Summarizing Product Reviews Using NLP Based Text Summarization

Ravali Boorugu, Dr. Gajula Ramesh, Dr. Karanam Madhavi

Abstract: Shopping was confined just to outdoor shopping few years ago when there were no websites for online shopping and no internet. But now Internet is available to everyone at fingertips with the advent of smartphones, tablets, laptops and even the cheaper rate to afford internet. This was the prime reason for the sudden booming of online shopping websites. Nowadays everyone loves online shopping. Everyone wishes to order products rather than buying directly from the shops. The primary thing a person will check before ordering the product is a review given by the customers who bought it already. It is becoming difficult for a user to go through various reviews of different products of a particular type and choose the best among them. Thus the need arises for the summarization of these reviews to the maximum extent possible, in order to make the user choose the best product from the whole lot. The process of minimizing the content of a given document without any loss in the meaning of the content is called as Text Summarization. It is grabbing attention of many NLP Researchers nowadays. Text Summarization is categorized based on Input type, Output type and Purpose. We will discuss in brief the various types of text summarization in detail in this paper. We propose seq2seq model for summarization. Its advanced version i.e LSTM is used along with the attention mechanism for increased accuracy. We used the latest word embedding model Conceptnet Numberbatch which is very much similar to GloVe but comparatively better than that. During classification we use 1D convolutional layer followed by max pooling layer, LSTM layer and then at the end by a fully connected layer.

Index Terms: Attention Mechanism, Conceptnet Numberbatch, LSTM, NLP Techniques, Product Review, Seq2Seq Model, Text Summarization

---

1. INTRODUCTION

THERE has been a continuous increase in the number of internet users every year. With increase in Internet users, comes a great deal of information that gets stored online every second. This amounts to storing of huge amount of data every second. It may contain useful and unnecessary data as well. It requires a large number of data centers to store this enormous amount of data. In addition to this, sometimes even the useful data becomes difficult to understanding due to the noise in them. So, there is a need for summarizing this data without losing the original meaning of the data and at the same time reducing the size of the data. Thus the process of Text Summarization comes into picture with its benefits spread over different fields such as Machine Learning, Natural Language Processing, Artificial Learning, Semantics etc., Few years earlier, when Internet has not reached the common man, online shopping was considered to be the worst ways of shopping. People never used to order online as it lacked the touch and feel scenario that we have when we go for shopping.

Gradually with the increase in the use and widespread availability of Internet, online shopping has gone to a different level we never expected.

With the increasing use of smart phones, increased number of online shopping sites, the pretty easy user interface, the quality of the items purchased, gradually online shopping has taken a head-start and reached an unexpected level today. Online Shopping has become a common thing these days as wide variety of products are available at a single place. The ease of ordering a product and getting the product delivered directly to home at a convenient date and time has attracted many people. The people are getting addicted to it. Along with these, the discount offers being offered by the Online shopping sites is making the people stick to online ordering. The wide varieties of items is an added advantage. People get items at a cheaper rate than the available market price. Everyone refers to the product reviews before buying a product. Then they can come to a conclusion of which is the best product to buy among the different products available. Suppose a user need to buy a laptop. Then he must go through different kinds of laptops available at his budget. He should make a note of different reviews for each product. He should consider the positive reviews and negative reviews for a particular product. He should even consider the rating given to each product. He must understand what he read and then only he can choose the best among the available laptops. This is a tedious and time taking task. Some even find it difficult to choose and approach the local vendors and get trapped to buy the same product at a higher price than what the online shopping is offering. In addition to this, some users' reviews are so long that the user could get the actual meaning of it only after closely going through the review. Thus there is a need for minimizing the review to a shorter representative sentence which depicts the same meaning as the whole content. It will be better if he could get a selection as well along with the representative sentence. Thus the text summarization and classification comes into picture that could make the summary of a review and thereby classify the product to be good enough to buy or no need to buy.

1.1 Problem Definition

Online shopping has become a common thing now a days. Product reviews, rating of the product, popularity of the product and quality of the product decides what product to buy from the whole lot. A person mainly relies on product reviews and rating of the product for buying a particular product from

---

- *Ravali Boorugu is currently pursuing masters degree program in software engineering in GRIET,Hyderabad,India , Ph-09704954527. E-mail: ravaliburugu@gmail.com*
- *Dr. Gajula Ramesh is currently working as Associate Professor in CSE Department, GRIET,Hyderabad,India PH-9440862112. E-mail: ramesh680@gmail.com*
- *Dr. Karanam Madhavi is currently working as Professor in CSE Department, GRIET,Hyderabad,India PH-9849364116. E-mail: bmadhaviranjan@yahoo.com*

the different types of the product varying only in terms of price and quality having approximately the same features. But for a particular product, it is difficult to go through all the reviews of the product. At the same time, he need to go through reviews and rating of all the products of same type and then need to come to a conclusion. This often becomes a problem when the reviews are in large number.

2 RELATED WORK

There are many prominent works in Text Summarization from the past few years. Earlier works dealt mainly with Single Document Text Summarization. Now that the technology has increased as well as computing power has increased which paved the path for faster, more effective and more accurate way of processing documents when compared with the earlier methods. Niladri Chatterjee, Amol Mittal and Shubham Goyal in [2] proposed an extractive based Text Summarization technique that makes use of Genetic Algorithms. In this paper, they represented the single document as a Directed Acyclic Graph. Weight is given to the each edge of the DAG based on a schema explained in the paper. They use an Objective function to express the standard of the summary in the terms such as ease of readability (readability factor), how closely sentences are related (cohesion factor) and topic relation factor. The Genetic Algorithm is intended to maximize the Objective function by selecting the prominent sentences from the whole text. Initially the Cohesion Factor i.e., how closely are the sentences related to each other is calculated. Then, the sentences that are similar to the input query should be given highest preference called as Topic Relation Factor is calculated. After calculating the aforementioned factors, we can determine the Objective Function (fitness function) of the summary. Then we use Genetic Algorithm to maximize the Objective function. Amol Tandel, Brijesh Modi, Priyasha Gupta, Shreya Wagle and Sujata Khedkar in [5] proposed a multi document summarization technique that will allow the customer to condense relevant data from multiple documents given as a single input. This method could save ample amount of time along with increased efficiency. They have inspired from the then existing approaches like Cluster based, Topic based and Lexical Chain based. LexRank prevents the score maximization of Sentences that are not relevant to the main theme of the document. Lower scores are given to the sentences that contain noisy data considering the fact that there will be no similitude with the cluster. In the initial phase they will extract the summary of each single document. Then generated metadata from those documents. This metadata is used to construct a graph that shows how the sentences are relevant to each other by considering each document as a node and the appropriate weights are given based on the similitude of the metadata generated earlier. Shivangi Modi & Rachana Oza in [10] discusses in detail about 3 single document techniques and 2 multi document techniques. Aditya Jain, Divij Bhatia, Manish K Thakur in [6] proposed a model which used Word Vector Embedding for Extractive Text Summarization. As per their paper, there are four prominent problems to deal with while extracting information. They are recognizing the most salient sentences from the document, removing the unnecessary information that is not relevant to the theme of the document, minimize the details and putting together the initially extracted information that is relevant into a condensed and organized report. To overcome the aforementioned challenges, they proposed a Word Vector Embedding approach to extract the prominent, then they used a Neural Network for Extractive Summarization by using Supervised Learning method. They tested on DUC2002 dataset and found that the results were more accurate when compared with the earlier summarizing methods. The results were satisfactory but can be improved if we increase the size of the dataset and theme diversity of the dataset and then implementing more effectual approaches like Sequence to Sequence Recurrent Neural Network for summarizing. Nithin Raphal, Hemanta Duwarah and Philemon Daniel in [11] provided a review on the prominent research performed on the abstractive text summarization. As there are two methods of summarization: Extractive and Abstractive methods. The Extractive method as said by Aditya Jain et el [6] will select the most important sentences from the document and make the summary out of it by maintaining the coherence between the sentences and sticking to the theme of the document. The Abstractive method on the other hand creates a summary by creating the phrases or sentences that may or may not be present in the document but could bring the complete meaning of the document. This is way more difficult than extractive technique used earlier. It is very much similar to what a human could generate after going through a document. Word embedding method and one hot vector methods failed to detect the similarly occurred word. This problem was resolved in Mikolov et al [3] [4] model in which they used continuous skip gram model which takes input word and can project the probable contextual words whereas on the contrary the continuous bag of words model is exactly the converse of the CSG method. They proposed various methods by which extractive summarization can be done, the preprocessing steps that are to be done in the initial phases, discussed about the latest research in this arena, the various kinds of architectures, mechanisms involved, supervised and reinforced learning & the advantages and disadvantages of various architecture. Yang Wei and Yang Zhizhuo in [7] "Query based Summarization using topic background knowledge" (2017). Basically a query oriented approach means to develop the summary based on the query given as an input. As the most of the queries doesn't hold the semantic details or information, the query based model is not effective. So Yang et al. proposed a model that will use the search engines to develop a background knowledge of the main theme of the document. Later they used the Page rank algorithm which contains the document information and cross document information. They applied this algorithm on the document to construct the summary of the document. They used the China search engine Baidu for the building of theme background knowledge. In the future works that may be extended to Google, Yahoo etc. and the results can be compared with the earlier results. In this way we can build a more accurate summary as there is a good knowledge of the background theme of the document. According to Shi Ziyan in [1], Summarization could not bring accurate results when the words has a lot of meanings. So there is need for the particular domain knowledge of the main theme of the document as well. This brings the domain-specific text summarization into limelight. But the problem arises when the referring is done inaccurately. Therefore this paper proposes a co-reference resolution algorithm to sort out this problem and bring accurate results. On the similar lines Paul Gigioli, Nikhita Sagar, Anand Rao, Joseph Voyles [9] "Domain-Aware Abstractive Text Summarization for Medical Documents" (2018) extended the domain specific summarization by adding deep reinforced abstractive summarization method which is capable of going through the biomedical abstracts and summarizing them into a single line summary. Priya Pawar, Siddhesha Tandel, Shweta Bore & Nikita Patil in [8] discussed about the importance of summarizing and classifying product reviews. They used hybrid classifiers such as SVM and Naïve Bayes. They also concluded that with the increase in the classifiers count the accuracy can also be increased.

2.1 Online Product Review Summarization

There were many other previous works on review summarization like Dr. B. Jayanti et al. have worked on a novel approach to generate summaries of text document. They used DTPE (Decision-Tree-Pattern-Extraction) algorithm for summarizing the text automatically. But the disadvantage here is that both the comments whether it may be positive or negative are being entered into the same range and it took much time for summarizing. So, later [8] "Online Product Review Summarization" by P Pawar, S Tandel, S Bore and Nikita Patel proposed this model to summarize product reviews. In this proposed system they used the hybrid-classifiers such as SVM and Naïve Bayes in sync with fuzzy logic.

2.2 Summarization of Customer Reviews for a Product on a website using Natural Language Processing

This is a very similar work to what we are proposing. It stands on the same lines of standard as that of our proposed word. Here they developed an android application that takes URL (link) of the product as an input and gives the output in terms of rating for different features. For example, consider a phone that is being kept for sale on Amazon shopping website. If we give the URL of the Amazon page displaying the phone for sale and then clicking on the get summary button initially validates the given URL and then brings the review data to the local storage. Then this data is being tokenized and then Part of Speech tagging is done with the help of the Word-Corpus for processing. Then Sentiment analysis training is done on the opinion and feature word list formed after the tagging is done. Here the reviews are categorized into negative, positive and neutral reviews. Then they are using the classifier such as Naïve Bayes classifier to further improve the analysis in which the training is given on a particular dataset consisting on huge number of reviews. This will improve the accuracy of the summary generated.

2.3 Summarizing Customer review based on product feature and opinion

They used K-NN which is a supervised ML algorithm for the classification of the reviews. Then for combining appropriate words and for identifying the features of the product, they used the syntactic rules. They used SentiWordNet for identifying the opinion sentences using the polarity score of the opinion words. Then they generated summary of the reviews based on the features available for a particular product. Summarization of Online product review can be achieved with higher accuracy by using Seq2Seq model's advanced version i.e., LSTM along with the attention mechanism for increasing accuracy. This model could bring more accurate and very close summary of the product review submitted by the customer for a particular product.

2.4 Proposed Model

The previous works had many drawbacks and therefore could not produce accurate results as expected. They also some had limitations due to unavailability of the data sets, classifier limitation, time constraint etc. Summarization of Online product reviews can be achieved with higher accuracy by using Seq2Seq model along with the attention mechanism for increasing accuracy. We use LSTM and GRU cells rather than the basic RNN cell for making accurate prediction of the summary. We are not using the previous word embeddings like Word2Vec and GloVe. We use the latest word embedding model Conceptnet Numberbatch which is very much similar to GloVe but comparatively is a better than that. We use the same word embedding even for the classification purpose. During classification we use 1D convolutional layer followed by max pooling layer, LSTM layer and then at the end by a fully connected layer. With this kind of approach there is higher chance of predicting the accurate classification whether the product is either good enough to buy or not good enough to buy. This model could bring more accurate and very close summary of the product review submitted by the customer for a particular product.

3 IMPLEMENTATION

3.1 System Architecture

This section describes about the workflow of this paper. The input text will go through different modules before getting the output. The initial text will be a dataset which needs to go through the text preprocessing step initially to make if free from noise and unwanted data. This makes sure that the data is clean and ready for the next step. The preprocessing itself has various steps like 1) Noise Removal, 2) Tokenization, and 3) Normalization. Noise Removal is often the foremost step in the preprocessing of the text which consists of removal of the file headers, the markup data like XML, HTML etc. and even extraction of the important information from the data formats such as JSON, CSV of XLS files. The second step in preprocessing is Tokenization which is breaking down of larger sentences into smaller phrases and then into words which eases the processing of text. Here we use NLTK for dealing with tokenization. We use word.tokenize() function to achieve this tokenization process. Consider the below example of how tokenization is done on a sample text.

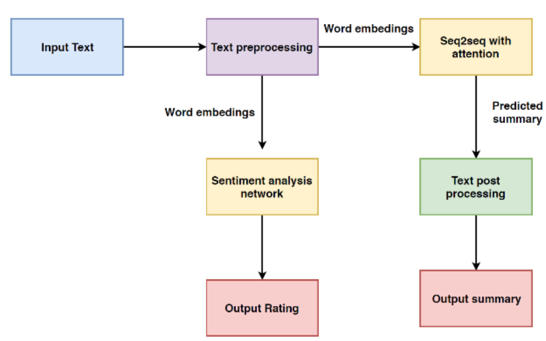

***Fig. 1**. System Architecture showing the workflow*

Eg: "Wow! This is an awesome view", Ravali exclaimed after watching the beach in Trivandrum.

The text after tokenize function is applied on it appears similar like:
[‘“’, ‘Wow’, ‘!’, ‘This’, ‘is’, ‘an’, ‘awesome’, ‘view’, ‘”’, ‘,’, ‘Ravali’, ‘exclaimed’, ‘after’, ‘watching’, ‘the’, ‘beach’, ‘in’, ‘Trivandrum’, ‘.’]
The next step in this preprocessing is Normalization. Normalization refers to all the steps in making the text into a homogenous level. It includes converting the whole text into upper case or lower case, removal of punctuation marks, conversion of the integers to their word equivalents etc. It brings the text into a homogenous level thus making the processing of words easy. After the preprocessing the tokenized words are being fed into a word embedding module which then recognizes the word meaning and the attributes that each word can carry. After applying the word embedding to the tokenized words, the output is then forwarded to the seq2seq module in which the actual summarization takes place. The seq2seq module is being explained in the further pages. The output of this module is the predicted summary which is the summarized form of the input text. It depicts the exact meaning as of the whole text but in a condensed version. The summarized text which is the output of the seq2seq module is then made to go through the post processing step before displaying the final summary. This step ensures that there are no errors in the summarized text and it carries the exact meaning as that of the whole text given as input. This is the final step in the summarization process. The output from the preprocessing module is to be given to the sentiment analysis network. This module helps in analyzing the review by considering the appropriate attributes for each word using the word embeddings. The attributes which help in analyzing the rating are to be taken and the other attributes are to be left. Then using the sentimental analysis network module, the rating is valued and this helps in choosing whether to buy a particular product or not. The output of this module is the final rating of each product by analyzing the whole summary of the product and rating of the product.

3.2 Summarization Architecture:

The deeper we go into the workflow of the summarization, the better we can understand how the summarization is actually done internally. As discussed earlier the output text after the preprocessing is done is taken and then it is fed to the word embedding module which then recognizes the word meaning and the attributes that each word can carry.

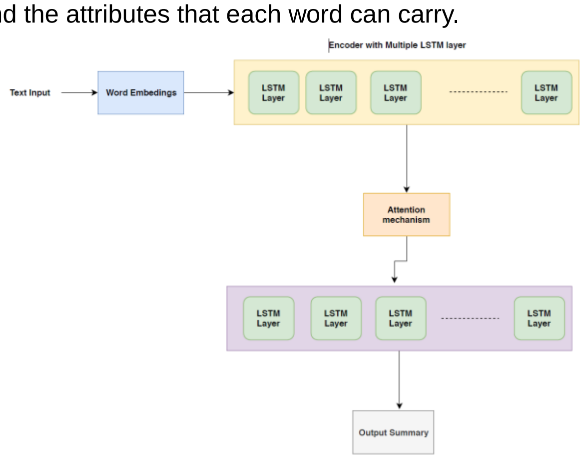

Fig. 2. *Summarization Architecture*

The text then is subjected to go through the Encoder with Multiple LSTM layers. Each LSTM layer contains the functions like sigmoid and tanh activation functions which could maintain the values of the words in range of {0 to 1} and {–1 to 1} respectively. We have gates like forget gate, reset gate, output gate and even the state of the cell. These gates manage which information should be passed to the next LSTM and which info to be forgotten. Apart from this the gates will also remember the relationship between words as we are passing the previous time step output (i.e. previous word) to next LSTM cell and so on. In this way the relationship will be extracted. The encoder will produce a fixed length representation of the text given as input. The words are then given to the Attention Mechanism module. It is very difficult to summarize a large text into some representative words which should depict the same meaning as that of the whole text. This could result in the loss of some important information which in turn could bring false results. This happens when the particular words have different meaning in the local and global context. So there is a need for special attention on those words for that particular context, this is what attention mechanism does. It is just a simple vector which takes the input from the encoder and generates a probability distribution for the input words and thereby giving the decoder global information pertaining to the specific words. This could result in selecting the representative words with more accuracy. In this context the attention mechanism gives special value to those words which could define the major meaning of the summary. Suppose there is a summary written on Chicken Burger. The words like delicious, spicy, good, tasty should be given more importance than the words like the, is, it. This is taken care by attention mechanism layer. The output from the attention mechanism layer is forwarded to the decoder module which again consists of many LSTM layers whose job now is to predict the words which are important. This decoder also contains the LSTM layers which here goes through the representation, an embedding of the word that is last produced and then uses these as the inputs for generating the words in the summary of the text. The LSTM cells have the same functionality as that of in the encoder but the as input changes in the decoder module the output will be a summary of the text without losing the meaning of the text.

3.3 Classification Architecture

The classification is often a tedious task. The output from the preprocessing module is given as input for the summarization module and as well as classification module. The input here is fed into the word embedding module in which the words are given a large number of attributes and depending on the context of the word being used, the exact meaning of the word is being taken. The word embedding we used is Conceptnet Numberbatch which is very much similar to GloVe but performs better than that and offers more attributes to the words than what the GloVe and Word2Vec offers. This Word embedding makes the classifier the words accurately and with more perfection. The output from the Word Embedding module is passed on to the next module Convolution 1D layer. The Convolution 1D later and Max pooling layer were used in Computer vision initially.

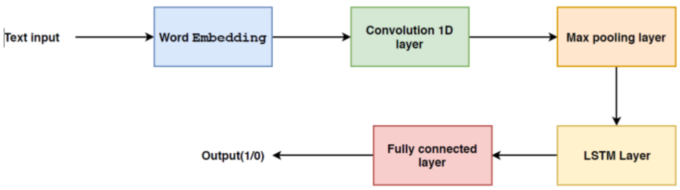

***Fig.3**. Classification Architecture*

But when they applied the same principles for the text classification the results were really promising. Therefore the use of CNN 1D and Max pooling layer is carried on to the text classification as well. The output from the Max pooling layer is given on to the LSTM layer and later to the fully connected layer which in turn produces the output as a positive or a negative review for a particular given product. The classification usually depends upon the review of the product and the rating of the product. The results depend on the data that we are providing as input for training the model. The larger the dataset, the higher will be accuracy The dataset taken into consideration is Amazon Food Reviews dataset. It consists of around 560000 reviews along with the summary for those reviews. It also has the columns such as ProductId, UserId, ProfileName, Score, Time, Text and Summary. The dataset contains reviews about various food products.

4 EXPERIMENTS AND RESULTS

The hardware configuration used for this method is a machine with 8GB RAM, i7 8th Gen processor with Nvidia GeForce GTX 1050 4GB graphics card. We used Python and Tensorflow for developing our models in the summarization and classification architecture. We used various libraries such as numpy, pandas, NLTK, Keras, RE, matplotlib, plotly etc. to code the models. Before experimenting of the dataset i.e. using the dataset. It is cleaned of all the unwanted characters from the text. The training of the model went on for nearly 15-18 hours per each training. The loss function got decreased with each iteration. You can see that in the figures. Each line in the figure depicts the decay of loss function for each iteration during the whole training time. We get the values such loss, epoch, seconds and batch for a set of particular batch undergone training during the span of 'T' seconds (T varies with each GPU and with each batch). We calculate the aver age loss for each iteration.

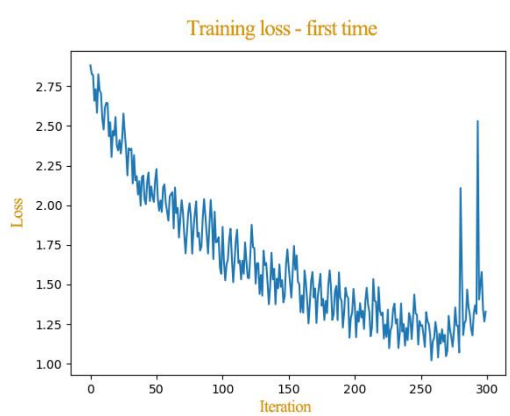

***Fig. 4**. Training the model for the first time*

The training of the model is designed such that the training get stopped whenever the average loss for three consecutive iterations are in increasing order. You can see that in the figure where the training got stopped at last. Each figure contains a line which indicates the loss decay (cross entropy loss function) when plotted with the number of iterations on the X-axis. The average loss got decreased from 5.268 to 1.181 during training the model for the first time.Then later during the second time it got decreased from 5.092 to 0.339 and during the third time it got further decreased from 5.147 to 0.263. We used the model after each training to summarize the reviews and the results were satisfactory. We observed that the summary the model produced got better with each training. You can observe the same trend in the Table shown above.

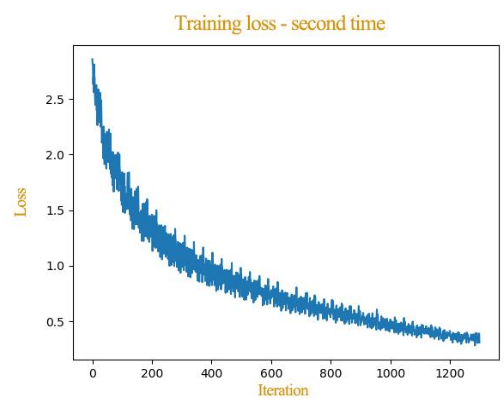

***Fig. 5.** Training the model for the second time*

The summaries were apt in each case but got extremely better with each training phase. This shows that further training will certainly make the model predict better summaries.

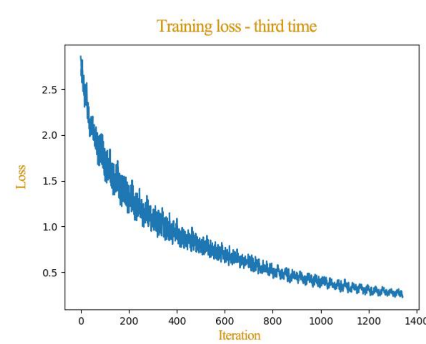

***Fig. 6.** Training the model for the third time*

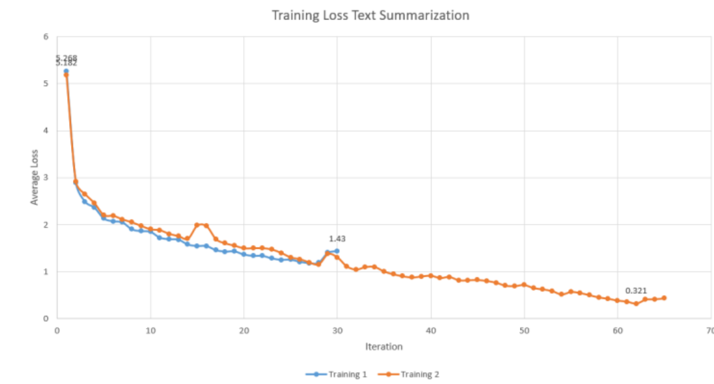

Fig. 7. *Average Training Loss*

When we take the average loss of the losses produced during each training model. The algorithm is designed in such a way that the training of the model will stop as soon the average loss occurs continuously in increasing order for 3 times. This is shown in the figure above.

TABLE 1
REVIEWS AND SUMMARIES GIVEN BY DIFFERENT MODELS

S.No	Review	Summary after Training Model 1st time (Avg Loss = 1.181)	Summary after Training Model 2nd time (Avg Loss = 0.339)	Summary after Training Model 3rd time (Avg Loss = 0.263)
1	I'm the only person in our house who enjoys a tuna SANDWICH so opening a can of tuna means some of it usually goes to waste. The small can of tuna salad (already prepared) makes one perfect sandwich, so unexpectedly, this snack ends up making a full meal for me. I use the can of tuna salad on two slices of bread. There is ample tuna for one sandwich. You can garnish it however you like. The tuna itself is already prepared, and is actually quite tasty. While I had originally purchased it as a snack, it's more beneficial to me as a sandwich.	great for salads hummus	tasty salads for sandwich	perfect for lunch
2	First ran across these lovely cookies from a friend living in Holland who brought a bag to the US as a gift. We loved them! So I wanted to see if it was possible to find them on my own here in the US. These tasted like the real deal. Can't wait to buy more!	these are great	nice biscuits	great cookies
3	This was a special Christmas gift for my BFF. The outer package was well wrapped and materials were perfect to keep the inner box in perfect condition. The macarons came boxed with a heavy weight lavender colored gift box (very elegant) with instructions for storage. Once they were unwrapped, the macrons were just beautiful (such lovely colors) and they tasted as exquisitely as they looked (wonderful flavors). You will certainly be thanks for such a lovely gift!!	great gift	colorful cookies	beautiful and delicious

TABLE 2
CLASSIFICATION OF PRODUCT REVIEWS

Sl No	Review	Rating	Output
1	McCann's Oatmeal is a good quality choice. Our favorite is the Apples and Cinnamon, but we find that none of these are overly sugary. For a good hot breakfast in 2 minutes, this is excellent.	5	1
2	No tea flavor at all. Just whole brunch of artifial flavors. It is not returnable. I wasted 20+ bucks.	1	0
3	I fed this to my Golden Retriever and he hated it. He wouldn't eat it, and when he did, it gave him terrible diarrhea. We will not be buying this again. It's also super expensive.	2	0
4	Always being a fan of ramen as a quick and easy meal, finding it on amazon for a decent price and having it delivered to your door by the case is an amazing situation for anyone to find themselves in.	4	1

After the model is trained, it is tested by giving a new review. The model could predict the exact output i.e. it is able to classify whether to buy a product or not in terms of 1 or 0 respectively. The above figure shows the sample review and its output as produced by the classifier

5 CONCLUSIONS

This paper explained in detail some of the remarkable works in the arena of text summarization. Summarization has always been a necessity since many years as there is a huge amount of information being released into the internet every day. This paper described all the major summarizations techniques and the prominent works that are being done on each technique. In this paper we used text summarization for summarizing the product reviews and even classifying the reviews so that we can come to a conclusion whether to buy a product or not. We used Seq2Seq model for doing the summarization. We have many optimization techniques available in this model like Attention mechanism, Beam search and Bucketing. We used attention mechanism optimization for getting higher accuracy. In the seq2seq model, instead of using the basic RNN cell, we replace that with the LSTM cell. We used LSTM layers during encoding and decoding modules during the summarization. For the classification of the product reviews, we used the 1D convolutional layer followed by the max pooling layer, LSTM layers and finally using a fully connected layer. This resulted in getting a classification with higher accuracy than existing approaches. As we used attention mechanism between the encoder and decoder, it resulted in predicting the accurate summary of the product.

6 FUTURE SCOPE

Our future work will focus on how to improve the accuracy by using the latest models in the field of Text Summarization. Google introduced a paper which revolutionized the entire machine learning world. Instead of using the Recurrent Networks such as LSTM, GRU etc., it uses Attention mechanism alone to do the summarization of text. This resulted in getting higher accuracy when compared with the Seq2Seq model. As a future work we can use this kind of attention mechanism alone in the Encoder and Decoder modules to get a higher accuracy. We can also use the state of the art Google's BERT- which uses Transformers in Encoding module in Bidirectional. This could bring in more accurate summary that what we have achieved using the Seq2Seq model.

7 REFERENCES

[1] Shi Ziyan "The Design and Implementation of Domain-specific Text Summarization System based on Co-reference Resolution Algorithm" 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery.

[2] Niladri Chatterjee, Amol Mittal and Shubham Goyal's "Single Document Extractive Text Summarization Using Genetic Algorithms" (2012)

[3] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean," Distributed Representations of Words and Phrases and their Compositionality," arXiv:1310.4546v1 [cs.CL], 2013

[4] Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean, "Efficient Estimation of Word Representations in Vector Space," arXiv:1301.3781v3 [cs.CL], 2013.

[5] Amol Tandel, Brijesh Modi, Priyasha Gupta, Shreya Wagle and Sujata Khedkar's "Multi-document text summarization - A survey" 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)

[6] Aditya Jain, Divij Bhatia, Manish K Thakur's "Extractive Text Summarization using Word Vector Embedding"(2017).

[7] Yang Wei and Yang Zhizhuo "Query based Summarization using topic background knowledge" 2017 13th International Conference on Natural Computation, Fuzzy

Systems and Knowledge Discovery (ICNC-FSKD)
[8] Priya Pawar, Siddhesha Tandel, Shweta Bore & Nikita Patil "Online Product Review Summarization" 2017 International Conference on Innovations in Information, Embedded andCommunication Systems (ICIIEC
[9] Paul Gigioli, Nikhita Sagar, Anand Rao, Joseph Voyles "Domain-Aware Abstractive Text Summarization for Medical Documents" published during 2018 IEEE BIBM.
[10] Shivangi Modi & Rachana Oza "Review on Abstractive Text Summarization Techniques (ATST) for single and multi documents" 2018 International Conference on Computing, Power and Communication Technologies.
[11] Nithin Raphal, Hemanta Duwarah and Philemon Daniel "Survey on Abstractive Text Summarization" 2018 International Conference on Communication and Signal Processing (ICCSP).