Sentiment Analysis using Attention Mechanism, Convolutional Neural Networks and Long Short-Term Memory

Sarang Ajay Patel[1], Sanket More[2], Dr. Pathan Mohd. Shafi[3]
[1,2]*Student- Department of Computer Science, MIT Art, Design and Technology University*
[3] *Professor- Department of Computer Science, MIT Art, Design and Technology University*

***Abstract*—Human sentiment analysis is a domain which has grown exponentially in recent years. There is a lot of work already published but still has scope of improvement and advancements. The aim of the project is to develop a AI/ML model which has basic and advanced capabilities of classifying the textual data, image files, etc. into accurate and automatic sentiment classification. The project mainly focuses on text and image sentiment analysis using a combination of two methods namely, LSTM and Attention-based mechanism. This research report is a deep analysis of working of this method and improving the accuracy of sentiment classification.**

The successful implementation of this project can introduce a wide variety of use-cases. It can have many applications ranging from customer feedback analysis, fake news identification, creditworthiness assessment. These applications may act as crucial factors, contributing towards growth of businesses and organizations.

***Index Terms*— Sentiment Analysis, LSTM, CNN, Attention Mechanism, VGG16, RNN**

I. INTRODUCTION

The project deals with classifying the textual data, image files, etc into a number of human sentiments. Accepting the data in textual format, image format, it analyzes using the AI/ML methodology mentioned and gives the results by classifying into different sentiments such as positive, negative or neutral.

The successful implementation of this project can benefit different organizations and businesses in improving their customer service, brand awareness, etc by analyzing the customer behavior or sentiment toward their products or services.

There is an increased need for tools like sentiment analysis in businesses, political campaigns, social media, etc. Sentiment analysis can help brands understand how their customers feel about them. They can keep an eye on their brand reputation by analyzing communities, forums and social media platforms. They can conduct surveys to understand what their customers feel.

The need for free and accurate sentiment analysis software with advanced features is high, as this would fill the gap of less accessibility of such software. Currently, the majority of the software offer free trial for a limited time period and then charge a hefty fee. Our product will be a differentiating platform creating healthy competition.

The objective of the project is to research on existing and advanced methodologies of sentiment analysis, identify the existing research gaps and research on ways to tackle them. This report describes all the aspects of the project. It covers the literature survey, software requirements specifications, followed by research on various gaps and methodologies.

II. LITERATURE SURVEY

Woldemariam [1] has tried to apply sentiment analysis to forum discussion posts. Using two approaches namely, lexicon based and RNTN, they have tried to check which is better at determining the sentiments. The lexicon method uses sentiment dictionaries with labeled sentences to train and the RNTN uses sentiment treebank. The results have shown that RNTN is doing 9.88% better at variable length texts.

Jie Li; Lirong Qiu [2] have tried to make use of emotional words and modifiers to get more accurate results for getting the sentiment of a sentence. This is achieved by dependency parsing and modified distance.

[3] The authors have tried to analyze the sentiment analysis of a popular Chinese application weibo. This research paper has tried to see the different approaches

that have been applied for this, check for the problems that exist and explore the future aspects of it.

Different approaches have been used for the image sentiment analysis but all of them took the whole image into consideration. To increase the accuracy, authors [4] have tried to implement the attention mechanism to highlight only the important aspects of the image. They have further used SVM in place of the softmax classifier on CNN.

III. RESEARCH GAPS

Through the study done through related work and comparison of existing work, gaps were identified in the existing work. One of the first things that were identified was that there are very limited sentiments namely positive, negative and neutral that have been worked on. Additional sentiments can be added and their analysis would give us a more clear picture.

Secondly, Most of the work was done on text based analysis. When in reality, people tend to use audios and images as well for reviewing and sharing their emotions. Therefore, work on these forms of input also would broaden our field.

Next, usually the sentiments are given for the inputs in a general way, but it doesn't tell us who the sentient is targeting. As there can be multiple subjects in the input but the sentiment might be only regarding one of them. This gives a wrong estimation.

Statements which portray neutral sentiments can have positive or negative impact on our subject. How these neutral statements affect our subject can be a crucial factor to look into.

Thus, a wide range of work has been done in the field using just the English language. But reviews, concerns are usually given in various languages. Limiting ourselves to just one language is leaving us far from the wide area of analysis. Therefore, including multiple languages can be a big plus.

IV. PROPOSED METHOD

a. Image Sentiment Analysis:

1. Image caption generation:

Image based sentiment analysis is an important field to look into as in today's world many people portray their emotions through images online.

For this task, we will be using an attention-based mechanism. Humans have a cognitive ability called attention. Attention helps us to ignore irrelevant information and only focus on the main information that is useful. This helps in saving time and resources. This same task of consciously ignoring irrelevant data can be done in neural networks through attention mechanism. With this mechanism we will divide our image into n parts. Then when the RNN tries to generate a word for our caption, the attention mechanism would focus only on certain parts of the image so that decoder only uses them to create the caption. We will be using local attention as it helps lower the computational cost. It does this by calculating the weight in neighboring windows from its position and weights it with the context vector.

For our convolutional neural network, we will be using VGG16. As we won't be classifying the images therefore, we don't need the softmax classifier. Thus, the VGG16 will be implemented without it.

i. Dataset will be used from kaggle. It will be flickr8k dataset.

ii. The images will all be converted into 224x224 size.

iii. The dataset will be divided into 70:30 for training and testing respectively.

iv. Using the captions in the training set we will tokenize and build our vocabulary.

v. VGG16 encoder, RNN and RNN decoder with attention mechanism are defined.

vi. Teacher forcing methodology will be implemented as it helps to learn the sequence quickly.

vii. Once the model is trained, captions will be generated.

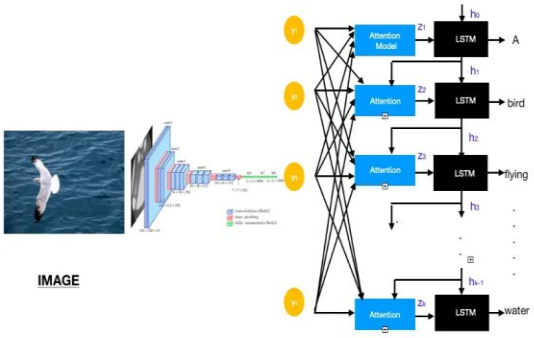

Fig. 1 Caption Generation using attention mechanism

2. Caption to sentiment:

LSTM stands for long short-term memory and it improves the basic RNN which suffers from short term memory.

RNN's work sequentially and keep feeding previous data along with new input. The problem is as we keep

moving forward the gradients start diminishing and therefore information is lost. LSTM was made to explicitly counter this problem.

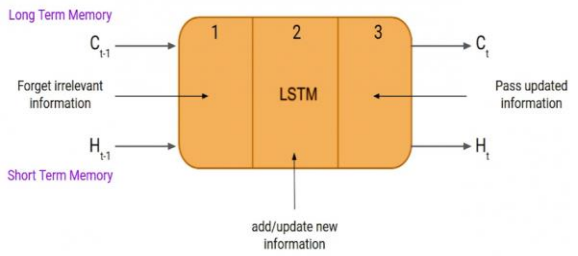

Fig. 2 LSTM Architecture

i. Load the captions that were generated in the previous step.
ii. Clean the data.
iii. convert the words into tokens.
iv. Next convert the tokens into vectors.
v. LSTM will be trained with the dataset.
vi. The data will be sent to the softmax function to get values between 0 & 1.
vii. Values closer to 0 indicate negative sentiment and those close to 1 show positive.

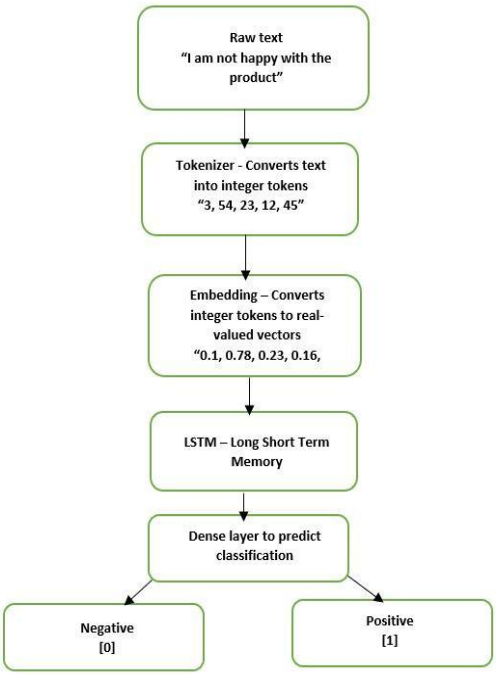

Fig. 3 Text To sentiment Analysis

b. Audio Sentiment analysis:
The speech to sentiment recognition is done using python and uses the "Speech Recognition" API and "PyAudio" library. This process is carried down in two sub-steps. The first step involves analyzing and converting the audio file into a textual format. This conversion takes place using the speech recognition library. The program takes the input as an audio file, from the user. This audio file is passed in the library and the text conversion is saved into a variable. The speech_recognition library supports wav, AIFF, AIFF-C, FLAC audio files.
The next step is classification of this textual data into various sentiments, say, positive, negative or neutral. For this purpose the LSTM model described in the previous section will be used.

VII. CONCLUSION

We have seen that Sentiment Analysis is used to analyze different types of inputs into various human sentiments like positive, negative, and neutral. The project discussed the existing research gaps, methodologies, use cases, etc. and presented solutions to tackle some of the research gaps. It mainly discussed the text-based, image-based and audio-based implementation of sentiment analysis.

VIII. FUTURE WORK

The above study contributes towards the study of Sentiment Analysis with its advanced functionalities. This algorithm could be further developed to add multiple use-cases, make it more efficient and faster. Further, there is scope of research and development of sentiment analysis in terms of adding multiple sentiments, identifying the subject which the sentiment is targeting, positive/negative impact of neutral statements, etc.

REFERENCE

[1] Y.Woldemariam, "Sentiment analysis in a cross-media analysis framework," 2016 IEEE International Conference on Big Data Analysis (ICBDA), 2016, pp. 1-5, doi: 10.1109/ICBDA.2016.7509790.
[2] J.Li and L. Qiu, "A Sentiment Analysis Method of Short Texts in Microblog," 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and

Ubiquitous Computing (EUC), 2017, pp. 776-779, doi: 10.1109/CSE-EUC.2017.153.

[3] Yuan, Zhang & Yuwei, Gu & Feng, Zhao & Reimei, Wang. (2019). A Literature Review of Sentiment Analysis on Chinese social media. Journal of Physics: Conference Series. 1314. 012140. 10.1088/1742-6596/1314/1/012140.

[4] P. Das, A. Ghosh and R. Majumdar, "Determining Attention Mechanism for Visual Sentiment Analysis of an Image using SVM Classifier in Deep learning-based Architecture," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 339-343, doi: 10.1109/ICRITO48877.2020.9197899.

[5] N. Mittal, D. Sharma and M. L. Joshi, "Image Sentiment Analysis Using Deep Learning," 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), 2018, pp. 684-687, doi: 10.1109/WI.2018.00-11.

[6] S. Jindal and S. Singh, "Image sentiment analysis using deep convolutional neural networks with domain specific fine tuning," 2015 International Conference on Information Processing (ICIP), 2015, pp. 447-451, doi: 10.1109/ INFOP. 2015.7489424.

[7] J. Yi, T. Nasukawa, R. Bunescu and W. Niblack, "Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques," Third IEEE International Conference on Data Mining, 2003, pp. 427-434, doi: 10.1109/ICDM.2003.1250949.

[8] K. Mouthami, K. N. Devi and V. M. Bhaskaran, "Sentiment analysis and classification based on textual reviews," 2013 International Conference on Information Communication and Embedded Systems (ICICES), 2013, pp. 271-276, doi: 10.1109/ICICES.2013.6508366.

[9] L. Kaushik, A. Sangwan and J. H. L. Hansen, "Automatic Sentiment Detection in Naturalistic Audio," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 8, pp. 1668-1679, Aug. 2017, doi: 10.1109/TASLP.2017.2678164.

[10] H. Yanagimoto and M. Shozu, "Multiple Perspective Caption Generation with Attention Mechanism," 2020 9th International Congress on Advanced Applied Informatics (IIAI-AAI), 2020, pp. 110-115, doi: 10.1109/IIAI-AAI50415.2020.00031.

[11] J. Sudhakar, V. V. Iyer and S. T. Sharmila, "Image Caption Generation using Deep Neural Networks," 2022 International Conference for Advancement in Technology (ICONAT), 2022, pp. 1-3, doi: 10.1109/ICONAT 53423.2022. 9726074.