Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-020-04827-6

RESEARCH ARTICLE-COMPUTER ENGINEERING AND COMPUTER SCIENCE

Improving Text Summarization using Ensembled Approach based on Fuzzy with LSTM

Minakshi Tomer[1,2] · Manoj Kumar[3]

Received: 24 October 2019 / Accepted: 23 July 2020

Abstract
Abstractive text summarization using attentional recurrent neural network (sequence-to-sequence) models have proven to be very effective. In this paper, a novel hybrid approach is presented for generating abstractive text summaries by combining fuzzy logic rules (which selects extractive sentences) with bidirectional long short-term memory (Bi-LSTM) which further produces abstractive summary. Bi-LSTM uses attention mechanism and Adam optimizer for updating network weights. The proposed approach utilizes fuzzy measures and inference to extract textual information from the document to find the most relevant sentences. These relevant sentences are given as input to Bi-LSTM to produce an abstractive summary of the significant sentences. The proposed FLSTM model is evaluated using ROUGE toolkit. The experiment is performed on standard datasets (i.e., DUC and CNN/daily mail). Another salient feature of this work is merging of DUC 2003–2004, DUC 2006–2007 datasets to generate a larger dataset to achieve better results. The FLSTM model is compared with other state-of-the-art models, and the empirical results suggested that the proposed FLSTM model outperforms all other models.

Keywords Automatic text summarization · Fuzzy rules · Feature extraction · Bi-LSTM

1 Introduction

Text summarization (TS) is a procedure that produces the condensed text from the source document. The purpose of text summarization is to find out the major concept and producing a text of shorter length compared to source document. One of the most challenging problems in the area of NLP is to locate the significant parts present in input document [1]. Text summarization obtains a text of shorter length with all pertinent information present from an abundance of text sources. As the amount of online data has grown tremendously with superabundant information from different sources, it makes very time consuming and difficult for users to obtain relevant information. Due to this if anyone searches for some particular information, a lot of data is retrieved which is impossible for a person to read thoroughly. This brings an exponential growth in the area of generating automatic text summary (ATS).

ATS can be majorly grouped into two different categories: extractive and abstractive on the basis of output [2]. Extractive text summarization approach produces summaries by extracting words, set of words, phrases, set of phrases or sentences from the source document. It consists of three main steps: intermediate representation of source document, calculating sentence score and sentence selection [3]. Such extractive summaries can be generated by using sentence selection, machine learning techniques, statistical approaches and soft computing approaches [4]. An abstractive approach rephrases the main content in the input text to produce summary which can correlate to summary generated by human. [5] categorized text summarization at different levels: surface level, entity level and discourse level. The aim of surface-level approach is to represent the salient features of source text which includes thematic features, cue words, phrases, location and background. Entity-level approach represents the connection between the entities in the source text and its relationship. These approaches represent similarity, proximity, co-occurrence, thesaural relationship among words, co-reference, logical relations between the entities. Finally, discourse-level approach represents global

✉ Minakshi Tomer
tomer.minakshi@gmail.com

1 USICT, Guru Gobind Singh Indraprastha University, New Delhi, India

2 Maharaja Surajmal Institute of Technology, New Delhi, India

3 Ambedkar Institute of Advanced Communication Technologies and Research, New Delhi, India

Published online: 26 August 2020

structure of the text which concerns with format of document, threads of topics and rhetorical structure of text. [2] explained other approaches for extractive text summarization: statistical based, topic based, graph based, discourse based and machine learning based. The deep learning algorithms outperforms most state-of-the-art approaches in area of text summarization [6,7] and NLP tasks such as semantic role labeling (SRL), named entity recognition (NER), POS tagging [8]. Fuzzy-based approaches for text summarization has been applied in most of the researches performed by [4,9,10]. Furthermore, there are some examples of hybridization of fuzzy set with neural network which generates better result [4,11]

In this paper, a hybrid fuzzy-based LSTM (FLSTM) algorithm which is a fusion of fuzzy logic and LSTM for the document summarization is proposed. The motive of this approach is to combine the essence of both extraction-based methods and abstractive methods. In this model, fuzzy logic is used for extraction-based summarization and LSTM neural network is used for abstractive summarization. The task of former method is to identify and select the most informative sentences that have the highest probability to be included for final generated summary. The selected sentences than pass through the LSTM layers. In deep neural network model, bidirectional LSTM is used as an encoder and a unidirectional decoder with beam search for inference is used as decoder.

The remaining section of this paper is categorized as follows: Sect. 2 reviews the literature work on text summarization using deep neural network and fuzzy approaches. Section 3 presents the proposed approach. Section 4 presents the experimental results. Section 5 concludes the paper and gives future scope.

2 Literature Survey

This section provides a review of deep learning-based and fuzzy logic-based automatic text summarization.

2.1 Deep Learning-Based Text Summarization

Majority of work in the field of extraction-based text summarization has been done [5] where a summary is generated from the sentences already existing in the source text. Apart from extractive summarization there is abstractive summarization which generates text similar to what a human will generate by rephrasing sentences with new words. Statistical approach, topic-based, graph-based, discourse-based and machine learning are the commonly used approaches for generating summaries [2,3,5,12]. A deep neural network architecture [13] used RNN as encoder as well as decoder for extractive text summarization.

In this area of abstractive text summarization, deep neural networks generate state-of-the-art results [8,14,15]. Sequence-to-sequence models have been used by many authors to generate summary which is abstractive in nature [6,7,16]. In [16], a convolutional neural network is utilized as an encoder and feed forward network as a decoder to produce a quality summary of DUC 2004 and Giga word dataset. [6] used RNN as decoder and achieved better results. [7] used RNN as an encoder as well as a decoder for abstractive text summarization, this paper has also highlighted the critical problems in the area of summarization. An extension to these models is done by adding the attention mechanism in [17] which considers the context cues in hidden states of encoder to decode the required sequence. The encoder-decoder model suffers from two main disadvantages: First incapability to handle OOV (out of vocabulary) words. Second unnatural summaries (repeated words) were often generated by the encoder-decoder model. Solution to OOV problem was resolved by applying a copy mechanism which can handle OOV words and coverage mechanism can resolve the repetition problem [18]. [15] proposed dual encoder for automatic text summarization (DEATS) which consists of two encoders (primary and secondary) and a decoder. In DEATS, the primary encoder and decoder conduct the encoding process, while secondary encoder generates encoding. This model generates results on CNN/daily Mail dataset and DUC 2004.

2.2 Fuzzy Summarization

Apart from deep learning-based techniques many researchers [10,19,20] successfully applied fuzzy logic for text summarization. In [21], linguistic and human conception is taken into consideration. A parser is designed using Visual C++ 6.0 for selecting sentences on basis of location and other attributes using fuzzy inference system. On the basis of six features selected by fuzzy logic each sentence is assigned a fitness value. The text is compressed and highly scored sentences compared to threshold value are selected as a summary. The dataset used by authors were three news articles and results were being compared with Microsoft office 2000 and Copernic summarizer. Three parameter precision, recall and F1 score were used for evaluation purpose. Accuracy of this method was found to be 96% of that of human reader.

Witte and Bergler [19] presented a cluster graph algorithm based on fuzzy set for the analysis of documents. The algorithm exhibits a flexible, easy to implement and integrate, context sensitivity. It shows how extraction of some common and distinctive topic of a document is more informative than keyword extraction. The dataset selected was DUC (2003/2004) and DUC (2005/2006). The result was evaluated using ROUGE package and compared to other system.

Kyoomarsi et al. [22] proposed a fuzzy summarization technique to generate extractive summary which was compared with the machine learning algorithms (C4.5 and Naïve bayes algorithm). Algorithms were applied on TOEFL text and the results were evaluated by five judges who were expert in English language. The results indicated that fuzzy method worked better than machine learning-based method summarization.

Suanmali et al. [23] proposed fuzzy logic method to extract the important sentences on the basis of sentence score. Sentence score was calculated using nine features. The characteristics were categorized using Gaussian membership function. DUC-2002 dataset is used for testing and training of the model.

Suanmali et al. [10] used the same text on DUC-2002 dataset with eight features and triangular membership function. The results were analyzed with baseline summarizer and Microsoft word 2007 summarizer. ROUGE is used for evaluation metric and it was observed that better results were produced by using fuzzy method.

Leite and Rino [24] described fuzzy-based knowledge produced by genetic algorithm (GA) for extractive summarization. SuPor-2 fuzzy is based on machine learning approach that employs a fuzzy-based ranking model for selecting relevant sentences. The eleven features were combined with the naïve Bayes classifier to verify if fuzzy classification fit well for text summarization. TeMario corpus were used for training the model and ROUGE toolkit for the evaluation. SuPor-2 fuzzy was compared with the other summarizer and results for ROUGE-1 is (0.74) and ROUGE-2 is (0.73) respectively.

Binwahlan et al. [25] introduced a hybrid model for text summarization. Hybrid model with eight features is based on three different models. In the first method, a binary tree is generated after applying MMI (maximal marginal importance) on the sentences of binary tree. Second method used PSO technique (swarm MMI diversity based) to get optimized weights for all the features. Finally, these weights are used to calculate sentence score using fuzzy swarm-based method. In fuzzy logic method sentence score is calculated through fuzzy inference system, the membership function used is trapezoidal. The best 'n' sentences with the highest score are extracted to represents summary.

Hannah et al. [26] proposed a model based on sentence-oriented approach rather than a feature based for text summarization. Sentence score is calculated on the basis of extracted features. Whether the sentence will be present in summary or not is based upon these features as a whole instead of any specific feature. It improves the nature of summarization since it extracts important sentences using score calculated by fuzzy inference system. DUC-2002 dataset is used for testing the model and ROUGE score is used as a metric for summary evaluation. The results are compared with feature-oriented approach, DUC-2002 and MS-Word.

In [27] introduced a novel hybrid model for automatic text summarization which is a combination of genetic algorithm, fuzzy logic and semantic role labeling called fuzzy genetic semantic. The results in the paper confirms that fuzzy GA-SRL experience notable improvement in nature of text summarization.

Dixit and Apte [28] presents an extractive summarization in which fuzzy techniques was used for sentence selection. The input text of 30 documents is fetched from news-based URL where sentence score is calculated for each sentence using fuzzy centroid method. Finally, sentences with the highest score are selected to represents summary. Results are better when compared with Corpermic summarizer and MS word 2007 summarizer.

Megala et al. [29] presents comparison of extractive text summarization by using fuzzy logic rules and neural network. Ten extracted features are applied as input to fuzzy system in which triangular membership function is used. A neural network was trained on corpus of legal judgements and modifies through feature fusions and finally highly ranked sentences were selected. The results show that average of f-measure, precision(P) and recall (R) of fuzzy logic is better than neural network method.

Kumar et al. [30] proposed a summarization of multi-document (news components) using relations based on fuzzy cross-document. It includes three main stages: sentence extraction, relation identification of cross-document and sentence scoring utilizing fuzzy rules. It identifies CST relations among sentences using genetic CBR which integrates genetic learning algorithm to the case base reasoning model. The model was trained on DUC-2002 dataset and evaluated using ROUGE measures where it is compared with the other models. This integration of these two techniques brings improvement to the results.

Megala et al. [31] presents a combination of fuzzy logic with conditional random field to produce the summary of legal judgement document. Eleven extracted features were used to compute the sentence score. Fuzzy logic extracts important sentences and CRF (Conditional Random Field) is used to perform the classification by identifying the rhetorical rules present in the legal document. This classification technique converts the paragraph summary into structured summary based on keywords.

Babar and Patil [32] describes a fuzzy logic extraction approach with eight extracted features for text summarization. Sentence score is calculated with all features which compares with the summary generated using SVD (singular value decomposition). In SVD-based approach important sentences are extracted based on importance of concept and sentences. The V^T matrix cell represents the significant sentences. Summaries from fuzzy and SVD are intersected and a set of common and uncommon sentences are extracted.

From set of uncommon sentences, ones with the high score is selected.

Abbasi-ghalehtaki et al. [33] proposed a model for extractive text summarization build on evolutionary algorithms, fuzzy logic and cellular learning automata (CLA). In this model, feature extraction of both word features and sentence features are performed. The particle swarm optimization technique is used to assign weights to extractive feature after which fuzzy rules were applied for scoring sentences. In this, artificial bee colony (ABC) and cellular learning automata (CLA)-based approach were also proposed for extracting diverse and relevant sentences. The training and testing are done on DUC-2002 dataset.

Chopade et al. [34] proposed a hybrid model consists of neural network and fuzzy rules for extractive text summarization. RBM (Restricted Boltzman machine) consists of one input, two hidden and one output layer for training the system. A seed table and a sentence table are created on which membership function is applied to calculate the score of sentences. Hybrid model is tested on 25 documents using ROUGE toolkit.

Azhari and Jaya Kumar [4] proposed a neuro-fuzzy approach based on classification for improving text summarization. The proposed model ANFIS was compared with the existing model based on neural network and fuzzy rule-based approach. Results were found to be better in terms of precision, recall and F-measure on DUC data set.

Lakshmi and Rani [35] proposed a multi-document text summarization combining fuzzy logic with deep learning algorithm. Features are extracted from preprocessed document which are then processed with a fuzzy classifier. Feature matrix is passed through the RBM (Restricted Boltzman Machine) layer by layer to generate efficient text summary. The model is trained and tested on DUC-2002 dataset. The proposed method gives better result in terms of precision, recall and F-measure.

Sahba et al. [36] proposed a novel model for text summarization based on fuzzy features and attention-based sequence-to-sequence model. This model benefit state-of-the-art approach to both summarization techniques extractive and abstractive. Benchmark dataset CNN/Daily mail is fed to fuzzy logic system for generating an extractive summary which is further applied to pointer generator network. The fuzzy logic is implemented using MATLAB toolkit and pointer -generator network using TensorFlow. Results are compared with the other system through evaluation metric ROUGE 1, ROUGE 2 and ROUGE L. Proposed model makes the sequence-to-sequence model immune to noise.

In [37] a new text processing tool KUSH is introduced to preserve the semantic cohesion. Two major concepts Textual Graph and Maximum Independent sets are combined for the generation of extractive summary. In this, a novel mathematical approach is mentioned in which nodes forming the Maximum Independent set are removed to form the summary on standard dataset DUC-2002 and DUC-2004.

In [38], Karci Enropy is utilized for the first time for extractive text summarization. The advantage of this method is that it does not require any training data and it selects the most relevant and informational sentences from text document. Experiments are performed on DUC-2002, DUC-2004 datasets. The results outperforms state-of-the-art model.

3 Proposed FLSTM-Based Text Summarization

In this section, a hybrid model is proposed. The proposed model combines extractive method and abstractive method for efficient document summarization. For extractive summarization, fuzzy logic is utilized which selects the most informative sentences which is further passed through LSTM neural network to generate the final abstractive summary. The method consists of four major steps executed in the following stages: (i) Text preprocessing, (ii) Feature extraction, (iii) Extractive summarization by fuzzy rules and (iv) Abstractive summarization using Bi-LSTM as depicted in Fig. 1.

3.1 Phase 1: Text Preprocessing

Preprocessing is the initial phase for text summarization. It includes identification of sentences present in a paragraph (segmentation); splitting the running text into words called tokens (tokenization); stop-word removal and eliminating irrelevant information like articles, adverbs and pronouns; normalizing the text which includes stemming (reducing various suffixes added to a word).

3.1.1 Segmentation and Tokenization

The preprocessing module includes different processes. It starts with the identification of different component sentences and then segmenting them. Further, the continuous set of characters is split into tokens (this process is referred as tokenization).

3.1.2 Stop Words Removal and Stemming

After tokenization irrelevant words which are also referred as stop words are removed. These are commonly used words which can be ignored (such as articles, prepositions, adverbs, pronouns). These words contribute nothing to the meaning of a sentence. Cleaning the text from these words is called stop word removal. Finally, a linguistic normalization is done to reduce words to their word stem (stemming).

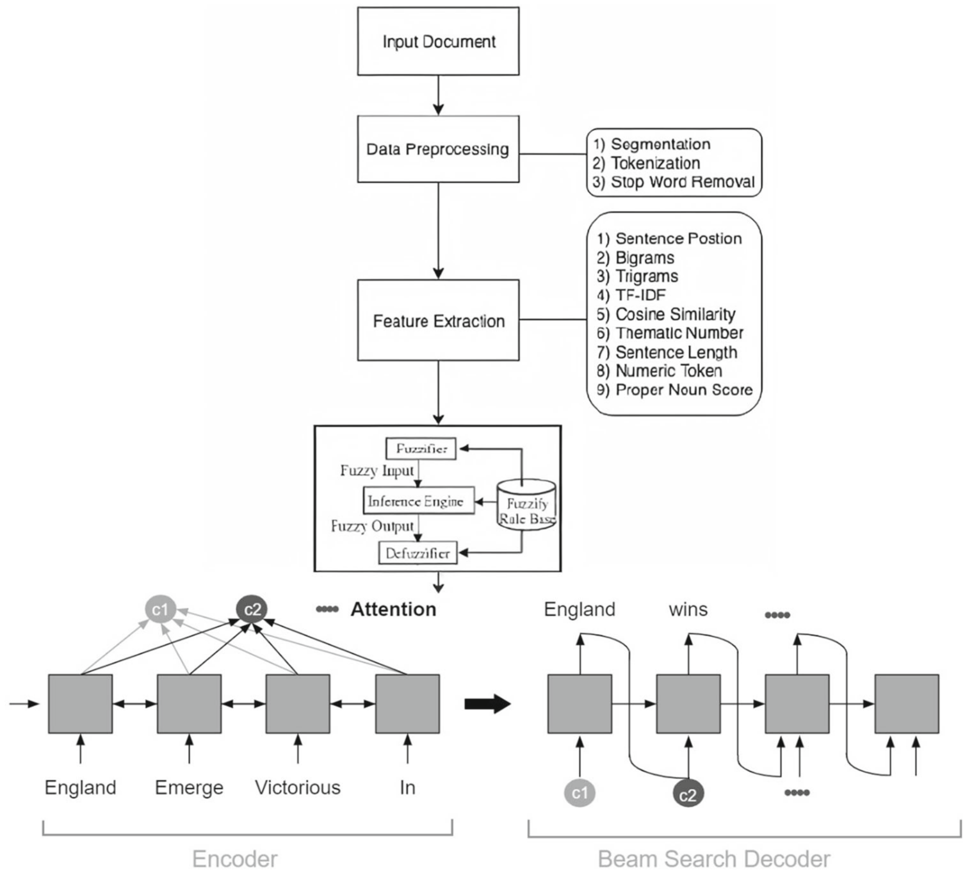

Fig. 1 Fuzzy LSTM (FLSTM)-based text summarization model

3.2 Phase 2: Feature Extraction

The preprocessed data is represented as a vector consisting of features. These features are utilized to ascertain the importance of each sentence. The summary is composed of sentences with high scores. Appropriate feature selection has a high impact on the quality of summary. Nine features are extracted for each sentence in the input data. The nine features are [28,32]: - Sentence Position, Bigrams, Trigrams, TF-IDF, Cosine Similarity, Thematic number, Sentence length, Proper Noun Score and Numeric Token.

3.2.1 Sentence Position

The key concept here is the sentence appearing in the beginning or end of input text are considered to be more important from the rest of the set of sentences. It is calculated using equation (1).

$$\textit{Sentence position} = \frac{\textit{Total number of sentences} - \textit{Index of sentence}}{\textit{Total number of sentences}} \tag{1}$$

3.2.2 Bigrams

Any order of two adjoining words present in the input string is Bigram. It is used to get the count of the set of two different adjacent words in the document.

3.2.3 Trigrams

Any sequence of three adjoining words in the input string is Trigram. It is used to get the count of the total set of three different adjacent words in the document.

3.2.4 TF-IDF

Initially, the sentences in the input text are converted to vectors before computing either the TF (term frequency) or IDF (inverse document frequency). TF-IDF weight is a statistical measure used to derive the importance of a word in the input text. TF is used for calculating relative frequency of words in the given sentence in its topic description, while IDF is a factor derived to diminish weight of the terms which occurs frequently in a set of document like ‘is,’ ‘and,’ ‘the,’ ‘of,’ etc., contributing nothing to the meaning of document and helps to increase weight of those terms which are occurring rarely. TF-IDF is calculated using equations (2-4).

$$TF(t) = \frac{Number\ of\ times\ term\ t\ appears\ in\ the\ document}{Total\ number\ of\ terms\ in\ the\ document} \tag{2}$$

$$IDF(t) = log\frac{Total\ number\ of\ documents}{Number\ of\ documents\ containing\ the\ term\ t} \tag{3}$$

$$TFIDF(t) = TF(t) * IDF(t) \tag{4}$$

3.2.5 Cosine-Similarity

The centroid of the sentence vector is calculated before working on cosine-similarity. Then the cosine function is used to identify the similitude among the different sentences with the centroid.

3.2.6 Thematic Score

Thematic words are the keywords present in the passage. These words are extracted from the document using RAKE [Rapid Automatic Keywork Extraction](Rose et al., 2010). Words extracted using this algorithm are given a high priority to be present in the summary.

3.2.7 Sentence Length

This is used to compare the length of every sentence with the size of longest sentence present in the input document. It is calculated using equation (5).

$$Sentence\ Length\ (L) = \frac{Length\ of\ sentence}{Length\ of\ longest\ sentence} \tag{5}$$

3.2.8 Proper Noun Score

Sentences containing proper noun are considered to carry more valuable information. Therefore, it is a good choice to select these sentences for the summary. It is obtained by dividing the proper nouns present in the given sentence to the length of that sentence.

$$Proper\ Noun\ Score\ (N) = \frac{Total\ number\ of\ proper\ nouns\ in\ sentence}{Total\ number\ of\ words\ in\ sentence} \tag{6}$$

3.2.9 Numeric Token

The text in which data is represented in the form of numbers is considered to be informative. It is obtained using the total number of numerical data present in the sentence divided by length of that sentence.

$$Numeric\ Token\ (S) = \frac{Total\ number\ of\ numeric\ data\ in\ sentence\ S}{Total\ number\ of\ words\ present\ in\ sentence\ S} \tag{7}$$

3.3 Phase 3: Fuzzy Rules

The text summarization model utilizes fuzzy logic approach to generate extractive summary. A fuzzy set over X is represented as $F = (x \cdot \mu_F(x)|x\epsilon X)$ where $\mu_F(x)$ denotes the membership function. This extractive text summarization method uses "IF/THEN" rules. The fuzzy-based system consists of a set of certain rules which are denoted by R (rules) where the i^{th} rule is defined as R^i: if m_1 is v_i^1 and m_2 is v_i^2 and $\ldots m_n$ is v_i^N then y is Y^i where v_i^N and Y^i are linguistic values. The first part of the rule "m_n is v_i^N" is known as the antecedent, whereas the second part "y is Y^i" is known as the consequent.

In the fuzzy set, antecedents are divided into three linguistic variables using triangular membership function into POOR, GOOD & AVERAGE. The consequents are also divided into three linguistic variables which are BAD, AVERAGE & GOOD using the triangular membership function. A set of five rules is produced (by combination of all input variables) in fuzzy rule base as mentioned in Table 1 and the fuzzy sets are represented using the Triangular membership function.

There are many different membership functions but this model employs triangular membership function. Also, this function is simpler and more flexible compared to other membership functions. The fuzzy system extracts the important sentences as shown in Fig. 2 from the document using the feature matrix which was calculated in the previous section. Then the output of the fuzzy is passed to the Bi-LSTM for abstractive text summarization.

3.4 Phase 4: Bi-LSTM

A sequence-to-sequence model with attention mechanism is used in FLSTM model. This model utilizes bidirectional encoder and a beam search decoder. The beam search decoder is used for inference to select the word with the highest probability as output by the decoder. Bahdanau Attention Mechanism with weight normalization is also used. A single vector as an input is passed to the decoder which has to accumulate all the information about the context. It is an issue when dealt with large sequences. Therefore, an attention mechanism is placed which enables the decoder to consider input sequence selectively. The standard seq2seq model is generally unable to accurately process long input sequences, since only last hidden state of the encoder RNN is used as the context vector for the decoder. On the other hand, the Atten-

Fig. 2 Example of extractive summary after application of Fuzzy Rules

[1] Taiwan share prices closed down percent Monday on wall street weakness and lack luster interim earnings from electronics manufacturing giant hon hai, dealers said.
[2] Shares closed down percent Monday following a weak lead from the united states and lower commodity dealers
[3] The united nations ' humanitarian chief John Holmes arrived in Ethiopia Monday to tour regions affected by drought, which has left some eight million people in need of urgent food aid.
[4] New Zealand share prices closed percent higher Monday in subdued trading ahead of us dealers said.
[5] Beijing is enjoying its best air quality in a decade thanks to steps taken for the authorities said Monday, amid a push by some locals for the anti-pollution measures to be made permanent.

Informative sentences (fuzzy classified as GOOD)

[1] The united nations ' humanitarian chief John Holmes arrived in Ethiopia Monday to tour regions affected by drought, which has left some eight million people in need of urgent food aid ."
[2] Beijing is enjoying its best air quality in a decade thanks to steps taken for the Olympics, authorities said Monday, amid a push by some locals for the measures to be made permanent

Table 1 Fuzzy Rules

IF	THEN
Sentence position = GOOD	Sentence = GOOD
Proper noun score = GOOD	
Numeric token score = GOOD	
Sentence length = GOOD	
Sentence position = POOR	Sentence = BAD
Proper noun score = POOR	
Numeric token score = POOR	
Proper noun score = POOR	Sentence = BAD
Thematic score = AVERAGE	
Cosine-similarity score = GOOD	Sentence = GOOD
Bi-tokens = GOOD	Sentence = AVERAGE
Tri-tokens = GOOD	
Numeric score = AVERAGE	
TF-IDF score = AVERAGE	

tion mechanism directly addresses this issue as it retains and utilizes all the hidden states of the input sequence during the decoding process. It does this by creating a unique mapping between each time step of the decoder output to all the encoder hidden states. This means that for each output that decoder makes, it has access to the entire input sequence and can selectively pick out specific elements from that sequence to produce the output.

The encoding model is composed of two distinct parts: embedding layer and RNN layers. The embedding layer as the first part of the model provides higher representative power for the words. The number of features gives the representation of each word in a sentence. Glove pretrained vectors is used to initialize word-embeddings. RNN layers are the second part of the model. In this model, several bidirectional RNN layers are stacked. The combined forward and backward layer outputs are used as input of the next layer.

Decoding model consists of two processes, training and inference. In this model, basic decoder is used for training purpose and beam search decoder for Inference purpose. Beam search considers multiple best options based on beamwidth using conditional probability, which is better than the sub-optimal greedy search or the normal approach.

Sequence loss function is a weighted SoftMax cross entropy loss function, which is particularly designed to be applied in time series model (RNN). The softmax classifier is a linear classifier that uses the cross entropy loss function. In other words, the gradient of the function tells a softmax classifier how exactly to update its weights using some optimization like gradient descent. Cross entropy measure is a widely used alternative of squared error. Adam Optimizer is used to optimize the performance of the model. Adam can be looked at as a combination of RMSprop and Stochastic Gradient descent with momentum. It uses the squared gradients to scale the learning rate like RMSprop and takes advantage of momentum by using moving average of the gradient instead of gradient itself like SGD with momentum.

FLSTM model is presented using Algorithm 1.

4 Experimental Results

For experimentation, benchmark DUC dataset and CNN/Daily Mail dataset is used. DUC was obtained from the Document Understanding Conference which is a series of summarization evaluations that have been conducted by the National Institute of Standards and Technology (NIST). ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metrics are used to evaluate results. The proposed ensembled model is implemented using TensorFlow. All the experiments have

Algorithm-1. Proposed Model of FLSTM

Input: Text
Ouput: Summarized Text
Step 1: The documents are preprocessed and reference summaries are used to obtain the salience score for every sentence;
Step 2: Construct a table for word embedding using pretrained word vectors;
Step 3: Calculate the nine-feature extraction
P = Calculate Sentence Position
B = Bigrams
T = Trigrams
F = TF-IDF
C = Cosine Similarity
H = Thematic Score
L = Sentence Length
N = Proper Noun Score
S = Numeric Token Score
Step 4: Partition Universe of Each Variant Fuzzification (P, B, T, F, C, H , L, N, S)
Step 5: Aggregate according to Fuzzy IF THEN Rules
Step 6: J = Defuzzification
Step 7: K = Bi-LSTM(J)

been conducted on Google Compute Engine with 1.2 GHz CPU, 64 GB RAM, Tesla K80 GPU and running Ubuntu 16.04.

4.1 Dataset

For our experiments the benchmark datasets from DUC and CNN/Daily Mail are used for our proposed FLSTM model. In this work, DUC 2003/2004/2006/2007 and CNN/Daily Mail dataset are used for training and testing the FLSTM model. Datasets DUC 2003/2004, DUC 2006/2007 were combined together. The advantage of combining the dataset is better abstractive summary is generated with larger dataset. The DUC and CNN/Daily Mail datasets consist of English news articles. The characteristics of these datasets are given in Table 2. Overall, 60% dataset is used for training and other 40% is used for testing the system.

4.2 Evaluation Metric

Text summarization can be evaluated in two categories: intrinsic or extrinsic method [39]. The intrinsic type evaluation is being conducted in this work. Standard evaluation ROUGE toolkit [40] computes n-gram overlap between the model generated summary and reference (ideal) summary. It is used for evaluation of text summarization. It includes measures to automatically calculate the quality of a summary by comparing it to other (ideal) summaries created by humans. ROUGE-N measures the N-gram units common between a particular model generated summary and ideal summary. ROUGE-L computes Longest Common Subsequence (LCS) metric. Rouge generates the precision, recall and F-measure metric.

Table 2 Characteristics of datasets

Dataset	Number of articles/ documents	Average words/ Length limit
DUC-2003	350	100 words
DUC-2004	500	100 words
DUC-2006	1300	250 words
DUC-2007	1200	250 words
CNN dataset	90, 000	–
Daily Mail dataset	1, 97, 000	–

$$Precision = \frac{Gold\ standard\ summary \cap Model\ generated\ summary}{Model\ generated\ summary} \quad (8)$$

$$Recall = \frac{Gold\ standard\ summary \cap Model\ generated\ summary}{Gold\ standard\ summary} \quad (9)$$

F-measure summary is subsequently given by:

$$F\text{-}measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (10)$$

We have used all 3 metrics for evaluation in this paper.

4.3 Experimental Setting

The batch size is set for the implementation to be 64 along with the hidden layer with the dimensions of 150. The model utilize Glove embedding with embedding size 300. The parameters of the network are initialized randomly using uniform distribution later Adam optimization is used to optimize the parameters of the network. Adam is used because it utilizes advantages of both AdaGrad and RMSProp. The learning rate was set to be 0.001 along with the number of epochs while training was 10. At the time of training we used a basic decoder with attention mechanism. At the time of testing we decoded the output summaries using a beam search decoder with beam width set to 10.

4.4 Results

To show the performance of FLSTM the proposed model is compared with RBM summarization model and other traditional models. The experimental results are shown in Table 3 and Table 4. The results show that FLSTM model outperforms state-of-the-art models on DUC 2003/2004 and DUC 2006/2007 datasets (Fig. 3). The result of proposed model

Table 3 Performance comparison of FLSTM on DUC 2003-2004

Model	Rouge-1			Rouge-2			Rouge-L		
	P	R	F-score	P	R	F-score	P	R	F-score
LSA	0.285	0.313	0.297	0.038	0.041	0.039	0.238	0.248	0.260
TextRank	0.334	0.398	0.362	0.068	0.079	0.073	0.258	0.306	0.279
LexRank	0.304	0.322	0.312	0.051	0.054	0.052	0.244	0.257	0.250
RBM	0.238	0.273	0.254	0.071	0.089	0.082	0.210	0.241	0.423
FLSTM	0.350	0.480	0.456	0.077	0.154	0.242	0.350	0.345	0.464

Table 4 Performance comparison of FLSTM on DUC 2006-2007

Model	Rouge-1			Rouge-2			Rouge-L		
	P	R	F-score	P	R	F-score	P	R	F-score
LSA	0.184	0.192	0.188	0.052	0.056	0.055	0.154	0.168	0.162
TextRank	0.216	0.220	0.217	0.066	0.071	0.068	0.172	0.185	0.179
LexRank	0.223	0.234	0.228	0.069	0.077	0.073	0.178	0.191	0.186
RBM	0.259	0.280	0.268	0.086	0.093	0.089	0.230	0.248	0.243
FLSTM	0.350	0.361	0.286	0.163	0.114	0.103	0.350	0.361	0.365

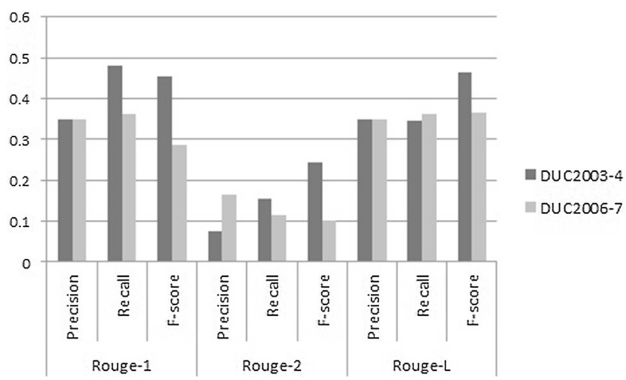

Fig. 3 Precision, Recall and F-Score of FLSTM for DUC 2003-2004 and DUC 2006-2007

as shown in Table 5 achieves better results compared to traditional models viz. [7,15,17,18,36,41,42] on Rouge-1 and Rouge-L scores and comparable results on Rouge 2 (Fig. 4). The ROUGE 1, ROUGE 2 and ROUGE L scores for all the metrics viz. Precision, Recall and F-score are clearly better for FLSTM compared to other traditional models. A sample summary generated by FLSTM on both datasets (DUC and CNN/Daily Mail) is presented in Figs. 5 and 6.

5 Conclusion and Future Scope

In this work, a hybrid (extractive+abstractive) approach is used for text summarization. Fuzzy logic is utilized for extractive summarization, whereas sequence-to-sequence deep learning model is used for abstractive summarization. The integration of fuzzy with deep learning model not only improves the results but also reduces the time required to train the abstractive model. The idea is to overcome the limitations of extractive and abstractive summarization. The limitation of extractive summarization is that it just selects the impor-

Table 5 Rouge scores of various summarization models on CNN/Daily Mail dataset

Model	Rouge-1	Rouge-2	Rouge-L
Fuzzy Logic System [36]	27.02	25.91	26.33
Sequence to sequence Neural Network [36]	39.53	17.28	36.28
Sequence to sequence + Attention [17]	31.34	11.79	28.10
Words-lvt2k-temp-att [7]	35.50	13.30	32.65
Summa Runner-abs [41]	37.50	14.50	33.40
RL+ML [42]	39.87	15.82	36.90
Pointer generator [18]	36.44	15.66	33.42
Pointer generator (pointer+coverage) [18]	39.53	17.28	36.38
DEATS [15]	40.85	18.08	37.13
FLSTM	40.96	15.22	38.63

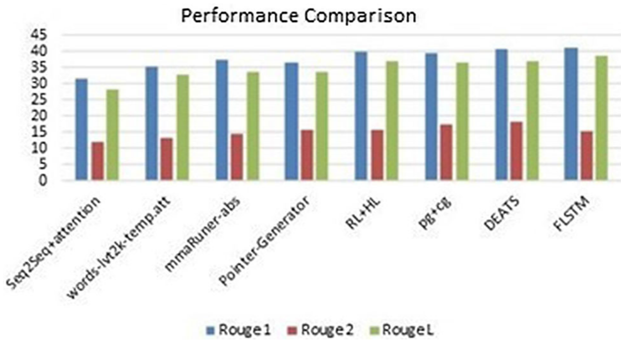

Fig. 4 Performance comparison of FLSTM with state-of-the-art models on CNN/Daily Mail dataset

tant sentences from the summary. In comparison, abstractive summary is better as it rephrases the sentences. The limitation of abstractive summarization is that it needs large databases to train and training consumes lots of time. Extract the important sentences with extractive summarization and passing these sentences for abstractive summarization can overcome the limitations of both approaches, i.e., we get abstractive summary with less time consumption.

The deep neural network model consists of an encoder and decoder in which encoder is bidirectional LSTM and decoder is beam search with attention mechanism. The final output generated from the proposed model is abstractive in nature.

The experimental results on DUC datasets (DUC 2003/2004 and DUC 2006/2007) and CNN/Daily Mail show that the proposed FLSTM model outperforms state-of-the art models in terms of precision(P), recall(R) and F-measure on ROUGE evaluation metric. For the CNN/Daily mail dataset a comparative analysis of FLSTM is presented with traditional methods. The results on Rouge-1, Rouge-L are better and comparable on Rouge-2. This work can be extended in future by enhancing the features and fuzzy rules or by replacing the existing deep learning model with a new model to obtain better results. Apart from fuzzy logic there are other extractive models which can be ensembled with deep learning models to improve the results.

Fig. 5 Comparison of gold truth summary with summaries from FLSTM on DUC dataset

Heading
Source document 1: Beijing is enjoying its best air quality in a decade thanks to steps taken for the Olympics, authorities said Monday, amid a push by some locals for the anti- pollution measures to be made permanent
Ground truth summary: Beijing 's best air in a decade due to Olympic measures
FLSTM Based summary: Beijing enjoys best air quality in decade
Source document 1: the united nations ' humanitarian chief john holmes arrived in Ethiopia Monday to tour regions affected by drought, which has left some eight million people in need of urgent food aid
Source document 2: the united nations ' humanitarian chief john holmes arrived in Ethiopia Monday to tour regions affected by drought, which has left some eight million people in need of urgent food aid
Ground truth summary: un 's top aid official arrives in drought-hit Ethiopia
FLSTM Based summary: un humanitarian chief visits Ethiopia
Source document 3: Clinton was going to the Pine Ridge Reservation for a visit with the Oglala Sioux nation and to participate in a conference on Native American homeownership and economic development. He also was touring a housing facility and signing a pact with Oglala leaders establishing an empowerment zone for Pine Ridge. But the main purpose of the visit -- the first to a reservation by a president since Franklin Roosevelt -- was simply to pay attention to American Indians, who are so raked by grinding poverty that Clinton's own advisers suggested he come up with special proposals geared specifically to the Indians' plight. At Pine Ridge, a scrolling marquee at Big Bat's Texaco expressed both joy over Clinton's visit and wariness of all the official attention: ``Welcome President Clinton. Remember Our Treaties," the sign read.
Ground truth summary: Clinton Focuses on American Indians
FLSTM Based summary: Clinton visits Indians

Fig. 6 Comparison of gold truth summary with summaries from FLSTM on CNN/Daily Mail dataset

Source-1
For the first time in a very long time, the well-intended but failed criminal penalties to protect public health and safety will be set aside. Adults who choose to use marijuana and obtain it through legal outlets will no longer be faced with the threat of criminal sanctions. People of color will no longer face the egregious inequities in how marijuana criminal penalties are imposed. Parents, as they help prepare their children for the choices, they face concerning marijuana, will no longer be hobbled by misinformation about the drug and the absence of effective supports to encourage abstinence.
Ground Truth Summary
Parents helping their children concerning marijuana, will no longer be hobbled up and people of color will no longer face the egregious inequities in how marijuana criminal penalties are imposed.
Model Summary
Parents, as they help prepare their children for the choices, they face concerning marijuana, will no longer be hobbled by misinformation about the drug and the absence of effective supports to encourage abstinence…People of color will no longer face the egregious inequities in how marijuana criminal penalties are imposed…
Source-2
I can understand the urge to react, to grasp at anything that might protect travellers. I too want air travel to be safe; hell, my husband is a pilot. But arming screeners at checkpoints well away from the airfield wouldn't be just another of the many precautions the airlines have taken to avert large-scale terrorism. It would simply be about protecting people from something that is everywhere in America: gun violence -- yes, at airports, and also at schools, at movie theatres, and malls.
Ground Truth Summary
It is simply about protecting people of America from the violence at airports, schools, malls and gun violence. But arming the screeners at each checkpoint is just another precaution taken by the airlines to avert terror attacks.
Model Summary
It would simply be about protecting people from something that is everywhere in America: gun violence—yes, at airport, and also at schools, at movie theatres, and malls…But arming screeners at checkpoints well away from the airfield wouldn't be just another of the many precautions the airlines have taken to avert large-scale terrorism…
Source-3
Following the overthrow of Libyan President Moammar Gadhafi, many Tuareg who had been fighting for Gadhafi 's forces reportedly returned to northern Mali, bringing their weapons with them. Last month, a Tuareg uprising triggered a military coup against Mali 's President Amadou Toumani Toure by officers dissatisfied with the government 's efforts to put down the insurrection. But in the disorder following the coup, the rebels seized large areas of the north.
Ground Truth Summary
In the uprising which occurred last month triggered a military coup against Mali's President by the dissatisfaction with the government efforts to put put down the insurrection. But this led the rebels to seize a large part of the North following the coup.

References

1. Hatzivassiloglou, V.; Klavans, J.L.; Eskin, E.: Detecting text similarity over short passages: Exploring linguistic feature combinations via machine learning. In: 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (1999)
2. Lloret, E.; Palomar, M.: Text summarisation in progress: a literature review. Artif. Intell. Rev. **37**(1), 1 (2012)
3. Nenkova, A.; McKeown, K.: Mining Text Data, pp. 43–76. Springer, Berlin (2012)
4. Azhari, M.; Jaya Kumar, Y.: Improving text summarization using neuro-fuzzy approach. J. Inf. Telecommun. **1**(4), 367 (2017)
5. Lloret, E.: Paper supported by the Spanish Government under the project TEXT-MESS (TIN2006-15265-C06-01) (2008)
6. Chopra, S.; Auli, M.; Rush, A.M.: Abstractive sentence summarization with attentive recurrent neural networks. In: Proceedings of the 2016 Conference of the North American Chapter of the Association

for Computational Linguistics: Human Language Technologies, pp. 93–98 (2016)
7. Nallapati, R.; Zhou, B.; Gulcehre, C.; Xiang, B.; et al.: Abstractive text summarization using sequence-to-sequence RNNS and beyond (2016). arXiv preprint arXiv:1602.06023
8. Young, T.; Hazarika, D.; Poria, S.; Cambria, E.: Recent trends in deep learning based natural language processing. IEEE Comput. Intell. Mag. **13**(3), 55 (2018)
9. Patil, P.D.; Kulkarni, N.: Text summarization using fuzzy logic. Int. J. Innovat. Res. Adv. Eng. **1**, 3 (2014)
10. Suanmali, L.; Salim, N.; Binwahlan, M.S.: Fuzzy logic based method for improving text summarization (2009). arXiv preprint arXiv:0906.4690
11. Goularte, F.B.; Nassar, S.M.; Fileto, R.; Saggion, H.: A text summarization method based on fuzzy rules and applicable to automated assessment. Expert Syst. Appl. **115**, 264 (2019)
12. Gambhir, M.; Gupta, V.: Recent automatic text summarization techniques: a survey. Artif. Intell. Rev. **47**(1), 1 (2017)
13. Cheng, J.; Lapata, M.: Neural summarization by extracting sentences and words (2016). arXiv preprint arXiv:1603.07252
14. Zhang, Y.; Er, M.J.; Zhao, R.; Pratama, M.: Multiview convolutional neural networks for multidocument extractive summarization. IEEE Trans. Cybern. **47**(10), 3230 (2016)
15. Yao, K.; Zhang, L.; Du, D.; Luo, T.; Tao, L.; Wu, Y.: Dual encoding for abstractive text summarization. IEEE Trans. Cybern. **50**, 985–996 (2018)
16. Rush, A.M.; Chopra, S.; Weston, J.: A neural attention model for abstractive sentence summarization (2015). arXiv preprint arXiv:1509.00685
17. Bahdanau, D.; Cho, K.; Bengio, Y.: Neural machine translation by jointly learning to align and translate (2014). arXiv preprint arXiv:1409.0473
18. See, A.; Liu, P.J.; Manning, C.D.: Get to the point: Summarization with pointer-generator networks (2017). arXiv preprint arXiv:1704.04368
19. Witte, R.; Bergler, S.: Fuzzy clustering for topic analysis and summarization of document collections. In: Conference of the Canadian Society for Computational Studies of Intelligence, pp. 476–488. Springer (2007)
20. Kyoomarsi, F.; Khosravi, H.; Eslami, E.; Davoudi, M.: Extraction-based text summarization using fuzzy analysis. Iran. J. Fuzzy Syst. **7**(3), 15 (2010)
21. Kiani-B, A.; Akbarzadeh-T, M.R.; Moeinzadeh, M.: Intelligent extractive text summarization using fuzzy inference systems. In: 2006 IEEE International Conference on Engineering of Intelligent Systems, pp. 1–4. IEEE (2006)
22. Kyoomarsi, F.; Khosravi, H.; Eslami, E.; Dehkordy, P.K.; Tajoddin, A.: Optimizing text summarization based on fuzzy logic. In: Seventh IEEE/ACIS International Conference on Computer and Information Science (icis 2008), pp. 347–352. IEEE (2008)
23. Suanmali, L.; Binwahlan, M.S.; Salim, N.: Sentence features fusion for text summarization using fuzzy logic. In: 2009 Ninth International Conference on Hybrid Intelligent Systems vol. 1, pp. 142–146. IEEE (2009)
24. Leite, D.; Rino, L.: genetic fuzzy automatic text summarizer Csbc 2009. Inf. Ufrgs. Br **2007**, 779 (2009)
25. Binwahlan, M.S.; Salim, N.; Suanmali, L.: Fuzzy swarm diversity hybrid model for text summarization. Inf. Process. Manag. **46**(5), 571 (2010)
26. Hannah, M.E.; Geetha, T.; Mukherjee, S.: Automatic extractive text summarization based on fuzzy logic: a sentence oriented approach. In: International Conference on Swarm, Evolutionary, and Memetic Computing, pp. 530–538. Springer (2011)
27. Suanmali, L.; Salim, N.; Binwahlan, M.S.: Fuzzy genetic semantic based text summarization. In: 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing, pp. 1184–1191. IEEE (2011)
28. Dixit, R.S.; Apte, S.: Improvement of text summarization using fuzzy logic based method. IOSRJCE **5**(6), 5 (2012)
29. Megala, S.S.; Kavitha, A.; Marimuthu, A.: Enriching text summarization using fuzzy logic. Int. J. Comput. Sci. Inf. Technol. **5**(1), 863 (2014)
30. Kumar, Y.J.; Salim, N.; Abuobieda, A.; Albaham, A.T.: Multi document summarization based on news components using fuzzy cross-document relations. Appl. Soft Comput. **21**, 265 (2014)
31. Megala, S.S.; Kavitha, A.; Marimuthu, A.: Text summarization system using fuzzy logic and conditional random field algorithm. Int. J. Comput. Sci. Inf. Technol. **1**(5), 863–867 (2015)
32. Babar, S.; Patil, P.D.: Improving performance of text summarization. Proc. Comput. Sci. **46**, 354 (2015)
33. Abbasi-ghalehtaki, R.; Khotanlou, H.; Esmaeilpour, M.: Fuzzy evolutionary cellular learning automata model for text summarization. Swarm Evolut. Comput. **30**, 11 (2016)
34. Chopade, H.A.; Narvekar, M.: Hybrid auto text summarization using deep neural network and fuzzy logic system. In: 2017 International Conference on Inventive Computing and Informatics (ICICI), pp. 52–56. IEEE (2017)
35. Lakshmi, S.S.; Rani, M.U.: Multi-Document Text Summarization Using Deep Learning Algorithm with Fuzzy Logic. In: 2018 IADS International Conference on Computing, Communications & Data Engineering (CCODE) (2018)
36. Sahba, R.; Ebadi, N.; Jamshidi, M.; Rad, P.: Automatic text summarization using customizable fuzzy features and attention on the context and vocabulary. In: 2018 World Automation Congress (WAC), pp. 1–5. IEEE (2018)
37. Uçkan, T.; Karcı, A.: Extractive multi-document text summarization based on graph independent sets. Egypt. Inf. J. 10.1016/j.eij.2019.12.002 (2020)
38. Hark, C.; Karcı, A.: Karcı summarization: A simple and effective approach for automatic text summarization using Karcı entropy. Inf. Process. Manag. **57**(3), 102187 (2020)
39. Steinberger, J.; Ježek, K.: Evaluation measures for text summarization. Comput. Inf. **28**(2), 251 (2012)
40. Lin, C.Y.: ROUGE: a package for automatic evaluation of summaries. In: Workshop on Text summarization branches out (WAS 2004) (2004)
41. Nallapati, R.; Zhai, F.; Zhou, B.: Summarunner: a recurrent neural network based sequence model for extractive summarization of documents. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
42. Paulus, R.; Xiong, C.; Socher, R.: A deep reinforced model for abstractive summarization (2017). arXiv preprint arXiv:1705.04304