All in One: A Multi-Task Learning for Emoji, Sentiment and Emotion Analysis in Code-Mixed Text

Anonymous ACL submission

Abstract

Code mixed language and emojis are being extensively used in social media to express opinions. In this paper, we propose a novel task that focuses on suggesting appropriate emojis in English-Hindi code-mixed sentences. We aim to exploit the dependency between emotion, sentiment, and emojis for building an end-to-end framework that can simultaneously identify the emotion, sentiment and emojis in code-mixed sentences. We introduce the *Code-Mixed Emoji, Emotion and Sentiment aware Dataset* (CMEESD) which is an extension of the *SemEval 2020 Task 9*. We establish strong baselines to predict the correct emojis by simultaneously identifying the emotion and sentiment of a given tweet. The sentiment and emotion prediction in turn helps for the appropriate emoji classification. Empirical results on the *CMEESD* dataset demonstrate that the proposed multi-task framework yields better performance over the single-task framework.

1 Introduction

Emoji is an essential aspect of our daily conversation and adds more meaning to the language. Recently with the extensive use of different social media platforms emoji prediction (Barbieri et al., 2018; Jin and Pedersen, 2018; Wang and Pedersen, 2018; Wu et al., 2018) has become an important task in the field of Natural Language Processing (NLP). Related tasks are often inter-dependent and correlated, therefore they perform better when they are handled simultaneously. We hypothesize that emojis are closely related to sentiment[1] and emotion[2]. This can be easily depicted through the example - *"Some people are just so selfish 🤮."*. This tweet, at the first glance, conveys that the person is extremely sad with some people's behaviour. But careful observation of the sentiment and emotion of the person helps us understand that the person is disgusted with these types of selfish people and has a negative sentiment during the tweet. This is where sentiment and emotion come into the picture. Sentiment, emotion, and emoji are highly intertwined, and one helps in understanding the other better.

No.	Utterances	Emoji	Sent	Emotion
1	*LoL @ Squirrel Tony stark ka dil toh ye leke ghum rha hai*	😂	Pos	Joy
2	*Happy Birthday Doctor sahab bhagwaan aapko khush or swasth rkhe or aap hme aise hi apni creativity se hansaate rhe*	😊	Pos	Joy
3	*@ PiyushGoyalOffc sir aam public ko dikat hoti hai safar ke dauran ticket nhi milne pe aur agent log 500hundred ka 1 dete hein*	😡	Neg	Anger
4	*@ Mastani4423509 Tu Safar Mera Tu Hi Meri Manzil .... Tere Bina Guzara Aye DiL Hai Mushkil .. I LOVE U @ iamsrk*	❤️	Pos	Joy
5	*@AbidSherAli Look who is talking Jo jhoot moot k bimaar ban k bahar baithay hain*	🤮	Neg	Disgust

Table 1: Some samples from CMSEED.

Monolingual discussions are not common in most parts of the world. It is more natural for people to transition between two (or more) languages while expressing themselves. The phenomena of code-switching or code-mixing occurs when a speaker regularly switches between two or more languages while speaking. This type of communication is quite typical among peers who are fluent in multiple languages. Even textual discussions, which mostly take place on social media sites, such as Twitter, Instagram, and Reddit, are mainly code-mixed. Some instances of code-mixed tweets are as follows: *"What a wicket ! @ iamamirofficial we love you . Kya kar diya apney"*. In every day life people often switch between languages while expressing their feelings or opinions making the text code-mixed in nature.

Although there are prior research that focused

[1]Determine the opinion (i.e., positive, negative, & neutral) expressed by a person for a topic, event, or product.

[2]Determine the emotion displayed by a person on a topic, event, or product (i.e., angry, disgust, fear, joy, sad, & surprise)

on determining the relationships between emoji and emotion (Shoeb and de Melo, 2020; Hussien et al., 2019; Hayati and Muis, 2019), emoji and sentiment (Tomihira et al., 2020; Al-Halah et al., 2019; Felbo et al., 2017; Chen et al., 2018b), but no attempt has been made so far that focuses on capturing the relationship between emoji, sentiment and emotion simultaneously in a multi-task framework. In Table 1, we present few examples from the CMEESD dataset. Sentiment and emotion are correlated hence have been known to improve the performance of each other when jointly performed. As emojis express emotions therefore by using the emotion information explicitly can help capturing the emoji correctly. Therefore, as sentiment helps in correct emotion classification(Gao et al., 2013; Sahay et al., 2018; Yu et al., 2018) which in turn assist in emoji predictions therefore it can be said that these tasks inherently are dependent on each other and when performed concurrently can improve the performance of each other.

The task becomes more challenging when code-mixed data is considered for implicitly capturing the relationship between emotion, emoji and sentiment for the correct prediction in a given tweet. In our current work, we build an end-to-end multi-task framework to leverage the sentiment and emotion information for solving the problem of emoji detection and vice versa. Further, to the best of our knowledge, this is the very first attempt at solving the emoji prediction with the help of sentiment and emotion together in multi-task framework in code-mixed data.

The main contributions and/or attributes of our work are as follows: **a).** We propose the task of emoji, emotion and sentiment prediction in code-mixed text capturing the relationship between them in a multi-task framework; **b)** We introduce a *Codemixed Emoji Emotion Sentiment aware Dataset* (CMEESD), an extension of *task 9 @SemEval2020* in terms of diverse emojis (i.e. positive and negative emojis), sentiment labels and emotion labels; and **c)** We establish strong multi-task baselines for predicting the emotion, emoji and sentiment simultaneously from a given code-mixed tweet.

2 Related Work

Review of the existing research (Barbieri et al., 2018; Jin and Pedersen, 2018; Wang and Pedersen, 2018; Eisner et al., 2016; Zhou and Wang, 2017; Al-Halah et al., 2019; Felbo et al., 2017; Chen et al., 2018b; Cappallo et al., 2018; Yeh et al., 2019; Chen et al., 2018a; Cowie et al., 2001) suggests that emoji, sentiment and emotion analysis are important areas in the field of Natural Language Processing (NLP). Recently, authors in (Barbieri et al., 2017) proposed several Long Short Term Memory (LSTM) based frameworks for single label emoji prediction. In (Barbieri et al., 2018; Jin and Pedersen, 2018; Wang and Pedersen, 2018), the authors proposed a classifier for multi-lingual emoji prediction for English and Spanish languages. The authors in (Eisner et al., 2016) released emoji2vec pre-trained embeddings. As emoticons are extensively used, therefore many researchers have focused on its usage in different works such as for emoji recommendation in instant messages (Guibon et al., 2018), emoji sense disambiguation (Wijeratne et al., 2017), understanding crisis events (Santhanam et al., 2019), building emotion classifiers (Hussien et al., 2019), sentiment analysis (Al-Halah et al., 2019; Felbo et al., 2017; Chen et al., 2018b) and emotional response generation (Zhou and Wang, 2017). Lately, (Ma et al., 2020) proposed transformer based network for multi-label emoji prediction. Recently, in (Chakravarthi et al., 2021) a Dravidian code-mixed data was proposed for identifying the sentiments and offensive languages. The dataset comprised of Tamil-English, Kannada-English, and Malayalam-English texts. In (Yadav and Chakraborty, 2020) methods that use different kinds of multilingual and cross-lingual embeddings to efficiently transfer knowledge from monolingual text to code-mixed text for sentiment analysis of code-mixed text was proposed. Lately, (Wang et al., 2016a) proposed a joint factor graph model for identifying emotions in code-mixed data. A Bilingual Attention Network (BAN) model was proposed in (Wang et al., 2016b) to aggregate the monolingual and bilingual informative words to form vectors from the document representation, and integrate the attention vectors to predict the emotion in code-mixed data.

Our current work differentiates from the existing works on emoji prediction as we aim to leverage the combined sentiment and emotion information for solving the problem of emoji detection, emotion classification and sentiment analysis in a multi-task framework in a code-mixed text. Further, to the best of our knowledge, this is the very first attempt at solving all the three tasks simultaneously in multi-task framework for code-mixed text.

3 Dataset

SemEval2020 task9 (Patwa et al., 2020)[3] dataset consists of approx. 20000 tweets, and each tweet is accompanied by one sentiment(positive, negative and neutral) .We then propose *Code Mixed Emoji Emotion Sentiment Aware Dataset* (CMEESD[4]) by annotating the *SemEval2020 task9* dataset with emotion (i.e., angry, disgust, joy, sad, neutral), and emoji (😡, 😠, 😞, 😫, 😊, 😐, 😢, ❤️, 🤮, 😘, 😂) labels . We show some samples from CMEESD in the Table 1. We divide the CMEESD (c.f. Table 2) into three sets i.e., train set, development set (dev set), and test set.

Statistics		**CMSEED Dataset**		
		Train	***Dev***	***Test***
Sentiment	*#Tweets*	14000	3000	3000
	#Positive	4,634	982	1000
	#Negative	4102	890	900
	#Neutral	5,264	1,128	1100
Emotion	*#Joy*	3029	886	786
	#Anger	2640	369	572
	#Disgust	951	103	131
	#Sad	2286	514	411
	#Neutral	5,264	1,128	1100

Table 2: Dataset statistics with sentiment and emotion.

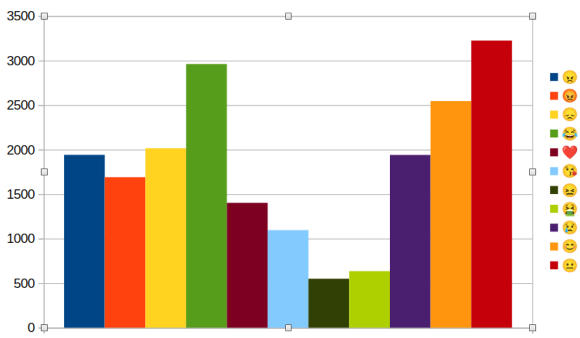

Figure 1: Emoji Distribution for full Dataset

3.1 Data Annotation

Due to the absence of emotion and emoji labels in CMEESD, we employ three annotators proficient in English and Hindi languages to label every tweet. For annotating the dataset, we consider Ekman's universal emotions, *viz.* Joy, Sadness, Anger and Disgust as emotion labels for the tweets along with neutral label for tweets having no emotions.

[3]https://competitions.codalab.org/competitions/26655#participate

[4]We will release the code and data.

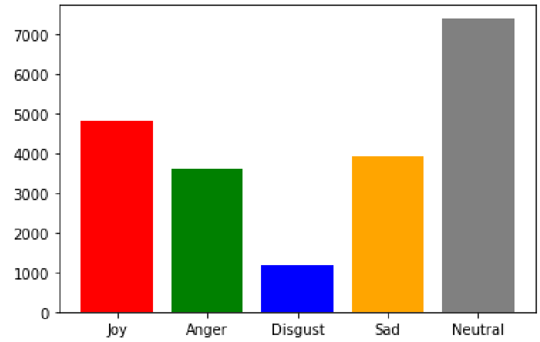

Figure 2: Emotion Distribution for full Dataset

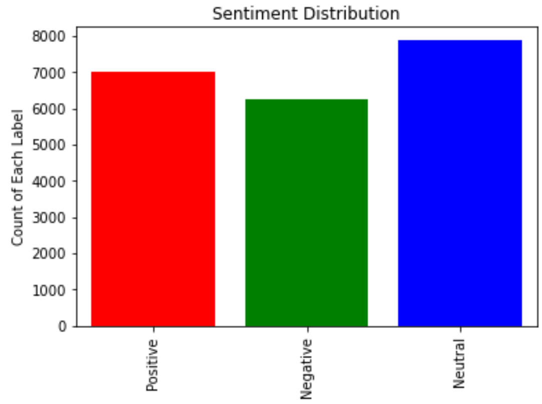

Figure 3: Sentiment Distribution for full Dataset

3.2 Inter-Annotator Agreement

A majority voting scheme was used for selecting the final emoji and emotion label. We achieve an overall Fleiss' (Fleiss, 1971) kappa score of 0.81 and 0.75, which are considered to be reliable.

The statistics of the CMEESD dataset are given in Table 2. We also show the distribution of Emoji, Emotion and Sentiment in the dataset as depicted in Figure 1, Figure 2 and Figure 3, respectively. In Figure 4, we present the correlation between the different labels of sentiment, emotion and emoji. From the figure it is evident that all the three tasks are highly correlated and dependent on one another.

4 Methodology

4.1 SentencePiece Tokenizer

SentencePiece (Kudo and Richardson, 2018) considers the tweets as a sequence of unicode letters. It uses byte-pair-encoding (BPE)(Sennrich et al., 2015) and the unigram language model(Kudo, 2018) to handle sentences as sequences of Unicode letters to make them sub-words. The byte pair en-

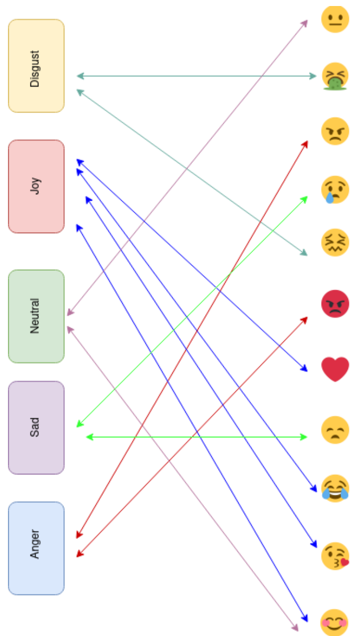

Figure 4: Emotion and Emoji dependency in the CMEESD dataset

coding initializes the vocabulary with every character in the corpus and learns a set of merge rules over time. Multiple subword segmentations are probabilistically sampled during training for the unigram language model.

4.2 Codemixed Embedding Generation

Pre-trained embedding have an issue with code-mixing that it will give more out-of-vocabulary (OOV) words(Pratapa et al., 2018). We train the word embedding with the available code-mixed corpus itself, but one important issue of any code-mixed data to decide which embedding model to be used for better performance. We, therefore, perform two different embedding and concatenate them to obtain the better output.

Char level word embedding: As we all know that code-mixed data have the challenge of OOV words, so we follow (Chiu and Nichols, 2016), for character level word embedding to extract the character level features. Because Recurrent Neural Network-based encodings do not significantly outperform CNNs while being computationally more expensive to train, we utilize a convolutional neural network (CNN) (followed by a max pooling layer) for simplicity of training.(Ling et al., 2015)

Contextual level word embedding for the contextual representation we used ELMO(Peters et al., 2018). Each token in ELMO is represented as a vector that functions as a function of the entire sentence (A word might therefore have various meanings depending on the context from which it was taken).

4.3 Baselines

We aim to leverage the sentiment and emotion information for solving the problem of emoji detection in a multi-task framework for code-mixed dataset, and vice versa. We use two strong baselines, defined as follows.

4.3.1 CM-BiLSTM:

BiLSTM (Schuster and Paliwal, 1997) is a sequence processing model which processes input in forward direction as well as in backward direction. For the embedding purpose, we use two embeddings which have been described previously in the embedding section.

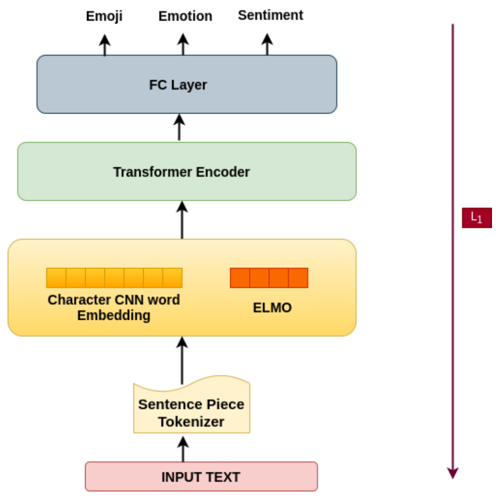

Figure 5: Architectural diagram CM-Trans

4.3.2 CM-Trans and CM-HTrans:

We discover that the attention mechanism is more effective at determining which parts of a phrase are necessary for capturing the sentiment (Patwa et al., 2020). As a result, for our code-mixed multitask emoji analysis, we picked the transformer model(Vaswani et al., 2017). We use transformer to capture contextualized representation for encoding the tweets for classification as shown in Figure 5. Conventionally, the input of the transformer encoder is basically the embedding of each word e_i in a given tweet $T = w_1, w_2, \ldots, w_n$, where w represents the words in a tweet having n number of words with e as their embedding along with the positional embedding PE_i of the word. But in our case, for improving the efficacy of the model, we

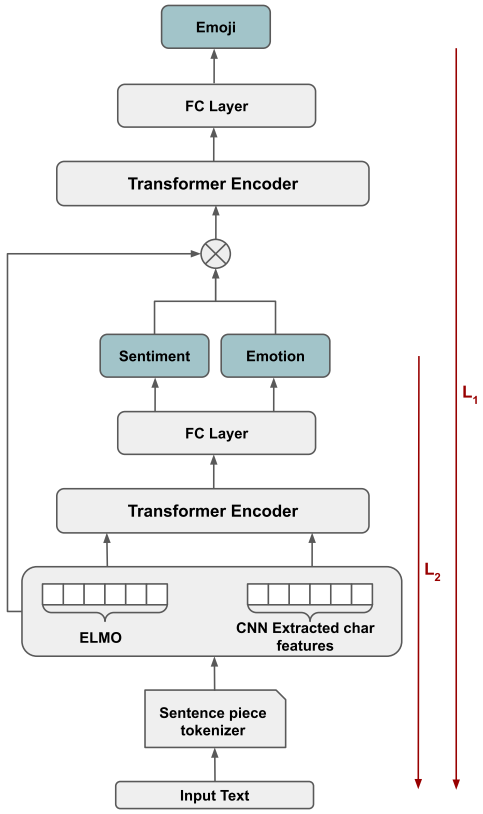

Figure 6: Architectural diagram CM-HTrans

utilize the embeddings from Elmo and CNN for effectively capturing the meaning of the code-mixed tweets (as discussed above).

Therefore, the input to the transformer encoder is $\mathcal{E}_i = e_{Elmo,i} + e_{CNN,i}$ as the concatenated embeddings of the word together with the PE_i positional embedding. The encoder creates a sequence of context vectors from the tweets through a succession of N_x encoder layers. Each layer has sub units- a Multi-head attention layer and a position-wise feedforward layer.

The encoder layers are a crucial module that handles all of the input sequence processing. We start by passing the source phrase and its mask to the multi-head attention layer, then dropout, apply a residual connection, and normalize it. We next apply dropout, a residual connection, and layer normalization to the encoded output sequence after passing it via a position-wise feedforward layer.

The Transformer model employs scaled dot-product attention, as shown in 1, in which the query Q and key K are merged using the dot product, followed by the softmax operation, and scaled by a scaling factor d_k before being multiplied by the value V. In the Transformer model, attention is a crucial unit since it aids in determining which portions of the sequence are significant.

$$Attention(Q, K, V) = softmax(\frac{KQ^{\mathrm{T}}}{\sqrt{d_k}})V. \quad (1)$$

The position-wise feedforward layer is the other major component of the encoder layer. The data is converted from hid dim to pf_d, with pf_d often being much greater than h_d. Before it is converted back into a h_d representation, the ReLU activation function and dropout are applied. The concept is based on infinitely large neural networks. The broad neural network provides more approximation capability and speeds up model optimization.

The output of the first FC layer i.e., *emotion* and *sentiment* information along with the word embeddings are fed as input again to the transformer encoder (CM-HTrans) (thereby forming a hierarchical framework) followed by FC layer to predict the final emoji of the given sentence as shown in Figure 6. The emotion and sentiment knowledge eventually helps in better emoji prediction.

4.3.3 Multi-task loss function

The main objective of our loss function is to teach the model how to weight the task specific losses. For this, we adopt a principled approach to multi-task deep learning that considers the homoscedastic uncertainty (Aleatoric uncertainty that is not reliant on the input data is known as task dependant or homoscedastic uncertainty. It is not a model output, but rather a number that is constant across all input data and changes between tasks. As a result, it is known as task-dependent uncertainty.) (Kendall et al., 2018) of each task while weighing multiple loss functions.

$$L_{total} = \sum_i W_i L_i \quad (2)$$

Where i defines the different tasks (Emoji, Emotion, Sentiment). We are updating the weights by using back-propagation for specific losses for each tasks.

5 Experimental results and analysis

We implement our proposed model in PyTorch, a Python-based deep learning library. We perform *grid search* to find the optimal hyper-parameters in Table 3. As evaluation metrics, we use accuracy and F1-score for the classification problems to show the performance of our proposed model. We use *Adam* as an optimizer. We use *Softmax* as a classifier for emoji, sentiment and emotion classification we use the multitask loss function as described previously in equation2.

5.1 Experimental Setup

We address three different tasks i.e. emoji, sentiment, and emotion analysis in a multi-task framework. We define the following experimental setups.

- **Emoji Classification (E^M)**
  - There are eleven different emojis in the CMEESD and only one emoji is associated with each tweet.
  - We use a one-hot vector to represent emoji classes corresponding to each tweet for the experiment.
- **Sentiment Classification (S_C)**
  - There are three sentiment classes i.e., positive, neutral, and negative. Only one sentiment class is associated with each tweet.
  - We use a one-hot vector to represent sentiment classes corresponding to each tweet for the experiment.
- **Emotion Classification (E_C):**
  - There are five emotion classes (i.e., angry, disgust, joy, sad, and neutral) and only one emotion is associated with each tweet.
  - We use a one-hot vector to represent emotion classes corresponding to each tweet for the experiment.

5.2 Result and Analysis

We solve three different problems, namely, emoji analysis, sentiment analysis, and emotion analysis. We evaluate our proposed approach for all the possible combinations of the tasks i.e., Uni task learning (UTL), Dual task learning (DTL), and Tri task learning (TTL)

5.2.1 Emoji Classification (E^M):

We show the emoji classification results in Table 7. For *TTL*, our model achieves 7.35 and 4.51 points improvement in F1-score compared to *UTL* and *DTL*, respectively. We see improvement of 6.78 and 3.48 improvement in accuracy also. We observe that the proposed approach yields better performance for the *TTL* than the *DTL* and *UTL*. This improvement implies that our proposed hypothesis is correct and very effective.

5.2.2 Sentiment Classification (S_C):

We show the sentiment classification results in Table 5. For *TTL*, our model achieves 5.28 and 2.85 points F1-score improvement compared to *UTL* and *DTL*, respectively. We see improvement of 5.41 and 3.37 improvement in accuracy also. We observe that the proposed approach yields better performance for the *TTL* than the *DTL* and *UTL*. Thus, we can say emoji and emotion E_C help to sentiment class (S_C).

5.2.3 Emotion Classification (E_C):

We show the emotion classification results in Table 6. Similar to sentiment classification, we observe that the proposed approach yields better performance for the *TTL* than the *DTL* and *UTL*.

Parameters	CMEESD
Transformer Encoder Layer	2
Embeddings	300
FC Layer	1024, Dropout=0.3
Activations	*ReLu* as activation for our model
Output	Softmax (E^M, S_C, E_C)
Optimizer	Adam (lr=0.003)
Model Loss	Cross-entropy (Classification)
Batch	32
Epochs	30

Table 3: Hyper-parameters for our experiments where *N*, *D*, E^M, S_C, and E_C stands for #neurons, dropout, emoji, sentiment classification, and emotion classification, respectively.

6 Error Analysis

In this section, we present the error analysis of our proposed multitask framework. We stated earlier that emoji, sentiment, and emotion are highly related to each other. To show the effect of these tasks on each other, we take some examples from CMEESD dataset (c.f. Table 4). Second tweet (T_1) in Table 4 "*Har bar jab batting line flop karti ha Sara kasor imam bichare par kuon dala jata hai ?*" has emoji 😡 with negative sentiment and Anger emotion. Our *TTL* predicts the emoji correctly while *DTL* fails to predict the correct emoji (😖) and emotion(Disgust). We observe that sentiment and emotion together help to predict the correct emoji. In other words, we can say sentiment and emotion also help each other.

Correct Prediction							
Code-Mixed Tweet	**English Tweet**	**Actual**			**Predicted**		
		Emoji	***Emotion***	***Sentiment***	***Emoji***	***Emotion***	***Sentiment***
@ NoorHSumra I wish my dad was still alive miss him a lot and I love ur dad's response khich ke maar saale ko	@ NoorHSumra I wish my dad was still alive miss him a lot and I love ur dad's response slap tightly	❤️	Joy	Positive	❤️	Joy	Positive
Har bar jab batting line flop karti ha Sara kasor imam bichare par kuon dala jata hai ?	Every time the batting line flops all the blame is put on Poor Imam?	😡	Anger	Negative	😡	Anger	Negative
Congress ki sarker mai cylinder he gayab ho gaya tha	During congress govt. cylinder went missing	🤮	Disgust	Negative	🤮	Disgust	Negative
Incorrect Prediction							
Code-Mixed Tweet	**English Tweet**	**Actual**			**Predicted**		
		Emoji	***Emotion***	***Sentiment***	***Emoji***	***Emotion***	***Sentiment***
tere ghamand k karan hi aaj congress k ye halat hai ... failure hai tu Bhai .. Tujhse na ho payega	Because of your pride, this is the condition of Congress today. .. you are failure..you cant do this	😡	Anger	Negative	🤮	Disgust	Negative
You better send me eid mubarak note in your voice beti...	You better send me eid mubarak note in your voice	😐	Neutral	Neutral	😡	Anger	Negative
I really wanna meet him to show my love with my hand.. saale ne jeena muskil kar rakha hai..	I really wanna meet him to show my love with my hands .. he made my life miserable..	😡	Anger	Negative	❤️	Joy	Positive

Table 4: Predictions of the proposed framework for Emoji, Emotion and Sentiment

Tasks		**Embeddings**		**CM-BiLSTM**		**CM-Trans**		**CM-HTrans**	
		CF	***ELMO***	***F1***	***Acc***	***F1***	***Acc***	***F1***	***Acc***
UTL	S^C	√	-	68.81	70.39	72.21	74.33	72.21	74.33
	S^C	√	√	69.29	71.32	73.98	75.87	73.98	75.87
DTL	$E^M + S_C$	√	-	69.43	72.91	74.69	76.71	74.69	76.71
	$E^M + S_C$	√	√	71.72	73.48	76.41	77.35	76.41	77.35
TTL	$E_C+E^M + S_C$	√	-	73.31	74.53	78.54	79.27	-	-
	$E_C+E^M + S_C$	√	√	74.89	77.49	**81.11**	**81.93**	-	-

Table 5: Results and ablation Study of our proposed framework for Sentiment Classification. Best model result is 0.75 for sentiment, Described in (Patwa et al., 2020)

Tasks		**Embeddings**		**CM-BiLSTM**		**CM-Trans**		**CM-HTrans**	
		CF	***ELMO***	***F1***	***Acc***	***F1***	***Acc***	***F1***	***Acc***
UTL	E^C	√	-	61.81	63.32	65.95	67.83	65.95	67.83
	E^C	√	√	63.30	64.72	67.15	68.25	67.15	68.25
DTL	$E^M + E_C$	√	-	64.19	66.91	68.73	70.82	68.73	70.82
	$E^M + E_C$	√	√	66.26	67.15	68.91	71.41	68.91	71.41
TTL	$S_C+E^M + E_C$	√	-	68.37	69.19	71.54	73.87	-	-
	$S_C+E^M + E_C$	√	√	70.71	72.51	**73.91**	**76.21**	-	-

Table 6: Results and ablation Study of our proposed framework for Emotion Classification

While in some tweets, TTL fails to predict the correct emoji. For example, sixth tweet (T_2) as given in Table 4 "I really wanna meet him to show my love with my hand..
saale ne jeena muskil kar rakha hai.." (I really wanna meet him to show my love with my hand.. that idiot has made my life hell..) has ❤️ emoji but *TTL* fails to predict 😡 emoji. Similarly, *TTL* predicts the emotion as *joy* while the correct emotion of the tweet is *anger*. As the emotion of the tweet is clearly understandable by the hindi code-mixed part of the tweet, therefore the inability to properly capture the Hindi meaning of the tweet lead to the misclassification.

7 Conclusion and Future Work

In this paper, we have proposed a *Code-Mixed Emoji, emotion and Sentiment aware Dataset* (CMEESD) which is an extension of the *SemEval 2020 Task 9* in terms of diverse emojis (i.e. positive and negative emojis), sentiment labels and emotion labels. We also propose several strong multi-task baselines (i.e., CM-BiLSTM, CM-Transformer, CM-HTransformer) and we proposed (CM-HTransformer) to simultaneously solve all the three problems, *viz.* emoji analysis, sentiment analysis, and emotion analysis. Empirical results on CMEESD dataset indicates that the proposed multi-task framework yields better performance over the single-task learning.

During our analysis, we found that more than one emoji is possible for a given tweet. So, we will try to make a group of emojis (multi-emoji) corresponding to each tweet and perform multi-label emoji prediction with sentiment and emotion in code-mixed tweets.

Tasks		**Embeddings**		**CM-BiLSTM**		**CM-Trans**		**CM-HTrans**	
		CF	***ELMO***	***F1***	***Acc***	***F1***	***Acc***	***F1***	***Acc***
UTL	E^M	√	-	61.34	62.18	63.72	65.43	64.61	66.73
	E^M	√	√	62.81	63.32	64.95	66.25	65.76	67.43
DTL	$S_C + E^M$	√	-	63.12	63.91	65.60	66.91	66.59	67.83
	$S_C + E^M$	√	√	63.85	64.31	66.43	67.21	67.73	68.91
	$E_C + E^M$	√	-	64.73	65.96	67.21	68.30	68.83	69.91
	$E_C + E^M$	√	√	66.37	66.91	67.87	69.41	68.60	70.73
TTL	$E_C+S_C + E^M$	√	-	66.19	67.21	69.54	71.91	71.13	73.24
	$E_C+S_C + E^M$	√	√	67.41	68.19	71.54	72.25	**73.11**	**74.21**

Table 7: Results and ablation Study of our proposed framework for Emoji Classification

8 Ethical Consideration

The dataset used in this paper is freely available and we extend the dataset by annotating (Emotion and Emoji) the dataset, and has been used only for the purpose of academic research. The annotation for extending the dataset was done by human experts, who are the regular employee of our research group. There are no other issues to declare.

References

Ziad Al-Halah, Andrew Aitken, Wenzhe Shi, and Jose Caballero. 2019. Smile, be happy:) emoji embedding for visual sentiment analysis. In *Proceedings of the*

IEEE International Conference on Computer Vision Workshops, pages 0–0.

Francesco Barbieri, Miguel Ballesteros, and Horacio Saggion. 2017. Are emojis predictable? *arXiv preprint arXiv:1702.07285*.

Francesco Barbieri, Jose Camacho-Collados, Francesco Ronzano, Luis Espinosa Anke, Miguel Ballesteros, Valerio Basile, Viviana Patti, and Horacio Saggion. 2018. Semeval 2018 task 2: Multilingual emoji prediction. In *Proceedings of The 12th International Workshop on Semantic Evaluation*, pages 24–33.

Spencer Cappallo, Stacey Svetlichnaya, Pierre Garrigues, Thomas Mensink, and Cees GM Snoek. 2018. New modality: Emoji challenges in prediction, anticipation, and retrieval. *IEEE Transactions on Multimedia*, 21(2):402–415.

Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Navya Jose, Shardul Suryawanshi, Elizabeth Sherly, and John P McCrae. 2021. Dravidiancodemix: Sentiment analysis and offensive language identification dataset for dravidian languages in code-mixed text. *arXiv preprint arXiv:2106.09460*.

Sheng-Yeh Chen, Chao-Chun Hsu, Chuan-Chun Kuo, Lun-Wei Ku, et al. 2018a. Emotionlines: An emotion corpus of multi-party conversations. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018, Miyazaki, Japan, May 7-12, 2018.*

Yuxiao Chen, Jianbo Yuan, Quanzeng You, and Jiebo Luo. 2018b. Twitter sentiment analysis via bi-sense emoji embedding and attention-based lstm. In *Proceedings of the 26th ACM international conference on Multimedia*, pages 117–125.

Jason PC Chiu and Eric Nichols. 2016. Named entity recognition with bidirectional lstm-cnns. *Transactions of the Association for Computational Linguistics*, 4:357–370.

Roddy Cowie, Ellen Douglas-Cowie, Nicolas Tsapatsoulis, George Votsis, Stefanos Kollias, Winfried Fellenz, and John G Taylor. 2001. Emotion recognition in human-computer interaction. *IEEE Signal processing magazine*, 18(1):32–80.

Ben Eisner, Tim Rocktäschel, Isabelle Augenstein, Matko Bošnjak, and Sebastian Riedel. 2016. emoji2vec: Learning emoji representations from their description. *arXiv preprint arXiv:1609.08359*.

Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. 2017. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. *arXiv preprint arXiv:1708.00524*.

Joseph L Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378.

Wei Gao, Shoushan Li, Sophia Yat Mei Lee, Guodong Zhou, and Chu-Ren Huang. 2013. Joint learning on sentiment and emotion classification. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 1505–1508.

Gaël Guibon, Magalie Ochs, and Patrice Bellot. 2018. Emoji recommendation in private instant messages. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, pages 1821–1823.

Shirley Anugrah Hayati and Aldrian Obaja Muis. 2019. Analyzing incorporation of emotion in emoji prediction. In *Proceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 91–99.

Wegdan Hussien, Mahmoud Al-Ayyoub, Yahya Tashtoush, and Mohammed Al-Kabi. 2019. On the use of emojis to train emotion classifiers. *arXiv preprint arXiv:1902.08906*.

Shuning Jin and Ted Pedersen. 2018. Duluth urop at semeval-2018 task 2: Multilingual emoji prediction with ensemble learning and oversampling. *arXiv preprint arXiv:1805.10267*.

Alex Kendall, Yarin Gal, and Roberto Cipolla. 2018. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7482–7491.

Taku Kudo. 2018. Subword regularization: Improving neural network translation models with multiple subword candidates. *arXiv preprint arXiv:1804.10959*.

Taku Kudo and John Richardson. 2018. Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. *arXiv preprint arXiv:1808.06226*.

Wang Ling, Tiago Luís, Luís Marujo, Ramón Fernandez Astudillo, Silvio Amir, Chris Dyer, Alan W Black, and Isabel Trancoso. 2015. Finding function in form: Compositional character models for open vocabulary word representation. *arXiv preprint arXiv:1508.02096*.

Weicheng Ma, Ruibo Liu, Lili Wang, and Soroush Vosoughi. 2020. Emoji prediction: Extensions and benchmarking. *arXiv preprint arXiv:2007.07389*.

Parth Patwa, Gustavo Aguilar, Sudipta Kar, Suraj Pandey, Srinivas PYKL, Björn Gambäck, Tanmoy Chakraborty, Thamar Solorio, and Amitava Das. 2020. Semeval-2020 task 9: Overview of sentiment analysis of code-mixed tweets. *arXiv e-prints*, pages arXiv–2008.

Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*.

Adithya Pratapa, Monojit Choudhury, and Sunayana Sitaram. 2018. Word embeddings for code-mixed language processing. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pages 3067–3072.

Saurav Sahay, Shachi H Kumar, Rui Xia, Jonathan Huang, and Lama Nachman. 2018. Multimodal relational tensor network for sentiment and emotion classification. *arXiv preprint arXiv:1806.02923*.

Sashank Santhanam, Vidhushini Srinivasan, Shaina Glass, and Samira Shaikh. 2019. I stand with you: Using emojis to study solidarity in crisis events. *arXiv preprint arXiv:1907.08326*.

M. Schuster and K.K. Paliwal. 1997. Bidirectional Recurrent Neural Networks. *Trans. Sig. Proc.*, 45(11):2673–2681.

Rico Sennrich, Barry Haddow, and Alexandra Birch. 2015. Neural machine translation of rare words with subword units. *arXiv preprint arXiv:1508.07909*.

Abu Awal Md Shoeb and Gerard de Melo. 2020. Can emojis convey human emotions? a study to understand the association between emojis and emotions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8957–8967.

Toshiki Tomihira, Atsushi Otsuka, Akihiro Yamashita, and Tetsuji Satoh. 2020. Multilingual emoji prediction using bert for sentiment analysis. *International Journal of Web Information Systems*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Zhenduo Wang and Ted Pedersen. 2018. Umdsub at semeval-2018 task 2: Multilingual emoji prediction multi-channel convolutional neural network on subword embedding. *arXiv preprint arXiv:1805.10274*.

Zhongqing Wang, Sophia Yat Mei Lee, Shoushan Li, and Guodong Zhou. 2016a. Emotion analysis in code-switching text with joint factor graph model. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(3):469–480.

Zhongqing Wang, Yue Zhang, Sophia Lee, Shoushan Li, and Guodong Zhou. 2016b. A bilingual attention network for code-switched emotion prediction. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 1624–1634.

Sanjaya Wijeratne, Lakshika Balasuriya, Amit Sheth, and Derek Doran. 2017. Emojinet: An open service and api for emoji sense discovery. In *Eleventh International AAAI Conference on Web and Social Media*.

Chuhan Wu, Fangzhao Wu, Sixing Wu, Zhigang Yuan, Junxin Liu, and Yongfeng Huang. 2018. Thu_ngn at semeval-2018 task 2: Residual cnn-lstm network with attention for english emoji prediction. In *Proceedings of The 12th International Workshop on Semantic Evaluation*, pages 410–414.

Siddharth Yadav and Tanmoy Chakraborty. 2020. Unsupervised sentiment analysis for code-mixed data. *arXiv preprint arXiv:2001.11384*.

Sung-Lin Yeh, Yun-Shao Lin, and Chi-Chun Lee. 2019. An interaction-aware attention network for speech emotion recognition in spoken dialogs. In *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2019, Brighton, United Kingdom, May 12-17, 2019*, pages 6685–6689. IEEE.

Jianfei Yu, Luis Marujo, Jing Jiang, Pradeep Karuturi, and William Brendel. 2018. Improving multi-label emotion classification via sentiment classification with dual attention transfer network. ACL.

Xianda Zhou and William Yang Wang. 2017. Mojitalk: Generating emotional responses at scale. *arXiv preprint arXiv:1711.04090*.