CURAJ_IIITDWD@LT-EDI-ACL2022: Hope Speech Detection in English YouTube Comments using Deep Learning Techniques

Vanshita Jha
Central University of Rajasthan, India
vanshitajha@gmail.com

Ankit Kumar Mishra
Goa University, India
ankitmishra.in.com@gmail.com

Sunil Saumya
Indian Institute of Information Technology Dharwad, India
sunil.saumya@iiitdwd.ac.in

Abstract

Hope Speech are positive terms that help to promote or criticise a point of view without hurting the user's or community's feelings. Non-Hope Speech, on the other side, includes expressions that are harsh, ridiculing, or demotivating. The goal of this article is to find the hope speech comments in a YouTube dataset. The datasets were created as part of the "*LT-EDI-ACL 2022:* Hope Speech Detection for Equality, Diversity, and Inclusion" shared task. The shared task dataset was proposed in Malayalam, Tamil, English, Spanish, and Kannada languages. In this paper, we worked at English-language YouTube comments. We employed several deep learning based models such as DNN (dense or fully connected neural network), CNN (Convolutional Neural Network), Bi-LSTM (Bidirectional Long Short Term Memory Network), and GRU(Gated Recurrent Unit) to identify the hopeful comments. We also used Stacked LSTM-CNN and Stacked LSTM-LSTM network to train the model. The best macro avg F1-score 0.67 for development dataset was obtained using the DNN model.The macro avg F1-score of 0.67 was achieved for the classification done on the test data as well.

1 Introduction

In recent years, the majority of the world's population has access to social media. The social media's posts, comments, articles, and other content have a significant impact on everyone's lives. People tend to believe that their lives on social media are the same as their real lives, therefore the influence of others' opinions or expressions is enormous (Priyadharshini et al., 2021; Kumaresan et al., 2021). People submit their posts to social networking platform and receive both positive and negative expressions from their peer users.

People in a multilingual world use a variety of languages to express themselves, including English, Hindi, Malayalam, French, and others

(Chakravarthi et al., 2021, 2020). While the most effective expression in real life is face or visual expression, which frequently delivers a much more efficient message than linguistic words, expressions in virtual life, such as social media, are frequently expressed through linguistic texts (or words) and emoticons. These words have a significant impact on one's life (Sampath et al., 2022; Ravikiran et al., 2022; Chakravarthi et al., 2022b; Bharathi et al., 2022; Priyadharshini et al., 2022). For example, if we respond to someone's social media post with "Well Done!", "Very Good", "Must do it again", "need a little more practise", and so on, it may instil confidence in the author. On the other hand, negative statements such as "You should not try it", "You don't deserve it", "You are from a different religion", and others demotivate the person. The comments that fall into the first group are referred to as "Hope Speech" while those that fall into the second category are referred to as "Non-hope speech" or "Hate Speech" (Kumar et al., 2020; Saumya et al., 2021; Biradar et al., 2021).

In the previous decade, researchers have worked heavily on hate speech identification in order to maintain social media clean and healthy. However, in order to improve the user experience, it is also necessary to emphasise the message of hope on these sites. To our knowledge, the shared task "*LT-EDI-EACL 2021*: Hope Speech Detection for Equality, Diversity, and Inclusion[1]" was the first attempt to recognise hope speech in YouTube comments. The organizers proposed the shared task in three different languages that is English, Tamil and Malayalam. Many research teams from all over the world took part in the shared task and contributed their working notes to describe how to identify the hope speech comments. (Saumya and Mishra, 2021) used a parallel network of CNN and LSTM with GloVe and Word2Vec embeddings and obtained a weighted F1-score of 0.91 for En-

---

[1] https://sites.google.com/view/lt-edi-2021/home

glish Dataset. Similarly, for Tamil and Malayalam they trained parallel Bidirectional LSTM model and obtained F1-score of 0.56 and 0.78 respectively. (Puranik et al., 2021) trained several fine tuned transformer models and identified that ULM-FiT is best for English with F1-score 0.9356. They also found that mBERT obtained 0.85 F1-score on Malayalam dataset and distilmBERT obtained 0.59 F1-score on Tamil dataset. For the same task a fine tuned ALBERT model was used by (Chen and Kong, 2021) and they obtained a F1-score of 0.91. Similarly, (Zhao and Tao, 2021; Huang and Bai, 2021; Ziehe et al., 2021; Mahajan et al., 2021) employed XLM-RoBERTa-Based Model with Attention for Hope Speech Detection. (Dave et al., 2021) experimented the conventional classifiers like logistic regression and support vector machine with TF_IDF character N-gram fearures for hope speech classification.

ACL 2022 will see the introduction of a new edition of the shared task "Hope Speech Detection for Equality, Diversity, and Inclusion." In contrast to *LT-EDI-EACL 2021*, this time the shared task *LT-EDI-ACL 2022* has been proposed in five different languages: English, Malayalam, Tamil, Kannada, and Spanish (Chakravarthi, 2020; Chakravarthi and Muralidaran, 2021; Chakravarthi et al., 2022a). The data was extracted via the YouTube platform. We took part in the competition and worked on the dataset of English hope speech comments. The experiments were carried out on several neural network based models such as a dense or multilayer neural network (DNN), one layer CNN network (CNN), one layer Bi-LSTM network (Bi-LSTM), and one layer GRU network (GRU), among deep learning networks. The stack connections of LSTM-CNN and LSTM-LSTM were also trained for hope speech detection. After all experimentation, it was found that DNN produced the best results with macro average F1-score of 0.67 on development as well as on test dataset.

The rest of the article is oraganized as follows. The next section 2 give the details of the given task and dataset statistics. This is followed by the description of methodology used for experimentation in Section 3. The results are explained in the Section 4. At the end, Section 5 talks about future scope of the research.
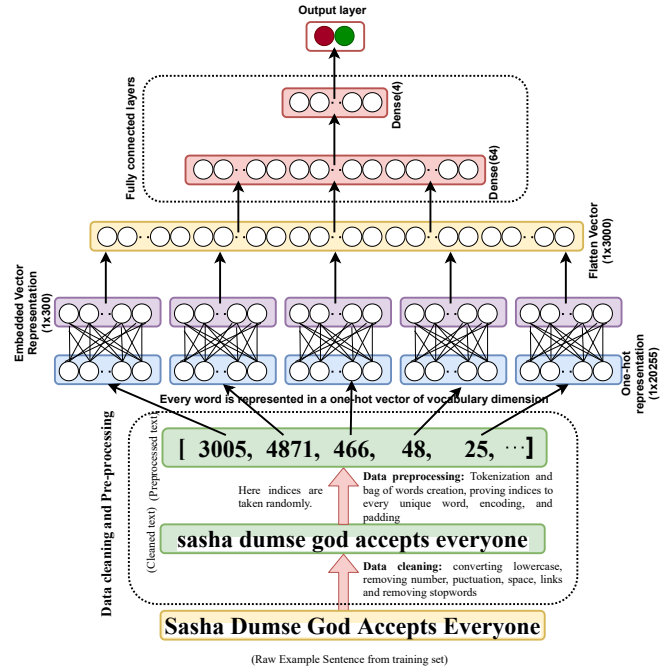

Figure 1: A DNN network for hope speech detection

2 Task and data description

At *LT-EDI-ACL 2022*, the shared task on Hope Speech Detection for Equality, Diversity, and Inclusion (provided in English, Tamil, Spanish, Kannada, and Malayalam) intended to determine whether the given comment was Hope speech or not. The dataset was gathered from the YouTube platform. In the given dataset, there were two fields for each language: comment and label. We only submitted the system for the English dataset. In the English training dataset, there were approximately 22740 comments, with 1962 labeled as hope speech and 20778 labeled as non-hope speech. There were 2841 comments in the development dataset, with 272 hope speech and 2569 non-hope speech comments. The test dataset contained 250 hope speech and 2593 non-hope speech comments. The English dataset statistics is shown in Table 1.

3 Methodology

Several deep learning models were developed to identify the hope speech from supplied English YouTube comments. The architecture of our best model DNN, as depicted in Figure 1, will be explained in this section. We also explain the architecture of stacked network LSTM-LSTM as shown in Figure 2.

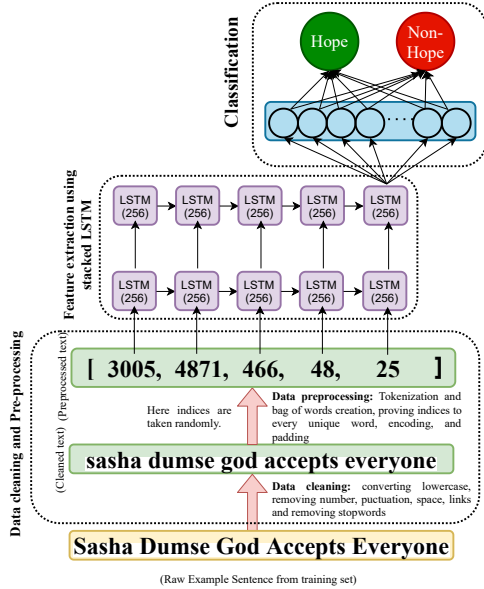|	Hopespeech	Non Hopespeech	Total
Training	1962	20778	22740
Development	272	2569	2841
Test	250	2593	2843
Total	2434	25940	28424

Table 1: English Dataset statistics

Figure 2: A stacked LSTM network for hope speech detection

3.1 Data cleaning and pre-processing

We used a few early procedures to convert the raw input comments into readable input vectors. We started with data cleaning and then moved on to data preprocessing. Every comment was changed to lower case during data cleaning. Numbers, punctuation, symbols, emojis, and links have all been removed. The *nltk* library was used to eliminate stopwords like the, an, and so on. Finally, the extra spaces were removed, resulting in a clean text. During data preprocessing, we first tokenized each comment in the dataset and created a bag of words with an index number for each unique word. The comments were then turned into an index sequence. The length of the encoded vectors was varied. After that, the encoded indices vector was padded to form an equal-length vector. In our case we kept the length of each vector as ten.

3.2 Classification Models

Several deep learning classification models were developed. We started with a multilayer dense neural network (DNN) as shown in Figure 1. After that,

a single layer CNN model and a single layer Bi-LSTM model were constructed. Finally, we built stacked LSTM-CNN and stacked LSTM-LSTM models shown in Figure 2. In Section 4, the results of each model are discussed. Regardless of the model, the feeding input and collecting output were the same in all instances. The whole process flow from input to output is depicted in Figures 1 and 2. As can be seen, there were three stages to the process: data preparation, feature extraction, and classification. The biggest distinction among the models was in the feature extraction criterion.

To demonstrate the model flow, a representative example from the English Dataset is used. The representative example "Sasha Dumse God Accepts Everyone." was first changed to lower case as "sasha dumse god accepts everyone.". During the data cleaning process, the dot(.) is eliminated. The lowercase text was then encoded and padded into a sequence list as "[3005, 4871, 466, 48, 25]". The index "3005" refers to the word "sasha", the index "4871" to the word "dumse," and so on. The sentence was padded into the length of ten.

After preprocessing the data, each index is turned into a one-hot vector (of 1x20255 dimension) with a size equal to the vocabulary. The resultant one-hot vector was sparse and high dimensional, and it was then passed through an embedding layer, yielding a low dimensional dense embedding vector (of 1x 300 dimension). Between the input and embedding layers, many sets of weights were used. We experimented with random weights as well as pre-trained Word2Vec and GloVe weights, but found that random weights initialization at the embedding layer performed better. As a result, we've only covered the usage of random weights at the embedding layer in this article. For abstract level feature extraction, the embedded vector was provided as an input to a stacked DNN or LSTM layer as shown in Figures 1 and 2 respectively. Finally, the collected features were classified into hope and non-hope categories using a dense (or an output) layer.

Table 2: Results of English Development dataset

Methods	Metrics	Non-Hope	Hope	Macro Avg	Weighted Avg
DNN	Precision	0.43	0.93	0.68	0.89
	Recall	0.38	0.95	0.66	0.89
	F1-score	0.40	0.94	0.67	0.89
CNN	Precision	0.39	0.93	0.66	0.88
	Recall	0.37	0.94	0.65	0.88
	F1-score	0.38	0.94	0.66	0.88
Bi-LSTM	Precision	0.39	0.94	0.66	0.88
	Recall	0.40	0.93	0.67	0.88
	F1-score	0.40	0.94	0.67	0.88
GRU	Precision	0.39	0.94	0.66	0.88
	Recall	0.38	0.94	0.66	0.88
	F1-score	0.38	0.94	0.66	0.88
LSTM-CNN	Precision	0.41	0.93	0.67	0.88
	Recall	0.35	0.95	0.65	0.89
	F1-score	0.38	0.94	0.67	0.87
LSTM-LSTM	Precision	0.34	0.94	0.64	0.88
	Recall	0.43	0.91	0.67	0.86
	F1-score	0.38	0.92	0.65	0.89

Table 3: Results of English test dataset

Metrices	Non-Hope	Hope	Macro Avg	Weighted Avg
Precision	0.40	0.94	0.67	0.89
Recall	0.38	0.94	0.66	0.90
F1 score	0.39	0.94	0.67	0.89

4 Results

All of the experiments were carried out in the Keras and sklearn environment. We used the pandas library to read the datasets. Keras preprocessing classes and the *nltk* library were used to prepare the dataset. All the results shown in Table 2 is on English development dataset. The initial experiment was with dense neural network (DNN). The three layers of dense network with relu activation (in the internal layer) and sigmoid activation at the output layer were trained with English comment dataset. Similarly, the experiments were performed with CNN, Bi-LSTM, GRU, LSTM-CNN, and LSTM-LSTM. The best result was obtained by DNN with macro average F1-score 0.67. The results of other models are shown in Table 2. Later, once the labels for test dataset was released by the organizers, we also collected the model performance on test dataset. The macro average F1-score achieved after categorising the test data from the DNN model was 0.67, which was the same as in the case of development data. Table 3 lists the test dataset results produced from the DNN model.

5 Conclusion

As part of the joint task *LT-EDI-ACL2022*, we presented a model provided by team *CURJ_IIITDWD* for detecting hope speech on an English dataset obtained from the YouTube platform. We used many deep learning algorithms in the paper and found that DNN with three hidden layers performed best on the development and test dataset with a macro average F1-score of 0.67. In the future, we can improve classification performance by training transfer learning models like BERT and ULMFiT ans so on.

References

B Bharathi, Bharathi Raja Chakravarthi, Subalalitha Chinnaudayar Navaneethakrishnan, N Sripriya, Arunaggiri Pandian, and Swetha Valli. 2022. Findings of the shared task on Speech Recognition for Vulnerable Individuals in Tamil. In *Proceedings of the Second Workshop on Language Technology for*

Equality, Diversity and Inclusion. Association for Computational Linguistics.

Shankar Biradar, Sunil Saumya, and Arun Chauhan. 2021. Hate or non-hate: Translation based hate speech identification in code-mixed hinglish data set. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 2470–2475. IEEE.

Bharathi Raja Chakravarthi. 2020. HopeEDI: A multilingual hope speech detection dataset for equality, diversity, and inclusion. In *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media*, pages 41–53, Barcelona, Spain (Online). Association for Computational Linguistics.

Bharathi Raja Chakravarthi and Vigneshwaran Muralidaran. 2021. Findings of the shared task on hope speech detection for equality, diversity, and inclusion. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 61–72, Kyiv. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Vigneshwaran Muralidaran, Ruba Priyadharshini, Subalalitha Chinnaudayar Navaneethakrishnan, John Phillip McCrae, Miguel Ángel García-Cumbreras, Salud María Jiménez-Zafra, Rafael Valencia-García, Prasanna Kumar Kumaresan, Rahul Ponnusamy, Daniel García-Baena, and José Antonio García-Díaz. 2022a. Findings of the shared task on Hope Speech Detection for Equality, Diversity, and Inclusion. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Vigneshwaran Muralidaran, Ruba Priyadharshini, and John Philip McCrae. 2020. Corpus creation for sentiment analysis in code-mixed Tamil-English text. In *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)*, pages 202–210, Marseille, France. European Language Resources association.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Thenmozhi Durairaj, John Phillip McCrae, Paul Buitaleer, Prasanna Kumar Kumaresan, and Rahul Ponnusamy. 2022b. Findings of the shared task on Homophobia Transphobia Detection in Social Media Comments. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*. Association for Computational Linguistics.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Rahul Ponnusamy, Prasanna Kumar Kumaresan, Kayalvizhi Sampath, Durairaj Thenmozhi, Sathiyaraj Thangasamy, Rajendran Nallathambi, and John Phillip McCrae. 2021. Dataset for identification of homophobia and transophobia in multilingual YouTube comments. *arXiv preprint arXiv:2109.00227*.

Shi Chen and Bing Kong. 2021. cs_english@ LT-EDI-EACL2021: Hope Speech Detection Based On Fine-tuning ALBERT Model. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 128–131.

Bhargav Dave, Shripad Bhat, and Prasenjit Majumder. 2021. IRNLP_DAIICT@ LT-EDI-EACL2021: hope speech detection in code mixed text using TF-IDF char n-grams and MuRIL. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 114–117.

Bo Huang and Yang Bai. 2021. TEAM HUB@ LT-EDI-EACL2021: hope speech detection based on pre-trained language model. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 122–127.

Abhinav Kumar, Sunil Saumya, and Jyoti Prakash Singh. 2020. NITP-AI-NLP@ HASOC-Dravidian-CodeMix-FIRE2020: A Machine Learning Approach to Identify Offensive Languages from Dravidian Code-Mixed Text. In *FIRE (Working Notes)*, pages 384–390.

Prasanna Kumar Kumaresan, Ratnasingam Sakuntharaj, Sajeetha Thavareesan, Subalalitha Navaneethakrishnan, Anand Kumar Madasamy, Bharathi Raja Chakravarthi, and John P McCrae. 2021. Findings of shared task on offensive language identification in Tamil and Malayalam. In *Forum for Information Retrieval Evaluation*, pages 16–18.

Khyati Mahajan, Erfan Al-Hossami, and Samira Shaikh. 2021. Teamuncc@ lt-edi-eacl2021: Hope speech detection using transfer learning with transformers. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 136–142.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Subalalitha Chinnaudayar Navaneethakrishnan, Thenmozhi Durairaj, Malliga Subramanian, Kogilavani Shanmugavadivel, Siddhanth U Hegde, and Prasanna Kumar Kumaresan. 2022. Findings of the shared task on Abusive Comment Detection in Tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Ruba Priyadharshini, Bharathi Raja Chakravarthi, Sajeetha Thavareesan, Dhivya Chinnappa, Durairaj Thenmozhi, and Rahul Ponnusamy. 2021. Overview of the DravidianCodeMix 2021 shared task on sentiment detection in Tamil, Malayalam, and Kannada. In *Forum for Information Retrieval Evaluation*, pages 4–6.

Karthik Puranik, Adeep Hande, Ruba Priyadharshini, Sajeetha Thavareesan, and Bharathi Raja Chakravarthi. 2021. IIITT@ LT-EDI-EACL2021-hope speech detection: there is always hope in transformers. *arXiv preprint arXiv:2104.09066*.

Manikandan Ravikiran, Bharathi Raja Chakravarthi, Anand Kumar Madasamy, Sangeetha Sivanesan, Ratnavel Rajalakshmi, Sajeetha Thavareesan, Rahul Ponnusamy, and Shankar Mahadevan. 2022. Findings of the shared task on Offensive Span Identification in code-mixed Tamil-English comments. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Anbukkarasi Sampath, Thenmozhi Durairaj, Bharathi Raja Chakravarthi, Ruba Priyadharshini, Subalalitha Chinnaudayar Navaneethakrishnan, Kogilavani Shanmugavadivel, Sajeetha Thavareesan, Sathiyaraj Thangasamy, Parameswari Krishnamurthy, Adeep Hande, Sean Benhur, Kishor Kumar Ponnusamy, and Santhiya Pandiyan. 2022. Findings of the shared task on Emotion Analysis in Tamil. In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*. Association for Computational Linguistics.

Sunil Saumya, Abhinav Kumar, and Jyoti Prakash Singh. 2021. Offensive language identification in dravidian code mixed social media text. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 36–45.

Sunil Saumya and Ankit Kumar Mishra. 2021. IIIT_DWD@ LT-EDI-EACL2021: hope speech detection in YouTube multilingual comments. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 107–113.

Yingjia Zhao and Xin Tao. 2021. ZYJ@ LT-EDI-EACL2021: XLM-RoBERTa-based model with attention for hope speech detection. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 118–121.

Stefan Ziehe, Franziska Pannach, and Aravind Krishnan. 2021. GCDH@ LT-EDI-EACL2021: XLM-RoBERTa for hope speech detection in English, Malayalam, and Tamil. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 132–135.