International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11, No.1 (Jan-2022)

http://dx.doi.org/10.12785/ijcds/110130

Text Summarization on Telugu e-News based on Long-Short Term Memory with Rectified Adam Optimizer

Kishore Kumar Mamidala[1*] and Suresh Kumar Sanampudi[2]

[1] *Department of Computer Science and Engineering, Vivekananda Institute of Technology & Science, Karimnagar, India*
[2]*Department of Information Technology, Jawaharlal Nehru Technological University Hyderabad College of Engineering Jagtial, India*

Received 3 Dec. 2020, Revised 13 Jul. 2021, Accepted 31 Dec. 2021, Published 9 Jan. 2022

Abstract: Text summarization is a natural language processing method that reduces the text in the article and provides the important information from the document. Few researches have been carried out in the text summarization in the Telugu language and have lower efficiency in Telugu Text summarization. In this research, the Long-Short Term Memory (LSTM) with Rectified Adam Optimizer (RAdam) method and focal loss function is proposed for text summarization in Telugu e-news data. The Eenadu Telugu e-news data of various categories are collected to evaluate the performance of the proposed LSTM with RAdam method. Tokenization method is applied in the pre-processing method to extract the important keywords from the input data. Focal loss function is applied between the cells of LSTM to handle the imbalance data problem. Modulation function in the focal loss function down weight the easy examples to focus on hard examples and effectively handles the imbalance data. The proposed LSTM with RAdam method has advantage of using Exponential Moving Average (EMA) for adaptive learning rate and rectify the variance. The proposed LSTM with RAdam method is evaluated in 10 categories of e-news data to analysis the performance. This proposed LSTM with RAdam method has 93.46% accuracy and existing LSTM method has 86.92% accuracy.

Keywords: Eenadu Telugu e-News, Long-Short Term Memory (LSTM) with Rectified Adam Optimizer (RAdam), Telugu Language, Text Summarization and Tokenization.

1. INTRODUCTION

Automatic text summarization is the process of extracting the key sentences from input documents to effectively represent the document and this is a potential solution to information overloading problem. Moreover, the text summarization reduces the length of the input document with its overall meaning and information [1]. Initially, the text summarization process relies on the document frequency method to represent the document based on the repeated words. Text summarization methods have included with a variety of features and heuristics in the process of content selection [2]. Text summarization is classified into two types, namely abstractive or extractive. The abstract text summarization methods involve in generating the summarization of text using a text or sentences which is not present in the original documents. The extractive text summarization methods use only the word or sentences from the original document [3]. Single document summarization method generates the summarization for the single text document and multiple document summarization method has the capacity to generate the text summarization for more than one documents [4]. The statistical based method measures the weight of key terms like title, keywords, position, cue words and total weight of sentences is used to determine the sentence importance. The linguistic method analysis the term relationship in the document using extract meaningful sentence, grammar analysis, part-of-speech and thesaurus usage [5].

Recently, deep learning methods like LSTM gained more attention in many Natural Language Processing (NLP) tasks such as text classification, sentiment analysis, language understanding and language translation. Followed by the success of achieving higher performance of Deep learning in NLP applications, many researches are applying deep learning in text summarization to analysis its capacity [6, 7, 8]. Text summarization state-of-art method depends on the machine learning and deep learning techniques. This kind of Text summarization method analysis the salient features of sentences in the document and extract the sentence of most salient value. As the deep learning methods are not involving in

E-mail: kishore.mamidala@gmail.com, sureshsanampudi@jntuh.ac.in

paraphrasing process and sentence construction, this have the advantage of language independent [9, 10]. LSTM method has the limitation of poor learning rate that affects the performance of the model.

In this study, the LSTM with RAdam is proposed for the text summarization in the Telugu language. The RAdam method is applied to select the optimal learning rate for the LSTM. Focus loss function is applied to handle the imbalance data problem in the summarization process. The Eenadu e-news data were collected to evaluate the performance of the proposed LSTM with RAdam method. Tokenization method is applied to extract the important information from input data. The LSTM with RAdam method performs text summarization in the tokenized data. The existing method has the lower efficiency in imbalance data. The proposed model applies focal loss function to handle the imbalance data. The objective of this research is presented as follows:

1. The LSTM with RAdam method is proposed in the tokenized data to perform text summarization in Telugu e-news data. The LSTM with RAdam method has the advantage of using the EMA method for adaptive learning rate and rectify the variance.
2. The Eenadu Telugu e-news data were collected to analysis the efficiency of proposed LSTM with RAdam in text summarization. The 10 categories of Telugu news were collected to analysis the efficiency.
3. The proposed LSTM with RAdam method is compared with LSTM and RNN to analysis the performance. The proposed LSTM with RAdam method is also compared with LSTM with Adam optimizer method.
4. Focal loss function applies between each cell of LSTM to handle the imbalance data problem in the classification.
5. Focal loss function classifies and balance the positive/negative examples in the class. Focal loss function reduces the down weight of easy examples and focus on hard negative in training process.

The formulation of this research paper is given as follows: text summarization researches are reviewed in the section 2, the tokenization and LSTM with RAdam method is explained in section 3, the experimental design information is given in section 4, the proposed LSTM with RAdam method performance is analysed in section 5 and conclusion is in section 6.

2. LITERATURE REVIEW

Text summarization is the process of reducing the text in article to extract the useful information from the articles. Many text summarization researches have been carried out to provide useful information in English Language and notable researches on the text summarization are reviewed in this section.

Naidu et al. [11] developed automatic keyword extraction algorithm based on Telugu POS tagging method for text summarization in Telugu Language. The Telugu e-newspaper datasets were used to evaluate the performance of this method. The probabilistic distribution was measured to train the model. The result shows developed method has higher efficiency in the text summarization in Telugu news articles. The feature selection and classification method can be used for the effective text summarization.

Elbarougy et al. [12] proposed graph based method for Arabic text summarization that represents the document sentences as a vertices and graph. A Modified PageRank method with initial score was applied to each node, which represents the number of nouns in the sentences. The cosine similarity was used to find the summarization based on the similarity of various sentences. The developed algorithm solved the problem of finding the noun in the sentences in Arabic language that suffered from the absence of capital and small letters. The developed method applied Modified PageRank algorithm in multiple iteration to improve performance of summaries generation. The feature selection method can be applied to increase the text summarization efficiency.

Gomez et al. [13] proposed Multi-Objective Artificial Bee Colony (MOABC) method for the automatic text summarization. The Document Understanding Conference (DUC) dataset is applied to analysis the efficiency of the method. This MOABC method has been applied for multi-document summarization and compared with existing methods. The exploitation process of MOABC method is needed to improve the efficiency of text summarization.

Diao et al. [14] proposed Hierarchical Attention based method with contextualized representation for the text summarization process. The contextual information is applied developed Hierarchical model to increases the regression in sentence for text summarization. The Bi-GRU is applied to analysis the word level and sentence level contextual relation and the contextual representation was allowed across linguistic context information. This analysis shows that the CRHASum method has the higher efficiency compared to existing method in text summarization. The developed model has lower efficiency in analysis the upper bound analysis and this requires to leverage more semantic information.

Radaideh and Bataineh [15] proposed hybrid single document text-summarization method. The approach consists of statistical features, domain knowledge and genetic algorithm to select the important points in Arabic political document. The two datasets such as KALIMAT and Essex were used to analysis the efficiency of developed method. The summary of the text is created based on the shortest path of the first and last sentences. The limitation of the method is summary first and last sentence must be similar to document first and last sentence that does not provide optimal solutions.

Liu *et al*. [16] proposed Rectified Adam (RAdam) method for rectify the variance of the adaptive learning rate. The developed model is evaluated for image classification, language modeling and neural network analysis. The RAdam method has the advantage of effectively update the learning rate of the classifier model based on the input data. The model test accuracy is low and this requires the more number of sample data for effective classification that tends to low performance in imbalance data.

A. Problem Definition and Solution

Many existing Text summarization methods failed to analyse the past and future data. Some of the existing methods select the irrelevant features for the Text summarization, which affects the overall performance of respective method. Many existing methods have been applied for the Text summarization in English language and only fewer studies is applied for the Telugu language. Some studies involve in applying the LSTM network for the Text summarization and achieves considerable performance. The learning rate and momentum parameter of LSTM network is needed to be optimized for effective performance. Telugu language has rich agglutinative characteristics and make it difficult for Text summarization.

Solutions: This research involves in applying the LSTM with RAdam Optimization method for the Text summarization of Telugu e-news data. The LSTM method has the advantage of analysis the past and future data for the Text summarization and it also holds the important feature for long duration. The RAdam optimization method optimizes the learning rate and moment value of the LSTM network to increases the efficiency of analysis in Telugu summarization. The proposed LSTM with RAdam method has the capacity to handle rich agglutinative characteristics for Text summarization.

3. PROPOSED METHOD

Text summarization is the process of reducing the text in the article to extract the important information. Few researches were carried out in the text summarization in the Telugu language and have lower performance. The LSTM with RAdam method is proposed for text summarization in Telugu language. The collected Eenadu Telugu e-news data were applied to analysis the efficiency of proposed LSTM with RAdam method.

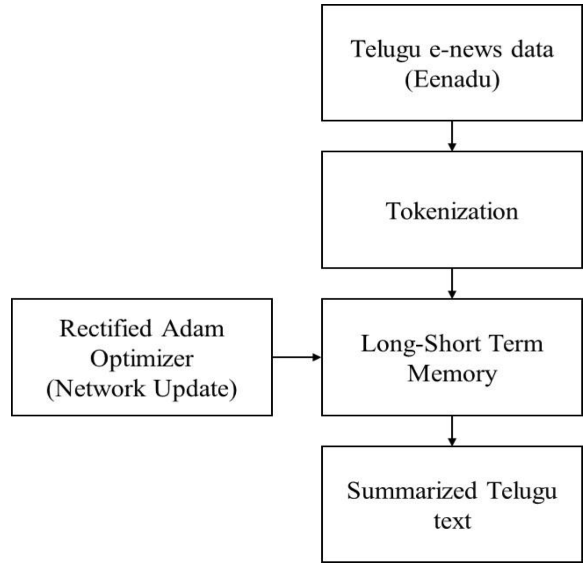

Figure 1. The block diagram of proposed LSTM with RAdam method

Tokenization method is applied in the input data to extract the important keywords from the information. The LSTM with RAdam method performs text summarization in the tokenized data. The block diagram of the proposed LSTM with RAdam method in text summarization is shown in Figure 1.

A. Tokenization

Tokenization method separates the input sentences from the dataset into small tokens and the input sentence may contain a word-space and delimiters. In this method, N-gram tokenizer is employed to eliminate the word-spaces and delimiters.

For example, a sentence from input data such as

"ఓపెనర్ రోహిత్‌శర్మ ఓ రికార్డును ఖాతాలో వేసుకున్నాడు. ఓపెనర్‌గా అతి తక్కువ ఇన్నింగ్స్‌ల్లో 7000 పరుగులు పూర్తి చేసుకున్న బ్యాట్స్‌మన్‌గా అతడు (137 ఇన్నింగ్స్) ఘనత సాధించాడు. దక్షిణాఫ్రికా ఓపెనర్ హాషీమ్ ఆమ్లా (147 ఇన్నింగ్స్) పేరిట ఉన్న రికార్డును బద్దలు కొట్టాడు. ఈ జాబితాలో సచిన్ (160), దిల్షాన్ (165) తర్వాతి స్థానాల్లో ఉన్నారు.".

After N-gram tokenizer is applied in input sentence, the result is obtained as

'▁ఓపెనర్', '▁రోహిత్', '▁శర్మ', '▁ఓ', '▁రికార్డు', 'ను', '▁ఖాతాలో', '▁వేసుకున్నాడు', '.', '▁ఓపెనర్', '▁గా', '▁అతి', '▁తక్కువ', '▁ఇన్నింగ్స్', '▁ల్లో', '▁70', '00', '▁పరుగులు',

'▁పూర్తి', '▁చేసుకున్న', '▁బ్యాట్', 'స్', '▁మన్', '▁గా', '▁అతడు', '▁(13', '7', '▁ఇన్నింగ్స్', '▁)', '▁ఘనత', '▁సాధించాడు', '.', '▁దక్షిణాఫ్రికా', '▁ఓపెనర్', '▁హష', 'ీమ్', '▁ఆ', 'మ్', 'లా', '▁(14', '7', '▁ఇన్నింగ్స్', '▁)', '▁పేరిట', '▁ఉన్న', '▁రికార్డు', 'ను', '▁బద్దలు', '▁కొట్టాడు', '.', '▁ఈ', '▁జాబితాలో', '▁సచిన్', '▁(16', '0),', '▁దిల్', 'షాన్', '▁(16', '5)', '▁తర్వాత', 'ీ', '▁స్థానాల్లో', '▁ఉన్నారు', '.'.

For further analysis, the result is kept in the form of an array.

B. Long-Short Term Memory with Rectified Adam Optimizer

Text summarization process requires the latest data and previous data, so LSTM network is applied in this text summarization process. The hidden layer of LSTM self-feedback method has the advantage of handling the long-term dependence problems [16], and the memory cell in LSTM unit stores the information that are updated by three gates namely input, output and forget gate [17, 18, 19]. The basic structure of LSTM is shown in Figure 2.

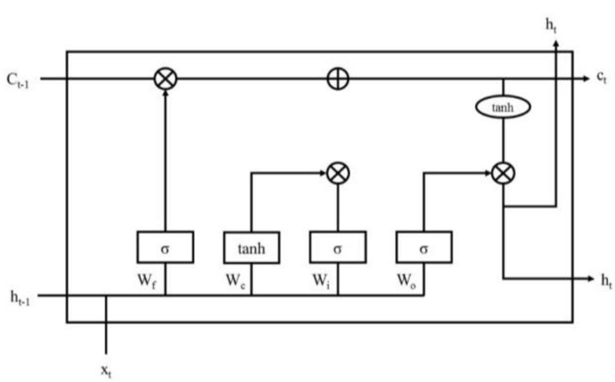

Figure 2. The basic structure of LSTM

Telugu input text data is denoted as x_t at time t in the LSTM cell and LSTM cell output cell is denoted as h_{t-1} at the previous moment. The memory cell value is C_t, the LSTM cell output is h_t. The working of LSTM is explained into following steps.

1. The candidate memory cell $\tilde{c}_t$ is calculated and the weight matrix W_c, then the bias b_c is calculated, as shown in Eq. (1).

$$C_t = \tanh\left(W_c.[h_{t-1}, x_t] + b_c\right) \quad (1)$$

2. The input gate i_t is calculated, the input gate controls the memory cell state value of current input data. The sigmoid function is denoted as σ, the weight matrix is denoted as W_i, and the bias is denoted as b_i, as shown in Eq. (2).

$$i_t = \sigma\left(W_i.[h_{t-1}, x_t] + b_i\right) \quad (2)$$

3. The forget gate f_t value is calculated and the memory cell state value of historical data is controlled by the forget gate, the weight matrix is denoted as W_f, the bias is denoted as b_f, as shown in Eq. (3).

$$f_t = \sigma\left(W_i.[h_{t-1}, x_t] + b_f\right) \quad (3)$$

4. The current moment memory cell value c_t and the last LSTM unit state value c_{t-1} is calculated, as shown in Eq. (4).

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (4)$$

Where "$*$" is the dot product. The memory cell update is based on the candidate cell and last cell state value, and input and forget gate is used to control memory cell.

5. The output gate o_t value is calculated, the memory cell state value is controlled by the output gate. The weight matrix is denoted as W_0, the bias is denoted as b_0 and sigmoid function is denoted as σ, as given in Eq. (5).

$$o_t = \sigma\left(W_0.[h_{t-1}, x_t] + b_0\right) \quad (5)$$

6. Finally, the LSTM unit h_t is calculated based on *tanh* activation function, as shown in Eq. (6).

$$h_t = o_t * \tanh\left(c_t\right) \quad (6)$$

The LSTM uses the memory cell and three control gates to store, reset, read and update the long-time information easily. The weight matrix dimension is adjusted to change the output dimension due to the LSTM internal parameters sharing mechanism. LSTM applies long-time delay between the input and feedback. This architecture memory cell state value maintains continuous error flow and this cause gradient neither explode nor explode. The current momentum memory cell c_t and learning rate of LSTM network is optimized using Rectified Adam Optimizer. The objective function θ of RAdam provide optimal value of learning rate W_i of LSTM network.

1) Focal Loss function

The Focal loss function is applied between each cell in the proposed LSTM with RAdam. The Focal loss function is developed for text summarization process to handle the imbalance data in the classes during training. The focal loss function is applied based on the Cross Entropy (CE) for binary classification, as shown in Eq. (7).

$$CE(c_t, y) = \begin{cases} -\log(c_t) & if\ y = 1 \\ -\log(1-c_t) & otherwise \end{cases} \quad (7)$$

The ground truth is denoted as $y \in \{\pm 1\}$ and probability estimation of model's $p \in [0, 1]$ for the class with label $y = 1$, defined as c_t, as shown in Eq. (8).

$$\mathrm{c_t} = \begin{cases} c_t & if\ y = 1 \\ 1-\mathrm{c_t} & otherwise \end{cases} \quad (8)$$

And rewrite $CE(c_t, y) = CE(c_t) = -\log(\mathrm{c_t})$

If loss function summed over a large number of easy examples, the small loss values can overwhelm the rare class.

Balanced Cross entropy

Applying a weighting factor $\alpha \in [0,1]$ for class 1 and $1-\alpha$ for class -1 is a common method for addressing class imbalance problem. The inverse class frequency is denoted as α and cross-validation is set to teat as a hyperparameter. The c_t is denoted with α_t and α-balanced CE loss as shown in Eq. (9).

$$CE(c_t) = -\alpha_t \log(c_t) \quad (9)$$

This loss is the extension to CE and in the focal loss, this is considered as experimental baseline.

Focal Loss Definition

The large class imbalance data in the training process overwhelms the cross entropy loss. The majority of loss is comprising in easily classified negatives and dominate the gradient. The α balance the positive/negative examples and this doesn't differentiate easy/hard examples. The reshape loss function is applied to provide less weight for easy examples and focus on hard negatives in training.

The cross entropy with tunable focusing parameter $\gamma \geq 0$ is added with Modulating factor $(1-\mathrm{c_t})^\gamma$. The focal loss is shown in Eq. (10).

$$FL(c_t) = -(1-c_t)^\gamma \log(c_t) \quad (10)$$

The focal loss is analysed for several value of $\gamma \in [0,5]$ and two characteristics of focal loss is analysed. If c_t is small and sample is misclassified, then the modulating factor is near 1 and the loss is unaffected. The factor goes to 0 if the $c_t \to 1$, and well-classified samples loss is down-weighted. The focusing parameter γ smoothly adjusts to down-weighted easy examples. If $\gamma = 0$, Focal Loss is equal to CE, and increases in γ is likely to increases the modulating factor effect.

For easy samples, the loss reduces for the modulating factor and increases the range to reduce loss value. If $\gamma = 2$ and $c_t = 0.9$, this would have $100\times$ lower loss compared to CE and if $c_t \approx 0.968$, this would have $1000\times$ lower loss. This helps to reduces the misclassification.

In practice, an α-balanced variant of the focal loss as shown in Eq. (11).

$$FL(c_{t+1}) = -\alpha_t(1-c_t)^\gamma \log(c_t) \quad (11)$$

This model is applied in RAdam-LSTM model to handle the imbalance data.

2) Rectified Adam Optimizer

Adam method selects the learning rate parameter adaptively. Similar to RMSprop and Adadelta, the Adam method stores a past squared gradient v_t with exponentially decaying average. In additional, the Adam method stores past gradients m_t and exponentially decaying average as shown in eq. (12-14). Similar to momentum, the gradients are computed with respect to the stochastic object at timestep t.

$$g_t \leftarrow \nabla_\theta f_t(\theta_{t-1}) \quad (12)$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \quad (13)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \quad (14)$$

The m_t and v_t denotes the first and last moment of the gradient, respectively. Adam are biased toward zero as m_t and v_t are set as of 0's vectors. If decay rate and initial time steps are zero(i.e. β_1 and β_2 are close to 1), the Adam are biased towards zero. In order to counteract these biases, the bias-corrected first and second moment estimates are computed, as shown in Eq. (15) & (16).

$$m_t^{new} = \frac{m_t}{1-\beta_1^t} \quad (15)$$

$$v_t^{new} = \frac{v_t}{1-\beta_2^t} \quad (16)$$

The exponential moving average (EMA) can be interpreted as an approximation to the simple moving average (SMA) [20], as shown in Eq. (17).

$$\rho\left(\frac{1-\beta_2 \sum_{i=1}^{t} \beta_2^{t-i} g_i^2}{1-\beta_2^t}\right) \approx \rho\left(\frac{\sum_{i=1}^{f(t,\beta_2)} g_{t+1-i}^2}{f(t,\beta_2)}\right) \quad (17)$$

Where the degree of freedom is denoted as ρ and estimation of ρ at t for quantitative analysis.

Similar to RMSprop and Adadelta, the parameters are updated that yields the Adam update rule, as shown in Eq. (18).

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{v^{new} + \in}} m_t^{new} \quad (18)$$

The default values of Adam are 0.9 for β_1, 0.999 for β_2, and 10−8 for $\in$. Adam shows the higher performance in practice compares to other adaptive learning-method analyzed [19]. The pseudo code for the LSTM-RNN with Adam Optimizer is shown as follows:

Pseudo Code: Algorithm: LSTM Algorithm with Rectified Adam Optimizer

1. Class LSTM-RNN [Telugu e-news data, _weights, _biases]

2. Collection of Telugu e-News data

3. $D=\{x_{i,n}, y_n \mid i \in f \; and \; n \in N\}$

//Database with N instances and f features

4. X_{train}, X_{test}, Y_{train}, Y_{test} ⟵ train_test_split(Telugu e-News data,test_size=0.3) // spitting the data into training data and testing data for the LSTM-RNN model

5. def model**:**

While θ_t is not converged **do**

6. $m_0, v_0 \leftarrow 0,0$ (Initialize moving 1st and 2nd moment)

7. $\rho_\infty \leftarrow 2/(1-\beta_2)-1$ (compute the maximum length of the approximated SMA)

8. while $t=\{1,...,T\}$ do

a. $g_t \leftarrow \Delta_\theta f_t(\theta_{t-1})$ (Stochastic Gradients calculation)

b. $v_t \leftarrow \beta_2 v_{t-1} + (1-\beta_2) g_t^2$ (Update exponential moving 2nd moment)

c. $m_t \leftarrow \beta_1 m_{t-1} + (1-\beta_1) g_t$ (Update exponential moving 1st moment)

d. $m_t \leftarrow m_t / (1-\beta_1^t)$ (Compute bias-corrected moving average)

e. $\rho_t \leftarrow \rho_\infty - 2t\beta_2^t / (1-\beta_2^t)$ (Compute the length of the approximated SMA)

f. if the variance is tractable, i.e., $\rho_t > 4$ then

i. $v_t \leftarrow \sqrt{\frac{v_t}{1-\beta_2^t}}$ (Compute bias-corrected moving 2nd moment)

ii. $r_t \leftarrow \sqrt{\frac{((\rho(t-4))(\rho_{t-2})\rho_\infty)}{(\rho_{\infty-4})(\rho_{\infty-2})\rho_t}}$

$\theta_t \leftarrow \theta_{t-1} - \alpha_t r_t m_t / v_t$ (Update parameters with adaptive momentum)

g. else

h. $\theta_t \leftarrow \theta_{t-1} - \alpha_t m_t$

i. (Update parameters with un-adapted momentum)

End while

Return; Resulting parameter θ_t

4. EXPERIMENTAL DESIGN

Text summarization is the process of providing the useful information from many documents to save user effect to find the relevant information. This research involves in applying the LSTM with RAdam optimizer for effective text summarization of Telugu e-Newspaper data. This section provides the detailed description about the dataset, metrics used, system requirement and parameter settings.

Figure 3. Eenadu Website Homepage

Dataset: Eenadu e-news data were collected using fastAPI to evaluate the performance of LSTM with RAdam in Text summarization. The size of Eenadu collected data is 0.91 GB and proposed method is evaluated in the data. The 150 articles with 1223 keywords were collected to evaluate the performance of the model. The Eenadu website homepage samples is presented in Figure 3.

Metrics used: The accuracy, precision, recall and F-Score were shown in Eq. (19-22) and the metrics used to analyse the efficiency of the proposed LSTM with RAdam method in Text summarization.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{19}$$

$$precision = \frac{TP}{TP + FP} \tag{20}$$

$$Recall = \frac{TP}{TP + FN} \tag{21}$$

$$F - Score = \frac{2TP}{2TP + FP + FN} \tag{22}$$

Parameter Settings: The epochs of the LSTM network is set as 1e-05, 6 layers are used in LSTM and dropout value is set as 0.2. The ReLU network layer is used with momentum 0.1.

System Requirement: The proposed LSTM with RAdam method is implemented in the system consists of Intel i7 process with 8 GB of RAM and is implemented in the Python 3.7 with keras library.

5. Experimental Results

Text summarization is the process of reducing the text in the article to provide useful information of the article. The LSTM with RAdam method is proposed for text summarization in Telugu language. The Eenadu Telugu e-news data is used to analysis the efficiency of the LSTM with RAdam in Text summarization. The tokenization method is applied to extract the important keywords from the data. The LSTM with RAdam method perform text summarization in input data. The following section provides the detailed information about the performance of the proposed LSTM with RAdam method in text summarization.

Sample text input to the proposed Model:

ౖ ౕ ౢ ,ϴ ͱ ̎ ు ౘ :Ç ౕ ʳ ̍ ѧ ƈ ̍ ̒ Ʈ _ ౘ ä ǀ ɹ ,̳ ƿ ᅢ ƙǄᏰe ª ⅄ ̬
ⁿ ϶ Ĕe| ᏰϴͰ ̎ ు ౘ ᵮƐɹ ʞ Ā (Ψ Ǵɥƃ)ᏰƇ ᆺdzᆺ̿ .ᅧ ᆺᅵ Ţ ᏰˀĐZƙ̧ ̈3
⅄ ƽ Ʈ ౖ ౕ ̍ ̊æ >϶ ꓷ ౕe| ̈3_ ౘ ᆺ̿ .

Summarized Output:

̳ ƿ ᅢ ƙǄᏰe ª ⅄ ̬

Figure 4. Sample input to the proposed model and summarized output

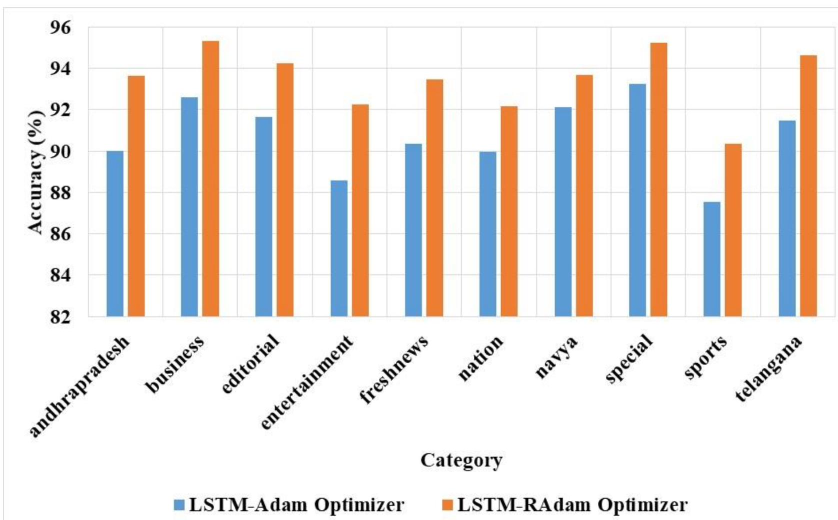

Figure 4. The accuracy of the proposed LSTM with RAdam method

The sample input text of the Eenadu data and the summarized output of the proposed model is shown in Figure 4. The sample input data denotes that the “Due to the corona virus, all league match are postponed” and the proposed model output denotes “All league match are postponed”.

The accuracy of the proposed LSTM with RAdam method is measured for the text summarization in Telugu e-news data as shown in Figure 5. The proposed LSTM with RAdam optimizer is compared with Adam optimizer in Text summarization. The proposed LSTM with RAdam method is analysed on the 10 category of Telugu e-news. The proposed LSTM with RAdam has the higher accuracy compared with LSTM with Adam optimizer. The proposed LSTM with RAdam method has the advantage of using EMA to measure the adaptive learning rate and rectify the variance. The proposed LSTM with RAdam method has the accuracy of 90.34% and LSTM with Adam optimizer has 87.56% accuracy.

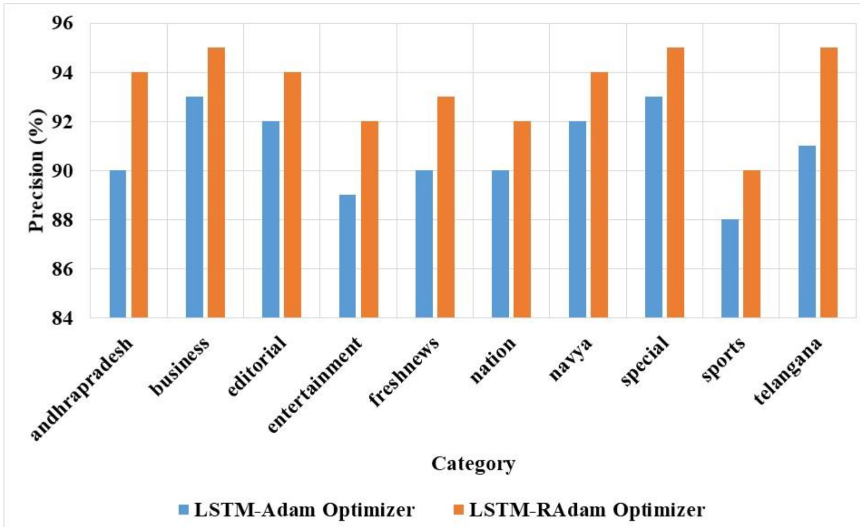

Figure 5. The precision of the proposed LSTM with RAdam method

The precision is measured for the proposed LSTM with RAdam method in Text summarization as shown in Figure 6. The proposed LSTM method with RAdam method is analysed for 10 category of Telugu e-news data. The proposed LSTM with RAdam has higher precision value compared to LSTM with Adam optimizer method in Text summarization. The proposed LSTM with RAdam method has the advantage of using EMA for measure adaptive learning and rectify the variance. The proposed LSTM with RAdam method has the precision of 95% and LSTM with Adam optimizer has 91% precision in Telangana news category.

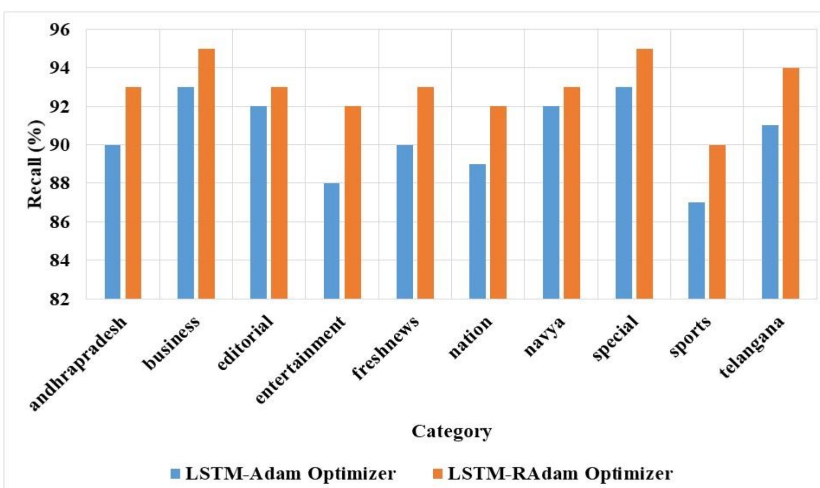

Figure 6. The Recall of the proposed LSTM with RAdam method

The recall value is measured for the proposed LSTM with RAdam method in Text summarization in Telugu e-news data as shown in Figure 7. The proposed LSTM with RAdam method is compared with LSTM with Adam optimizer in 10 categories of Telugu e-news data. The proposed LSTM with RAdam has higher recall value than LSTM with Adam optimizer. The proposed LSTM with RAdam has higher recall value due to the EMA is used for adaptive learning rate and rectify the variance. The proposed LSTM with RAdam method has the recall value of 95% and LSTM with Adam optimizer has 93% in business news category.

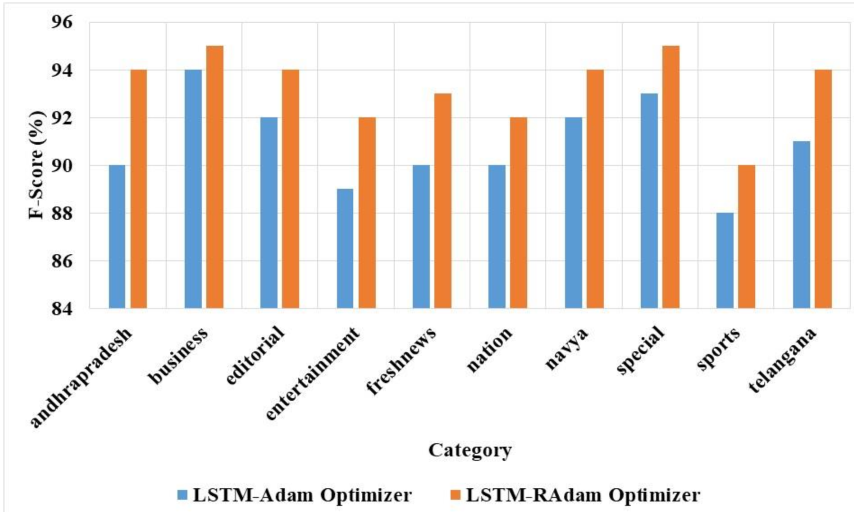

Figure 7. The F-Score of the proposed LSTM with RAdam method

The F-Score is measured for the proposed LSTM with RAdam method for the text summarization in Telugu e-news data as shown in Figure 8. The proposed LSTM with RAdam method is evaluated in 10 categories of Telugu e-news data. The proposed LSTM with RAdam has higher F-Score compared with LSTM with Adam optimizer. The LSTM with RAdam method has the advantage of using EMA method for adaptive learning rate and rectify the variance. The proposed LSTM with RAdam method has the F-Score of 94% and LSTM with Adam optimizer has 91% in Telangana news category.

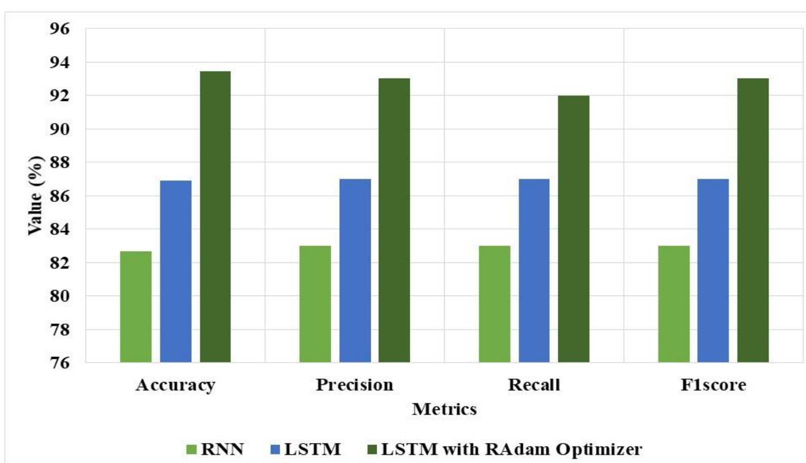

Figure 8. Comparative analysis of the proposed LSTM with RAdam method

TABLE I. COMPARATIVE ANALYSIS

Method	Accuracy (%)	Precision (%)	Recall (%)	F1score (%)
RNN	82.67	83	83	83
LSTM	86.92	87	87	87
LSTM with RAdam Optimizer	93.46	93	92	93

TABLE II. AVERAGE RESPONSE TIME OF THE PROPOSED MODEL

Method	Category	Length	Response Time
LSTM ADAM	Andhrapradesh	100	0.41
		200	0.7
		300	0.93
	Business	100	0.49
		200	0.67
		300	0.98
	Editorial	100	0.32
		200	0.65
		300	0.98
	Entertainment	100	0.39
		200	0.62
		300	0.97
	Freshnews	100	0.39
		200	0.6
		300	0.99
	Nation	100	0.48
		200	0.7
		300	0.95
	Navya	100	0.44
		200	0.69
		300	0.97
	Special	100	0.5
		200	0.62
		300	1
	Sports	100	0.43
		200	0.69
		300	0.9
	Telangana	100	0.5
		200	0.66
		300	0.71
LSTM RADAM	Andhrapradesh	100	0.1
		200	0.57
		300	0.78
	Business	100	0.08
		200	0.58
		300	0.79
	Editorial	100	0.03
		200	0.54
		300	0.7
	Entertainment	100	0.09
		200	0.6
		300	0.78
	Freshnews	100	0.13
		200	0.52
		300	0.8
	Nation	100	0.04
		200	0.54
		300	0.7
	Navya	100	0.09
		200	0.54
		300	0.79
	Special	100	0.12
		200	0.59
		300	0.73
	Sports	100	0.07
		200	0.57
		300	0.8
	Telangana	100	0.18
		200	0.55
		300	0.72

TABLE III. COMPARATIVE ANALYSIS IN TELUGU TEXT SUMMARIZATION

Metrics	Keyword extraction algorithm [11]	Proposed LSTM-RAdam with Focal loss function
Accuracy (%)	90.7	93.46
Precision (%)	79.5	93
Recall (%)	76.1	92
F-Score (%)	77.7	93

The LSTM with RAdam is compared with standard LSTM and RNN method as shown in Figure 9. The comparison shows that LSTM with RAdam method has higher efficiency compared with LSTM and RNN method. The proposed LSTM with RAdam method has the advantage of using EMA method for adaptive learning rate and rectify the variance. The proposed LSTM with RAdam method has the accuracy of 93.46% and LSTM method has 86.92% accuracy.

The comparative analysis of the proposed LSTM-RAdam method with LSTM and RNN method is shown in Table I. The comparison shows that the proposed LSTM with RAdam method has higher efficiency than LSTM and RNN method. The proposed LSTM with RAdam method has the accuracy of 93.46% and LSTM method has 86.92% accuracy. This shows that the proposed LSTM with RAdam method has the higher performance in Text summarization of Telugu e-news.

The average response time of the proposed and LSTM Adam model in Telugu text summarization is shown in Table II. The result shows that proposed LSTM-RAdam with Focal loss function has lower average response time compared to existing LSTM-Adam method in text summarization. The proposed model eliminates the process of optimization for class with more positive samples, while existing model process learning rate optimization for large data.

The proposed LSTM-RAdam with focal loss function is compared with existing keyword extraction algorithm [11] in Telugu text summarization, as shown in Table III. The result shows that proposed model has higher performance compared to existing keyword extraction algorithm. The proposed model has the advantage of adaptive learning rate for the LSTM model and efficient in handling imbalance dataset. The proposed LSTM-RAdam with Focal loss function has accuracy of 93.46 % and existing keyword extraction algorithm has 90.7 % accuracy.

6. CONCLUSION

Text summarization is the process of reducing the number of text in the article to extract important information from the article. The LSTM with RAdam method is proposed for the text summarization in Telugu e-news. The Eenadu Telugu e-news data were collected to evaluate the performance of the proposed LSTM with RAdam method. Tokenization method is applied to extract the important keywords from the input data. The LSTM with RAdam method is applied for text summarization in the tokenized data. The LSTM with RAdam method has the advantage of using EMA method for adaptive learning rate and rectify the variance. The proposed LSTM with RAdam method has been evaluated for 10 News categories. The proposed LSTM with RAdam has higher performance compared to RNN, LSTM and LSTM with Adam optimizer in Telugu Text Summarization. The result shows that the proposed LSTM with RAdam has the 93.46% accuracy and existing LSTM has 86.92% accuracy. The future work of the proposed method involves in analysing the multiple document of news.

REFERENCES

[1] M. Mohd, R. Jan, and M. Shah, “Text document summarization using word embedding,” Expert Syst. Appl. vol. 143, pp. 112958, April 2020.

[2] M. Moradi, G. Dorffner, and M. Samwald, “Deep contextualized embeddings for quantifying the informative content in biomedical text summarization,” Comput. Methods Programs Biomed. vol. 184, pp. 105117, Feburary 2020.

[3] J.M. Sanchez-Gomez, M.A. Vega-Rodríguez, and C.J. Perez, “Experimental analysis of multiple criteria for extractive multi-document text summarization,” Expert Syst. Appl. vol. 140, pp. 112904, Feburary 2020.

[4] V. Gupta, and N. Kaur, “A novel hybrid text summarization system for Punjabi text,” Cognit. Comput. Syst. vol. 8(2), pp. 261-277, April 2016.

[5] M.A. Tayal, M.M. Raghuwanshi, and L.G. Malik, “ATSSC: Development of an approach based on soft computing for text summarization,” Comput. Speech Lang. vol. 41, pp. 214-235, January 2017.

[6] A. Joshi, E. Fidalgo, E. Alegre, and L. Fernández-Robles, “SummCoder: An unsupervised framework for extractive text summarization based on deep auto-encoders,” Expert Syst. Appl. vol. 129, pp. 200-215, September 2019.

[7] J.M. Sanchez-Gomez, M.A. Vega-Rodríguez, and C.J. Pérez, “Parallelizing a multi-objective optimization approach for extractive multi-document text summarization,” J. Parallel Distrib. Comput. vol. 134, pp. 166-179, December 2019.

[8] S. Hou, and R. Lu, “Knowledge-guided unsupervised rhetorical parsing for text summarization,” Information Systems, vol. 94, pp. 101615, December 2020.

[9] B. Mutlu, E.A. Sezer, and M.A. Akcayol, “Candidate sentence selection for extractive text summarization,” Information Processing & Management, vol. 57(6), pp. 102359, November 2020.

[10] Z. Liang, J. Du, and C. Li, “Abstractive Social Media Text Summarization using Selective Reinforced Seq2Seq Attention Model,” Neurocomputing, May 2020.

[11] R. Naidu, S.K. Bharti, K.S. Babu, and R.K. Mohapatra, “Text summarization with automatic keyword extraction in telugu e-newspapers,” In Smart Computing and Informatics, pp. 555-564, March 2017.

[12] R. Elbarougy, G. Behery, and A. El Khatib, “Extractive Arabic Text Summarization Using Modified PageRank Algorithm,” Egyptian Informatics Journal, vol. 21(2), pp. 73-81, July 2020.

[13] J.M. Sanchez-Gomez, M.A. Vega-Rodríguez, and C.J. Pérez, "Extractive multi-document text summarization using a multi-objective artificial bee colony optimization approach," Knowledge Based Syst. vol. 159, pp. 1-8, November 2018.

[14] Y. Diao, H. Lin, and L. Yang, "CRHASum: extractive text summarization with contextualized-representation hierarchical-attention summarization network," Neural Comput. Applic. vol. 32, pp. 11491–11503, January 2020.

[15] Q.A. Al-Radaideh, and D.Q. Bataineh, "A hybrid approach for arabic text summarization using domain knowledge and genetic algorithms," Cognit. Comput. Syst. vol. 10(4), pp. 651-669, August 2018.

[16] L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, and J. Han, "On the variance of the adaptive learning rate and beyond," arXiv preprint arXiv:1908.03265, August 2019.

[17] J. Cao, Z. Li, and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," Physica A: Statistical Mechanics and its Applications, vol. 519, pp. 127-139, April 2019.

[18] O. Yildirim, U.B. Baloglu, R.S. Tan, E.J. Ciaccio, and U.R. Acharya, "A new approach for arrhythmia classification using deep coded features and LSTM networks," Comput. Methods Programs Biomed. vol. 176, pp. 121-133, July 2019.

[19] Y. Liu, X. Zhang, F. Huang, X. Tang, and Z. Li, "Visual question answering via Attention-based syntactic structure tree-LSTM," Applied Soft Computing, vol. 82, pp. 105584, September 2019.

[20] D.P. Kingma, and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, December 2014.

Kishore Kumar Mamidala is Associate Professor in the Department of Computer Science at Vivekananda Institute of Technology & Science, Karimnagar since June 2008. Prior to this he was an Assistant Professor in the CSE department at Dr.VRK College of Engineering since 2006. He received B.Tech in CSE from JPNCE (affiliated to JNTUH), M.Tech. in Computer Science from JBIET (affiliated to JNTUH) in 2008, and pursuing Ph.D. in Computer Science from the JNTU, Hyderabad.

Mr Kishore Kumar has been involving in research activities to develop the systems which can detect useful, new, and timely sentence-length updates about a developing event. In Temporal Summarization, he works on extracting key sentences to produce the automatic text summaries of the input documents, and hence is a potential solution to the information overload problem. He has published 10 research articles and also a book chapter on his research activities titled " Mining Behavioural Data: Data Mining Technologies using Machine Learning Algorithms".

Suresh Kumar Sanampudi is presently working as Assistant Professor & Head in the Department of Information Technology, Jawaharlal Nehru Technological University Hyderabad. His research interests include Artificial Intelligence, Natural Language Processing, Information Retrieval, Algorithms and Information Security.