

Enterprise Chat Platform using Machine Learning Techniques

Malvika R¹, Vikram K Kharvi², Akhil Bidhuri³, Bhaskar Kumar⁴, Dr Annapurna D⁵

Student, Department of Information Science and Engineering, PESIT Bangalore South Campus, Hosur Road,
Electronic City, Karnataka, India^{1, 2, 3, 4}

Professor, Department of Information Science and Engineering, PESIT Bangalore South Campus, Hosur Road,
Electronic City, Karnataka, India⁵

ABSTRACT: Enterprise chat platform is a web application designed on the principles of machine learning. It allows each on boarded enterprise to communicate with their employees by using group chats. The system can be accessed from either a desktop or a mobile device. The proposed chat platform will include two additional features apart from basic chat features like chatting, calling and file sharing namely sentiment analysis and text summarization. Text summarization, a NLP feature, creates precise and short summaries for a given source document. The Encoder-Decoder RNN architecture developed has proven to be effective when applied to a text summarization problem. Sentiment Analysis intends to gage a person's opinion on a project or help determine the outlook of a speaker regarding some contextual data. The CNN technique is an applicable algorithm to classify and interpret sentiment of data.

KEYWORDS: Enterprise, Sentiment Analysis, Text Summarization.

I. INTRODUCTION

Current business environments have necessitated the need for utilising online chat features like chatbots or live chat to engage in positive interactions among companies and clients.

Multinational corporations and banks utilise chat bots to fulfil tasks that do not require a specific skill base apart from promoting healthy interactions. In saying that, the development of chat platforms in the blooming enterprise industry has been very minute. In the present day, there exists no platform that can help with daily corporate tasks like scheduling meetings, follow up on team activities, fill out a timesheet, conducting training for employees or help with team recruitment. Utilizing a chat platform to accomplish these tasks can lead to significant enhancements in employee collaboration, while providing huge savings for the organization [1].

Sentiment Analysis of text messages is gaining importance as businesses are becoming increasingly dependent on data, and much of this data is coming from platforms like emails, social media and online surveys. Extracting necessary information from this large content of data is difficult and time consuming. Sentiment analysis in an enterprise chat platform allows to identify the attitude and emotion of an employee or client based on the subject matter in a conversation. Classification and analysis of the sentiment will be performed through the proposed Convolution Neural Network (CNN) algorithm.

The introduction of text summarization in an enterprise chat platform makes the platform user friendly. Summarization of text is accomplished by shrinking the original set of text messages so as to give a summary of the conversation but the significance of the messages is not changed. In a situation where in a new employee has joined a team group chat, he will just receive a summarized overview of the team strategy over chat instead of scrolling through the entire chat conversation. Extractive text summarization and abstractive text summarization are the two methodologies of the Text summarization feature. The extractive summarization method chooses parts of phrases, words and sentences of the specific text documents to generate the text summary. In contrast, understanding the central idea of the text document followed by representing the same in a natural language utilizing a natural language generation method to create

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

summaries [2] is the main idea behind the abstractive approach.

Additionally, apart from the two main features explained, this chat platform will have features for team and company managers to view a dashboard that will display statistics regarding the number of messages sent and received, number of team members and the number of active users. Additionally, a sentiment data chart is updated simultaneously according to the percentage of positive, negative or neutral sentiments present in the message.

This paper concentrates on a machine-learning technique to create a suitable web application for the enterprise chat communication using a proposed CNN algorithm with a focus on sentimental analysis and RNN algorithm for abstractive text summarization.

Paper is organized as follows. Section II deals with the existing work and platforms present currently in the market. The proposed model is introduced in Section III followed by Experimental Setup in Section IV. The results are revealed and analyzed in Section V. Finally Section VI presents conclusion.

II. LITERATURE SURVEY

In the paper titled 'Instant Messaging: No Longer Just Chat'[3] it was understood that IM has begun to outgrow its consumer-based origins and has established a presence in the corporate world. It has grown significantly since its launch changing the mindset of majority of its users who considered it just a casual chat service, which has been the technology's role since its inception.

There are several methods that can be utilized for text sentiment analysis. In the paper titled 'Twitter Sentiment Analysis' [4], the author has compared a number of deep learning and machine learning methods to implement an analysis of sentiment classification on a twitter dataset. He endeavored to classify the polarity of the various tweets based on its positive or negative nature. In some conditions where the tweet had both positive and negative elements; the more dominant sentiment was picked as the final label. The methods compared were Naïve Bayes, Maximum Entropy, Support Vector Machine (SVM), Convolution Neural Network (CNN), Recurrent Neural Network (RNN), Decision Tree and Random Forest. It was concluded that neural methods performed better than the classifiers.

Sentiment in a document can also be identified using RNN. According to paper titled 'Investigation of Recurrent Neural Networks in the field of Sentiment Analysis' [5], the author analyzed the performance of three RNNs namely vanilla RNNs, Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) on an Amazon health product dataset and concluded that GRU was the best option in terms of accuracy, but for future scope, CNN based hybrid models or RNNs based highway networks can be experimented.

Agarwal et al. in the paper "Sentiment Analysis of Twitter Data" [6], made use of the twitter to tackle the task of sentiment analysis. They classified sentiment into three classes namely, positive, negative and a neutral class. The unigram model, feature based model and a tree kernel based model were the three types of models chosen for experimentation. The tree kernel based model, necessitated the design of a new tree representation for tweets. Regarding the number of features, the unigram model uses over 10,000 features but the feature based model uses only about a 100 features. They concluded that the features that combine prior polarity of words with their parts-of-speech tags are essential for the task of classification. The tree kernel based model performed better than the other two models [7].

A major setback with sentiment analysis is regarding messages that are sarcastic in nature. Extracting the intended sentiment from such messages proves to be an arduous task. Aditya Joshi in his paper "Automatic Sarcasm Detection" [8], stated that there are three highlights in the research history of sarcasm detection that is required for sentiment analysis. These are semi-supervised pattern extraction, context beyond target text and hashtag based supervision. Semi supervised pattern was used to identify the implied sentiment. Large scale annotated datasets were created with the help of context beyond the target text and hashtag-based supervision.

Text summarization when implemented improves usability of platform. This particular feature explained in paper titled 'Abstractive text summarization using LSTM-CNN based deep learning'[9] explains how new sentences can be created by utilizing LSTM-CNN to study the more fine-grained fragments and condense them into a shorter representation while making sure to preserve information content and the overall meaning.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

A. Existing Model

There has been an emerging trend in the recent years where a vast number of chat applications which are present help people to connect with one other across different mediums, like Hike, WhatsApp, Telegram, etc. [10].

Table I, shows that most of the existing chat platforms have the ability to call and message other users but they lack analytical features like text summarization, sentiment analysis or data visualization built on it that helps improve business processes.

TABLE I
FEATURES OF EXISTING CHAT PLATFORMS

Enterprise chat platform features	Existing Chat Platforms				
	WhatsApp	Slack	Telegram	Facebook Chat	Skype
Messaging	Yes	Yes	Yes	Yes	Yes
Calling	Yes	No	No	Yes	Yes
Sentiment Analysis	No	No	No	No	No
Text Summarization	No	No	No	No	No
Data Visualization	No	No	No	No	No

III. PROPOSED MODEL

A. Sentiment Analysis

Sentiment Analysis is implemented using Convolution Neural Networks (CNN). Twitter dataset was utilized for training CNN, as the content is similar to an enterprise chat data. CNN does not accept characters; hence, the word vectors of each word were grouped in a tweet and summation or average of word vectors were taken to get one vector representation of each tweet in a numerical format. The unsupervised learning algorithm called GloVe was used to get vector format of each word.

For experimentation case, 25 dimensional vector representations were taken using GloVe and over 15 lakh labelled data from twitter dataset for training.

Fig.1 explains the different layers in CNN that are used to build the proposed model.

- 1) Convolution Layer: It has a set of filters that are used to perform linear multiplications by taking a part of input data at a time.
- 2) Pooling Layer: Average and Max pooling layer used to limit the quantity of parameters and dimensionality of each vector map.
- 3) ReLu Layer: The layer called Rectified Linear Unit is used to eliminate negative parameters from activation map.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

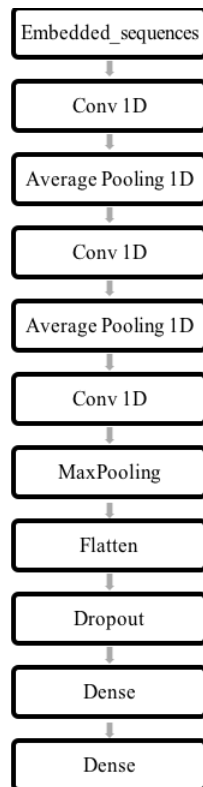


Fig 1. Layers of CNN

- 4) Dropout Layer: Dropout reduces overfitting of parameters. It makes sure that all individual nodes are removed from the network and only relevant nodes are present after each stage of training.
- 5) Flatten Layer: Flatten converts a 2D convolution layer to 1D layer.
- 6) Dense Layer: Dense or fully connected layer connects every node in one layer to every node in subsequent layers.

The more the CNN is trained for sentiment analysis on the enterprise platform; the conversations that employees engage in can be used as input to the training phase. This improves the accuracy with which sentiment of conversation is predicted.

B. Text Summarization

News Headline dataset was targeted as an applicable dataset to train and test text summarization. The Encoder-Decoder RNN is used to create short meaningful summaries of text document. It is a way of organizing sequence-to-sequence RNNs for prediction problems that have many inputs and outputs. The RNN module is used to parse through all the messages in the enterprise chat between the team members and generate a summary with a similar intent. The architecture has two components namely an Encoder and a Decoder.

The Encoder reads input sequence and encodes it to a context vector representation. The Decoder will read encoded input

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

and generate an output sequence as depicted in Fig.2.

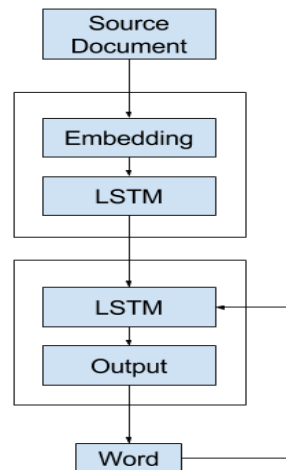


Fig. 2 RNN Architecture

To summarize the text of the source document, the encoder initially creates a context vector portrayal of the document. At each stage of output sequence created, the context vectors are sent to decoder that enables it to develop the equivalent interior state. This is utilized to construct the words in the output sequence so it is prepared to produce the following word in the arrangement. The procedure is iterated for each word in the output until the entire document is read.

IV. EXPERIMENTAL SETUP

The enterprise chat platform is a web application that helps companies engage with each other in the form of chat. The proposed chat platform architecture is shown in Fig.3.

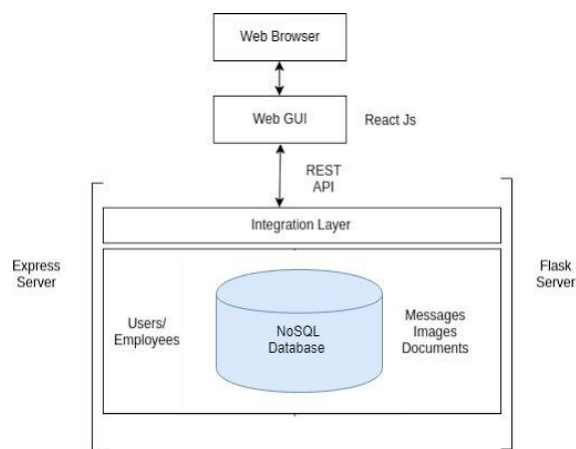


Fig. 3 System Architecture

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

In an enterprise scenario, Fig.4, the manager or HR of a company should be able to login to the chat platform over a network. He has the option to login to an existing company or onboard a new company. The chat platform has different views for the different hierarchy present in an enterprise.

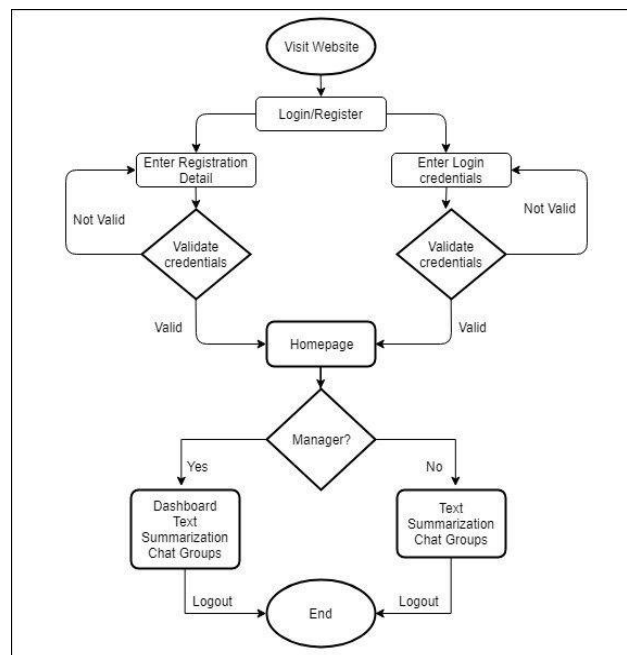


Fig. 4 Detailed flow of Chat Platform

The team members of a company on the other hand can only chat with their teammates and view summarized text.

V. RESULT AND ANALYSIS

A twitter dataset and enterprise data set generated was utilised to train and test the sentiment analysis. It was evaluated using classification accuracy.

In the proposed model, 80% of the data set for sentiment analysis was used for training and 20% for testing. It resulted in an accuracy of over 79%.

TABLE II
ACCURACY COMPARISON OF CNN WITH EXISTING MODELS

Algorithms	Accuracy
CNN	79%
RNN	38.3% [11]
Naïve Bayes	71.3% [12]
SVM	76.6% [12]

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

Table II compares the accuracy of sentiment model using CNN with the accuracy's of existing sentiment analysis models using other algorithms like RNN, Naïve Bayes and SVM.

Duyu Tang [11] made use of the YELP 2015 dataset by dividing it in the ratio of 80/10/10 for training, development and testing respectively and recorded an accuracy of 38.3% for RNN. The sentiment analysis for Naïve Bayes and SVM was observed by Balakrishnan Gokulakrishnan on a public stream of tweets from the Twitter microblogging website [12]. He along with his team concluded that SVM yielded a better accuracy when compared to Naïve Bayes. Overall, it is observed that the proposed CNN algorithm gave a much higher accuracy than RNN and Naïve Bayes whereas it had a slight improvement when compared with SVM.

Text Summarization module was evaluated using Loss Function instead of accuracy, the reason being that for a task like summarization it can generate multiple summaries which all have same intent.

TABLE III
CALCULATED VALUE OBTAINED BY RNN FOR SUMMARIZATION

Evaluation Metrics	Experimental Values obtained by RNN
Loss Function	0.5

Summarization module using proposed RNN algorithm was trained and tested on the news headline dataset and it resulted in a loss function of 0.5, (Table III). Loss function was used in the training period to find the best parameter values for the model. The lower the loss, the better the model.

VI. CONCLUSION

We have implemented an Enterprise Chat Platform using machine learning techniques. It includes text summarization and sentiment analysis of messages using RNN and CNN algorithms respectively. Our proposed algorithm for sentiment analysis successfully determined the sentiment of messages sent and received and simultaneously updated the dashboard for the manager to view. The algorithm for text summarization detected the messages received and created a short meaningful summarization for the users to view.

REFERENCES

- [1] K. Chidambaram, The Expanding Role of Chatbots in Enterprise Collaboration. Available: <https://www.cognizant.com/whitepapers/the-expanding-role-of-chatbots-in-enterprise-collaboration-codex2575.pdf>, 2019.
- [2] M. Kirmani, N. Manzoor Hakak, M. Mohd, and M. Mohd, "Hybrid Text Summarization: A Survey," in *Soft Computing: Theories and Applications (Advances in Intelligent Systems and Computing)*, pp. 63-73, 2019.
- [3] S. Ortiz, "Instant Messaging: No Longer Just Chat," *Computer, Article Summaries* vol. 34, pp. 12-13, 1994. IEEE Computer Society
- [4] A. F. Ansari, A. Seenivasan, A. Anandan, and R. Lakshmanan, "Twitter Sentiment Analysis," 2017.
- [5] K. Baktha and B. K. Tripathy, "Investigation of recurrent neural networks in the field of sentiment analysis," in *2017 International Conference on Communication and Signal Processing (ICCSP)*, pp. 2047-2050, 2017.
- [6] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of Twitter data," In *Proceedings of the Workshop on Languages in Social Media (LSM '11)*, Association for Computational Linguistics, 2011.
- [7] A. Kumar and T. M. Sebastian, "Sentiment Analysis on Twitter," *IJCSI International Journal of Computer Science Issues*, vol. 9, no. 4, 2012.
- [8] Aditya Joshi, Pushpak Bhattacharyya, and M. J. Carman, "Automatic Sarcasm Detection," *ACM Computing Surveys*, vol. 50, pp. 1-22, 2017.
- [9] S. Song, H. Huang, and T. Ruan, "Abstractive text summarization using LSTM-CNN based deep learning," *Multimedia Tools and Applications*, vol. 78, no. 1, pp. 857-875, 2019/01/01 2019.
- [10] S. Karthick, R. J. Victor, S. Manikandan, and B. Goswami, "Professional chat application based on natural language processing," in *2018 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC)*, pp. 1-4, 2018.
- [11] Duyu Tang, B. Qin, and T. Liu, "Document Modeling with Gated Recurrent Neural Network for Sentiment Classification," *Association for Computational Linguistics*, pp. 1422-1432, 2015.
- [12] B. Gokulakrishnan, P. Priyanthan, T. Ragavan, N. Prasath, and A. Perera, "Opinion mining and sentiment analysis on a Twitter data stream," in *International Conference on Advances in ICT for Emerging Regions (ICTer2012)*, pp. 182-188, 2012.