmentation. Although some studies propose machine-learning models for extracting aspect phrases, there is still a gap in their accurate and full evaluation, with only one or two-word embedding. Making decisions based on a single performance parameter is not always sufficient. In this paper, an attempt is made to solve the aforementioned issues. Three different word embedding techniques glove, fasttext, and word2vec are used in combination with recursive neural network and sequence labeling.

2. Related Work

Sentiment analysis, also widely recognized as opinion mining [5], is an important area of research in Natural Language Processing (NLP). Aspect-level sentimental analysis is a fine-grained classification method. The focus of early research in this area was mostly on collecting characteristics, such as those from emotion lexicons and bag-of-words, in order to create a sentiment classifier [6]. Among these techniques are rule-based [7] and statistical approaches [8] correspondingly and are time and manpower-consuming processes. Deep neural network (DNN) approaches have gained popularity in recent years and DNNs are broadly categorized as convolutional, recursive, recurrent neural network respectively, and memory networks as depicted in Figure 2. The increased significance is due to their ability to create dense vectors of phrases without the need for handmade characteristics [9]. These sequences are weak word characterizations with significant semantic content. In addition, by using the attention mechanism the representation of phrases can be improved with a focus on the key part of a particular sentence[10-11].

In recent years, sentiment analysis has received a lot of attention. Text attitudes are evaluated and gathered in this field. In research, there are a few sub-areas that are relevant. The reference [12] focuses on sentiment analysis at the aspect level (product feature). The noun phrases in the sentences are the characteristics of the products. In order to determine the features of the entities, it is important to identify the aim and gather emotions on entities. The study's thorough summary is presented in such a way that tremendous progress has already been made in identifying the sentiment's target. Aspect detection and extraction are the foundations of modern solutions. A performance report and evaluation of the data sets are provided in a though research. A variety of existing techniques try to standardize the evaluation methodology by using shared data values. Future studies will focus on sentiment analysis, which will primarily focus on aspects -of centric reputation of online items. One of the most significant subtasks in aspect-based sentiment analysis is extracting aspect terms. Previous research has indicated that adopting dependency tree structure representation for this job is promising. Most dependency tree topologies, on the other hand, only have one directed propagation. To extract dependency structure characteristics from the provided words, first present a unique bidirectional dependency tree network in [13].

The important concept is to explicitly include both bottom-up and top-down representations on the provided dependency syntactic tree. The embedded representations are then combined with BiLSTM + CRF to train both tree-structured and sequential features in order to address the aspect term extraction problem. On four benchmark SemEval datasets, experimental results show that the proposed model outperforms state-of-the-art baseline models. The challenge of extracting named entities and their associated sentiment information in a combined way is the subject of [14]. The major finding in entity-level sentiment analysis (also known as targeted sentiment analysis) work is that each named entity is contained within a sentiment scope, which substantially determines the sentiment information associated with the entity. Such sentiment scopes, on the other hand, are rarely clearly labeled the data, and their durations are often unlimited. As a result, the article presents a unique technique that can directly describe the latent sentiment scopes, as opposed to previous algorithms that portray this problem as a straightforward sequence labeling assignment. It demonstrates that on common datasets proposed method outperforms previous techniques based on conditional random fields (CRFs) and more recent work based on neural networks.

Aspect-based sentiment analysis (ABSA) attempts to anticipate a document's polarity in relation to a certain aspect object in [15]. Although neural network designs have proven effective in predicting the overall polarity of phrases, aspect-specific sentiment analysis is still a work in progress. The article offers a unique approach for incorporating aspect information into the neural model in this work. More precisely, the construct word-aspect interactions to include aspect information into the neural model. Authors exhibit a unique model, Aspect Fusion LSTM (AF-LSTM), which learns to attend based on associative connections between sentence words and aspects, allowing the model to focus on the proper words in response to an aspect term. Other state-of-the-art methods that use naïve concatenations to represent word-aspect similarity suffer from this problem. Instead, they use circular convolution and circular correlation to describe the similarity between aspects and words, which are elegantly incorporated into a differentiable neural attention framework. Finally, the proposed model is differentiable from beginning to end and has a strong relationship with convolution-correlation (holographic-like) memories. The presented neural model outperforms ATAE-LSTM by 4% to 5% on average across several datasets, achieving state-of-the-art performance on benchmark datasets. In [16], the authors investigate the dynamic nature of residents' subjective evaluations and corresponding attitudes to the event, based on a systematic assessment of how residents' trust in government(s) and attachment to a marquee event influence their evaluations of the event's impacts and subsequent attitudes toward the event's hosting. The data clearly indicate that citizens' faith in government(s), connection to the event, perceptions of the event's consequences, and eventual support for the event have evolved in a predictable manner over time, in keeping with the confirmation bias hypothesis. Furthermore, studies show that an individual's direct experience with the event affects the relationships between their cognitive/affective assessments and attitudes regarding the event, with a shift in attention to cognitive evaluations following the event.

Many natural language processing tasks benefit from representing words as real-value vectors and using them as input to deep neural networks. Currently, several research employs character-based vectors as a lower-level representation. For the topic of aspect-based sentiment analysis reference [17] discusses how to combine multiple representations of the input. This will propose a model that combines several Convolutional Neural Networks (CNNs), with one CNN handling each distinct representation of the input. The focus on three types of representation in this paper: word embeddings from two techniques (Word2Vec and GloVe), as well as one-hot character vectors.

Jiang et al. [18] will be the first to highlight the relevance of targets based on Twitter sentiment analysis, and they also show that 40 percent of categorization mistakes in classical classifiers are driven by their inability to include the target's characteristics. They demonstrate that by including target-dependent characteristics into standard classifiers, they can outperform target-independent classifiers. The support vector machine is used as a classifier for sentiment analysis by Kiritchenko et al. [19]. It is based on the surface, lexicon, and parse characteristics and achieves adequate accuracy. Due to its discrete property, this method will not provide high accuracy. The [20] proposes phrase latent Dirichlet allocation as a probabilistic generating model. It is used to alleviate the difficulty posed by relying solely on latent Dirichlet distribution by neglecting the order of the words.

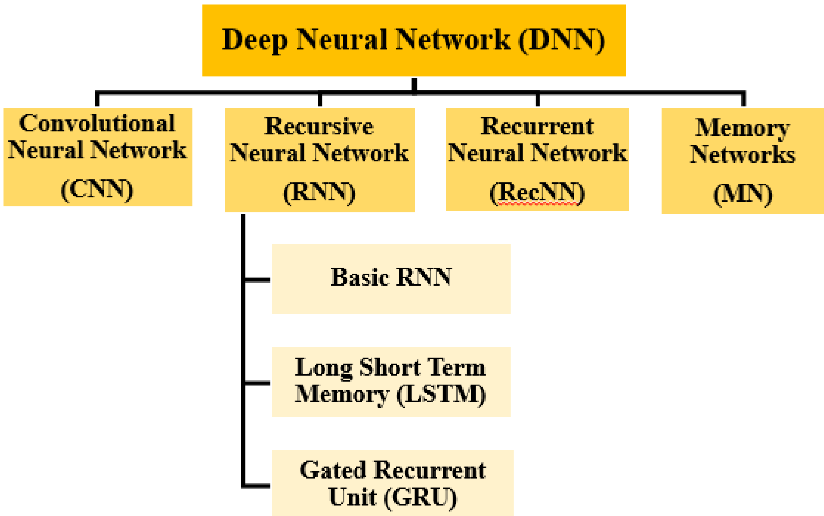

Figure 2. Classification of DNNs.

ATAE-LSTM, a hybrid LSTM, has a mechanism of attention, as suggested by Wang et al. [21]. The model creates aspects that help to calculate the weights of attention. Chen et al. [22] propose RAM, which uses multi-core technology to build a memory using two-way LSTM. Ma et al.[23] establish a framework with a two-way attention mechanism, which dynamically focuses on context and aspect words. Song et al. [24] introduced a network of attention encoders that prevent repetition to model relations across contexts and aspects. However, these attention works represent each component individually in one phrase, which may result in some sentiment dependence information being lost in the case of many aspects. The attention mechanism is employed to emphasize aspect-related words while downplaying aspect-irrelevant ones.

Jakob and Gurevych [25] presented the first of several techniques to extract aspects via sequence tagging. They have built a CRF model based on features such as token knowledge, POS, short dependence path, phrase length, and sentence details. Toh and Wang[26] extended this task at SemEval 2014 with additional handcrafted functions in the DLIREC system. Aspect extraction, on the other hand, does not employ sequence tagging techniques as frequently as Named Entity Recognition tasks [27-28]. Lample et al. [27] suggested a bidirectional LSTM-based neural architecture with a conditional random field. Ma and Hovym [28] proposed a bidirectional LSTM, CNN, and CRF neural network design. As a result, we find various techniques to sequence tagging in NER but few implementations in aspects extraction.

The progress in this field reveals the usage of the word and character embedding with deep learning techniques to extract emotions and opinions. Word embedding is a text vectorization approach that converts words in a lexicon to constant scalar vectors. It should be noted every word length in the embedding vector reflects the hidden characteristic of a word. These vectors were found to encode language patterns and symmetries. Word2Vec [29] is the very earliest and most well-known word embedding technique. The whole computer vision model incorporates two methods: the Continuous Bag-of-Words (CBOW) framework and the Skip-gram model (SG). Global Vector (GloVe) [30] is another outstanding word embedding technique that is learned using a global word-word co-occurrence matrix. FastText [31] is a third commonly utilized method. It is built on the Skip-gram framework, which represents every word as a packet of character n-grams. One may use this technique to construct a representation of words for terms that did not exist in the training data. Furthermore, word embedding Is extended to the character level embedding. This type of embedding was being proven to be beneficial for code-mixed languages and for dealing with the out-of-vocabulary (OOV) issues in applications such as part-of-speech (POS) identification, speech recognition [32], vocabulary test [33], and learning-based identification [27]. One of the earliest techniques for sentiment analysis using char embedding using convolution networks was reported in [34].

The proposed model achieves state-of-the-art performance in aspect category identification and sentiment classification tasks, according to our experiments.

3. Models

RNN may employ arbitrarily lengthy sequenced information in principle, but only a few earlier stages can be examined in practice. The variable length of input must be processed recurrently or recursively since the number of input layers in a neural network is fixed. The RNN does this by separating the variable length of input into tiny bits of similar length, which are then fed into the network. Subsequently, RNN is fed one word at a time until the last word in the statement and produces comparable output. The fundamental RNN model is as depicted in Figure 3.

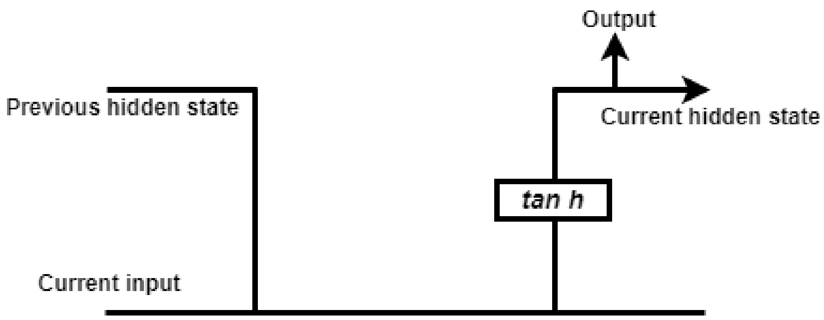

Figure 3. Fundamental RNN model

Long-term short memory networks - often known as "LSTMs" - are unique recursive neural networks (RNNs) that are smart enough to learn long-term dependencies. Hochreiter & Schmidhuber [35] presented LSTM networks, and many individuals were polished and popularised in their work. It works effectively on a wide range of issues and is now extensively utilized. The RNNs are the concepts used to solve the present task by providing the information of the previous task based on user applications. At certain applications, past information may not be required on the contrary RNNs may depend on past information more intently for example "In Japan, I grew up... I'm proficient in English." Recent data proves that the following term is presumably the name of a language, but we also need the framework of France from farther away if we are to narrow down which language. In this kind of situation, inevitably, the gap between pertinent data and the moment at which it is required is becoming quite wide. If the gap grows greater, RNNs may fail to learn to link the data leading to long-term dependencies. The LSTMs were developed to resolve these long-term dependencies by employing a precise memory cell. This network design employs unique neuronal portals to govern the part of the input to the memory cell and the fraction to be forgotten from the prior condition. Further to enhance the performance and training of the LSTMs, a Bidirectional LSTMs (BiLSTM) network has been introduced [36]. This network divides the state neuron to - forward and backward. As the name indicates, the forward section $\vec{b}$ is accountable for the positive direction of the word order and the negative section $b^{\leftarrow}$ is responsible for the reverse direction of the forward section. The general structure of BiLSTM is depicted in Figure 4. The output of the BiLSTMs is a concatenation of each section. The following equations are used to realize the LSTM networks.

$$i_t = \sigma(W_i\, b_{t-1} + U_i a_t + y_i) \quad (1)$$

$$f_t = \sigma\left(W_f\, b_{t-1} + U_f a_t + y_f\right) \quad (2)$$

$$\underline{C_t} = (W_c\, b_{t-1} + U_c a_t + y_c) \quad (3)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \underline{C_t} \quad (4)$$

$$O_t = \sigma(W_o\, b_{t-1} + U_o a_t + y_o) \quad (5)$$

$$b_t = o_t \odot tanh\, tanh\, C_t \quad (6)$$

In aforementioned equations, σ and $\odot$ are element-wise sigmoid and product function respectively, a_t and b_t are the input and hidden state vector at time t respectively, different weight matrices of the gates for input is U, W and y are weight matrices for hidden state and bias vectors respectively. The input vectors $(a_1, a_2, \dots \dots . a_n)$ are initialized randomly.

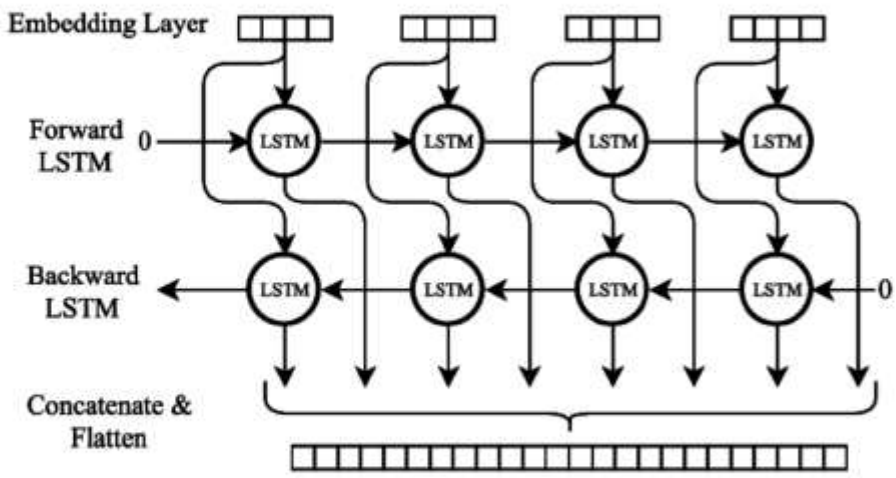

Figure 4. The general structure of BiLSTM [27]

The general architecture of the proposed aspect extraction and character embedding model is as shown in below Figure 5. This model comprises of long short-term memory (LSTM) network and a sequential model for fine-tuning the output of the LSTM.

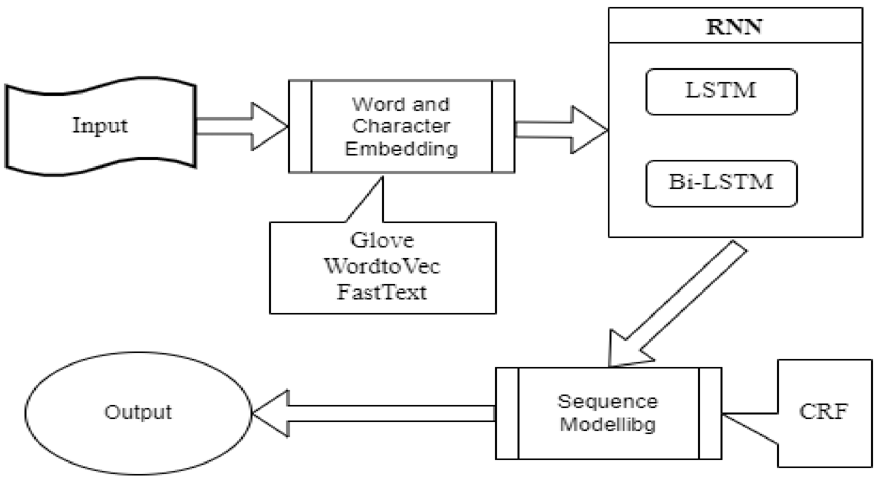

Figure 5. Proposed architecture of aspect extraction.

Sequence modeling is the process of forecasting the next word and/or character to appear. The current output is calculated using previous inputs and it does not depend on the length of words. Some of the sequential modelings are hidden Markov model (HMM), maximum entropy Markov model (MEMM), conditional random fields (CRF). The HMMs take into account only the observed state, and there is a target difference and predicted target function respectively. These issues can be resolved by using MEMMs. It uses conditional probability and not joint probability as in the case of HMM. MEMMs predict target function by taking account of the full observed sequence and adjacent states. The probability of input will not take into account leads to reduced workload and improved learning consistency amongst the target and estimated function correspondingly. The MEMMS is realized using the following equation.

$$P\left(\frac{Y_{1:n}}{X_{1:n}}\right) = \prod_{i=1}^{n} \frac{\exp(W^T f(y_i,\ y_{i-1}, X_{1:n}))}{Z(y_{i-1}, X_{1:n})} \qquad (7)$$

This model chooses the state which has fewer convertible states and resulting in the labeling bias problem. The increased complexity in CRFs provides the solution to the problem of MEMMs by implementing the normalization at global variance. CRFs are characterized by the following equation

$$P\left(\frac{C}{a}\ ; W, d\right) = \frac{\prod_{1=1}^{n} \alpha_i\ ((\ C_{i-1}, C_{i,}\ a))}{\sum_{C' \epsilon (\frac{c}{a})} \alpha_i (C'_{i-1,}\ C'_{i,}\ a)} \qquad (8)$$

In above expression α_i $=exp\ exp\ (W^T_{y',y} z_i + b_{y',y})$, a is the input sequence vector, C is label sequence vector, W is the weight vector, d is the bias label vector concerning the (C', C). Therefore, on top of LSTM, the inclusion of a sequence model improves the overall performance of the system.

Experimental Setup

Aspect extraction experiment is carried out using sequence labeling and in combination with the LSTMs as listed in Table 1. Data used in this experiment is Semeval-2014. The model parameters were selected based on the experimental values available in the literature. These models are based on the well-liked deep learning techniques CNN and RNN. The initial models have the names M-1, M-2, M-3. Different hybrid models with combination of M-1 & M-2 as M-1&2 and combination of M-1 & M-3 as M-1&3 with different pre-trained word embedding techniques used are Glove, Wordtovec, FastText.

Table 1. Models used in the experiment for word and character embedding

	Models
M-1	CRF
M-2	LSTM
M-3	BiLSTM
Hybrid-1(M- 1&2)	LSTM CRF
Hybrid- 2(M- 1&3)	BiLSTM CRF

The standard performance metrics are used for proposed model evaluations and are precision, recall, and F1-score, as described below. The parameters required to calculate these performance metrics are listed in Table 2.

$$Precision = \frac{True\ positive}{True\ positive + false\ positive} \qquad (9)$$

$$Recall = \frac{True\ positive}{True\ positive + false\ negative} \qquad (10)$$

$$F1 - score = \frac{(2)(Precision)(Recall)}{Precision + Recall} \qquad (11)$$

Table 2. Classification matrix.

Category	Positive	Negative
Positive	True Positive (TP)	False positive (FP)
Negative	False negative (FN)	True Negative (TN)

Baseline Models

We test the performance of our projected models to a variety of baselines to validate performance:

CNN based on dependency tree [37]: To capture grammatical information, a convolution layered neural network with dependency trees is used.

Memory interactions [38]: Deep multi-task training approach relies on LSTM. It conducts aspect and opinion extraction tasks collaboratively using memory interactions.

Deep CNN [39]: Poria et al. used a deep convolutional neural network with Glove.840B word embedding.

Recursive neural conditional random fields [40]: CRF input, they employed tree-structured features and a recursive neural network. RNCRF-O was prepared without the use of opinionated labels. RNCRF-F was trained using opinion labelling and even some custom characteristics.

Embedding word and dependency path [41]: Word embedding, linear context embedding, and dependency path embedding were used to improve CRF.

4. Results

The average performance of the models with and without word embedding approaches for the Semeval-2014 restaurant dataset of the restaurant at the word character level is shown in Table 3. It can be noticed that the inclusion of sequence model and word embedding techniques to the LSTM model overall performance is improved. The highest F1-score is achieved in the combination of models consisting of BiLSTM, CRF, and word embedding technique Glove.840B (Hybrid -2 & Glove.840B).

Table 3. Performance evaluation of different models for sentiment analysis

Models	Precision	Recall	F1-Score
M-1	76.3	74.62	75.45
Hybrid-1	83.9	80.86	82.35
Hybrid-2	83.5	80.73	82.09
M-2 & Glove.840B	83.01	81.07	82.03
M2 & Glove.6B.300	79.4	77.31	78.34
M-2 word2vec	79.3	76.57	77.91
M-2 FastText	81.5	76.71	79.03
M-3 FastText	83.67	82.16	82.91
Hybrid -1 & Glove.840B	86.01	82.52	84.23
Hybrid -1 & Glove.6B.300	82.04	84.23	83.12
Hybrid -1 & word2vec	82.59	83.45	83.02
Hybrid -1 & FastText	82.99	85.14	84.05
Hybrid -2 & Glove.840B	87	83.20	85.06
Hybrid -2 & Glove.6B.300	83.9	82.59	83.24
Hybrid -2 & word2vec	81.2	84.19	82.67
Hybrid -2 & FastText	87.03	83.08	85.01

The superiority of the BiLSTM-based model over the LSTM standard is confirmed by experiments. Figure 6 has really been inspected and confirmed. The models are compared with the architectures LSTM and BiLSTM through all pre-trained words embedding systems analysed. The added CRF layer over the neural network architecture significantly improved the performance of all models.

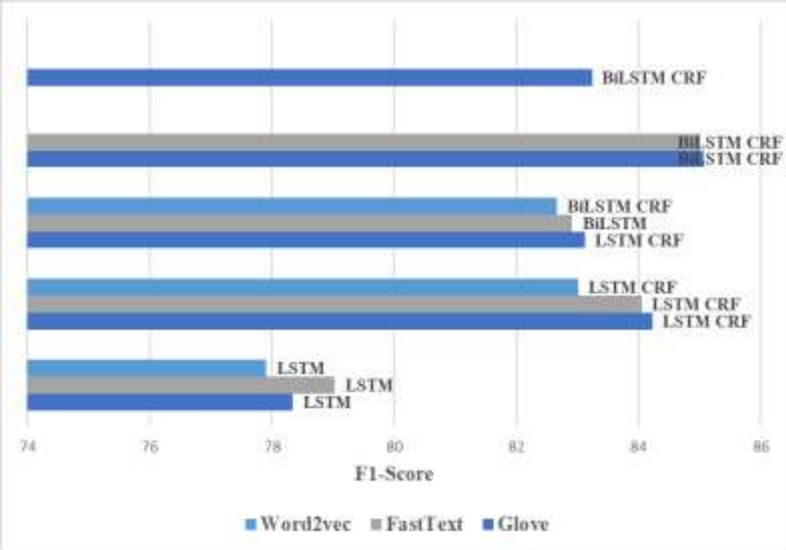

Figure 6. Performance comparison of different word embedding techniques

Finally to validate the performance of the proposed models, compared with the existing approaches available in the literature are shown.

Table 4. Performance comparison of proposed models with five highest F1-Score and previous works on SemEval 2014

Previous Works		**Proposed Models**	
Models	**F1-score**	**Models**	**F1-score**
CNN based on dependency tree [37]	83.97	Hybrid -2 & Glove.840B	85.06
Memory interactions [38]	77.58	Hybrid -2 & Glove.6B.300	83.24
Deep CNN [39]	82.76	Hybrid -1 & Glove.840B	84.23
Recursive neural conditional random fields [40]	82.73	Hybrid -1 & FastText	84.05
Embedding word and dependency path [41]	84.97	Hybrid -2 & FastText	85.01

5. Conclusion

In this paper, LSTM based different models are developed for aspect-level sentiment analysis. Different word embedding approaches are incorporated to analyze which word embedding methods are best suited for this experiment. On top of LSTM architecture, a sequence model is introduced such as MEMMS, CRFs. The addition of these sequence models shows the improved performance in extracting the aspect-based sentiment analysis. We used various customizations of LSTM-based architectures and different word representations to analyze sequence labelling approaches for aspect term extraction comprehensively. We tried to compare several pre-trained word embeddings and language models, and it's clear that the appropriate embeddings are critical for the ultimate model's effectiveness. The LSTM CRF Glove.840B, BiLSTM CRF Glove.840B, LSTM CRF FastText, BiLSTM CRF FastText models exhibit the best F1-score as 84.23%, 85.06%, 84.05%, and 85.01% respectively amongst the proposed models. Therefore, it is very important to evaluate the word embedding techniques and models to obtain optimum results for particular applications.

Conflicts of Interest

The authors declare no conflict of interest

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

[1] Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.

[2] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.

[3] Kumar, A., & Teeja, M. S. (2012). Sentiment analysis: A perspective on its past, present and future. International Journal of Intelligent Systems and Applications, 4(10), 1.

[4] Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. Electronics, 9(3), 483.

[5] Abid, F., Alam, M., Yasir, M., & Li, C. (2019). Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. Future Generation Computer Systems, 95, 292-308.

[6] Rao, D., & Ravichandran, D. (2009, March). Semi-supervised polarity lexicon induction. In Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009) (pp. 675-682).

[7] Ding, X., Liu, B., & Yu, P. S. (2008, February). A holistic lexicon-based approach to opinion mining. In Proceedings of the 2008 international conference on web search and data mining (pp. 231-240).

[8] Jiang, L., Yu, M., Zhou, M., Liu, X., & Zhao, T. (2011, June). Target-dependent twitter sentiment classification. In Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies (pp. 151-160).
[9] Dong, L., Wei, F., Tan, C., Tang, D., Zhou, M., & Xu, K. (2014, June). Adaptive recursive neural network for target-dependent twitter sentiment classification. In Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 2: Short papers) (pp. 49-54).
[10] Li, X., Bing, L., Lam, W., & Shi, B. (2018). Transformation networks for target-oriented sentiment classification. arXiv preprint arXiv:1805.01086.
[11] Ma, X., Zeng, J., Peng, L., Fortino, G., & Zhang, Y. (2019). Modeling multi-aspects within one opinionated sentence simultaneously for aspect-level sentiment analysis. Future Generation Computer Systems, 93, 304-311.
[12] Chiranjeevi, P., Santosh, D. T., & Vishnuvardhan, B. (2019). Survey on sentiment analysis methods for reputation evaluation. In Cognitive Informatics and Soft Computing (pp. 53-66). Springer, Singapore.
[13] Luo, H., Li, T., Liu, B., Wang, B., & Unger, H. (2019). Improving aspect term extraction with bidirectional dependency tree representation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 27(7), 1201-1212.
[14] Li, H., & Lu, W. (2017, February). Learning latent sentiment scopes for entity-level sentiment analysis. In Thirty-First AAAI Conference on Artificial Intelligence.
[15] Tay, Y., Tuan, L. A., & Hui, S. C. (2018, April). Learning to attend via word-aspect associative fusion for aspect-based sentiment analysis. In Proceedings of the AAAI conference on artificial intelligence (Vol. 32, No. 1).
[16] Chi, C. G. Q., Ouyang, Z., & Xu, X. (2018). Changing perceptions and reasoning process: Comparison of residents' pre-and post-event attitudes. Annals of Tourism Research, 70, 39-53.
[17] Pham, D. H., & Le, A. C. (2018). Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis. International Journal of Approximate Reasoning, 103, 1-10.
[18] Dinsoreanu, M., & Bacu, A. (2014, October). Unsupervised Twitter Sentiment Classification. In KMIS (pp. 220-227).
[19] Kiritchenko, S., Zhu, X., Cherry, C., & Mohammad, S. (2014, August). NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. In Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014) (pp. 437-442).
[20] Jo, Y., & Oh, A. H. (2011, February). Aspect and sentiment unification model for online review analysis. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 815-824).
[21] Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2016, November). Attention-based LSTM for aspect-level sentiment classification. In Proceedings of the 2016 conference on empirical methods in natural language processing (pp. 606-615).
[22] Chen, P., Sun, Z., Bing, L., & Yang, W. (2017, September). Recurrent attention network on memory for aspect sentiment analysis. In Proceedings of the 2017 conference on empirical methods in natural language processing (pp. 452-461).
[23] Ma, D., Li, S., Zhang, X., & Wang, H. (2017). Interactive attention networks for aspect-level sentiment classification. arXiv preprint arXiv:1709.00893.
[24] Song, Y., Wang, J., Jiang, T., Liu, Z., & Rao, Y. (2019). Attentional encoder network for targeted sentiment classification. arXiv preprint arXiv:1902.09314.
[25] Jakob, N., & Gurevych, I. (2010, October). Extracting opinion targets in a single and cross-domain setting with conditional random fields. In Proceedings of the 2010 conference on empirical methods in natural language processing (pp. 1035-1045).
[26] Toh, Z., & Wang, W. (2014). Dlirec: Aspect term extraction and term polarity classification system. In Association for Computational Linguistics and Dublin City University.
[27] Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., & Dyer, C. (2016). Neural architectures for named entity recognition. arXiv preprint arXiv:1603.01360.
[28] Ma, X., & Hovy, E. (2016). End-to-end sequence labeling via bi-directional lstm-cnns-crf. arXiv preprint arXiv:1603.01354.

[29] Quoc Le, Tomas Mikolov, and Tmikolov Google Com. Distributed Representations of Sentences and Documents.32, 2014.
[30] Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).
[31] Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. Transactions of the Association for Computational Linguistics, 5, 135-146.
[32] Ling, W., Luís, T., Marujo, L., Astudillo, R. F., Amir, S., Dyer, C., ... & Trancoso, I. (2015). Finding function in form: Compositional character models for open vocabulary word representation. arXiv preprint arXiv:1508.02096.
[33] Ballesteros, M., Dyer, C., & Smith, N. A. (2015). Improved transition-based parsing by modeling characters instead of words with LSTMs. arXiv preprint arXiv:1508.00657.
[34] Zhang, X., Zhao, J., & LeCun, Y. (2015). Character-level convolutional networks for text classification. Advances in neural information processing systems, 28, 649-657.
[35] http://www.bioinf.jku.at/publications/older/2604.pdf
[36] Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural networks, 18(5-6), 602-610.
[37] Ye, H., Yan, Z., Luo, Z., & Chao, W. (2017, May). Dependency-tree based convolutional neural networks for aspect term extraction. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 350-362). Springer, Cham.
[38] Li, X., & Lam, W. (2017, September). Deep multi-task learning for aspect term extraction with memory interaction. In Proceedings of the 2017 conference on empirical methods in natural language processing (pp. 2886-2892).
[39] Poria, S., Cambria, E., & Gelbukh, A. (2016). Aspect extraction for opinion mining with a deep convolutional neural network. Knowledge-Based Systems, 108, 42-49.
[40] Wang, W., Pan, S. J., Dahlmeier, D., & Xiao, X. (2016). Recursive neural conditional random fields for aspect-based sentiment analysis. arXiv preprint arXiv:1603.06679.
[41] Yin, Y., Wei, F., Dong, L., Xu, K., Zhang, M., & Zhou, M. (2016). Unsupervised word and dependency path embeddings for aspect term extraction. arXiv preprint arXiv:1605.07843.

Authors' Profiles

Usha G R received a BE degree in Information Science and Engineering from SJMIT, Chitradurga in the year 2007, and MTech in Computer Science and Engineering from UBDTCE, Davanagere in 2010. She has 13 years of teaching experience. She organized various workshops in the department and also published many papers in reputed journals and international conferences. Her subjects of interest include Big Data Analytics, Machine Learning, and Artificial intelligence.

J V Gorabal received BE in Computer Science & Engineering from BEC, Bagalkot, MTech in Computer Science & Engineering, KBNCE, Kalburgi, and Ph.D. in Computer Science & Engineering, from JNT University Anantapuram respectively. He had 25 years of teaching experience. Currently, working as a Professor in the Department of Computer Science & Engineering, ATME College of Engineering, Mysuru. He published many papers in national and international journals and attended conferences and workshops. His current research interests include Image Processing and Artificial Intelligence.