Intensity-Valued Emotions Help Stance Detection of Climate Change Twitter Data

Apoorva Upadhyaya , Marco Fisichella , Wolfgang Nejdl
L3S Research Center, Leibniz Universität Hannover, Germany
{upadhyaya, mfisichella, nejdl}@l3s.de

Abstract

Our study focuses on the United Nations Sustainable Development Goal 13: Climate Action, by identifying public attitudes on Twitter about climate change. Public consent and participation is the key factor in dealing with climate crises. However, discussions about climate change on Twitter are often influenced by the polarised beliefs that shape the discourse and divide it into communities of climate change deniers and believers. In our work, we propose a framework that helps identify different attitudes in tweets about climate change (deny, believe, ambiguous). Previous literature often lacks an efficient architecture or ignores the characteristics of climate-denier tweets. Moreover, the presence of various emotions with different levels of intensity turns out to be relevant for shaping discussions on climate change. Therefore, our paper utilizes emotion recognition and emotion intensity prediction as auxiliary tasks for our main task of stance detection. Our framework injects the words affecting the emotions embedded in the tweet to capture the overall representation of the attitude in terms of the emotions associated with it. The final task-specific and shared feature representations are fused with efficient embedding and attention techniques to detect the correct attitude of the tweet. Extensive experiments on our novel curated dataset, two publicly available climate change datasets (ClimateICWSM-2023 and ClimateStance-2022), and a benchmark dataset for stance detection (SemEval-2016) validate the effectiveness of our approach.

1 Introduction

Changing climate has catastrophic effects on all ecosystems in the world, including humans. As climate change continues to worsen, social media platforms such as Twitter plays a critical role in creating awareness among the public. Despite the scientific evidence of the current climate crisis [Pörtner *et al.*, 2022], the public remains skeptical, leading to polarized discussions about climate change on Twitter, which often result in misinformation, fake news, and bias that influence public attitudes towards climate change [Jang and Hart, 2015; Zhou and Shen, 2021]. Recent articles[1,2] published on the Verge news (an American technology news website) and the Euro news (European television news network) also claim that there has been a sharp increase in the number of tweets and retweets referring to skeptical content towards climate change and these climate deniers frequently abuse, use hate speech and promote violent behavior towards climate activists and scientists. It is, therefore, needful for government organizations, technology companies, and concerned authorities to recognize and interfere in order to stop such climate change deniers' spread of content that can lead to violent activities, disbelief in government policies, and mitigation of climate change. Since stance detection helps to understand the viewpoint of the author whether it is in favor or against the target topic, we perform stance detection to help identify climate change denier and believer tweets.

Existing literature on climate change has largely focused on examining the polarising effects of climate change conversations [Jang and Hart, 2015; Tyagi *et al.*, 2020b; Falkenberg *et al.*, 2022], while some of the more recent climate-specific studies have conducted stance identification. [Vaid *et al.*, 2022] developed the dataset and proposes a BERT embedding architecture to detect the stance of tweets. Other models use basic model components and lack efficiency and more advanced architecture [Chen *et al.*, 2019; Kabaghe and Qin, 2020]. [Upadhyaya *et al.*, 2022a] used sentiment analysis for the stance task but suffered from the drawback of not properly detecting similar sentiments in tweets from supporters and deniers. Moreover, several previous works performed stance detection on the SemEval 2016 dataset including 364 climate-related tweets [Reveilhac and Schneider, 2023; Fu *et al.*, 2022; Wang and Wang, 2021], but this limited number of tweets did not help the models to focus on the climate change domain, which is currently one of the biggest crises facing humanity. Therefore, these works motivated us to develop an efficient model that can detect the climate change attitude of tweets by incorporating advanced compo-

[1]https://www.theverge.com/2022/12/5/23494220/elon-musk-twitter-climate-misinformation-rise-analysis

[2]https://www.euronews.com/green/2022/11/17/how-climate-disinformation-is-spreading-after-elon-musks-twitter-takeover

nents and other auxiliary tasks to avoid the drawbacks of previous works.

Emotions and their intensity values have helped various domains [Upadhyaya and Chandra, 2022; Abro *et al.*, 2022]. Moreover, emotional and sentiment aspects in climate change conversations often shape discussions and influence public opinion toward the climate crisis [Salama and Aboukoura, 2018; Upadhyaya *et al.*, 2023b]. Therefore, we focus on emotion recognition and emotion intensity prediction as auxiliary tasks to identify the different attitudes of tweets based on the presence of different emotions. Even in the case of similar emotions, the intensity of those emotions, whether high or low, can further help to identify the attitude. The following are examples of tweets on climate change with their emotion and intensity values: ***(i.)Deny:*** "Another fucking hypocrite. There is no climate crisis just the biggest scam in history...#climatehoax *(Anger:0.58, Disgust:0.67, Sadness:0.31)*"; "oh fuck off with your climate bullshit..*(Anger:0.82, Disgust:0.71)*". ***(ii.)Believe:*** "This is what you get when you #votegreen–hope for a better world...*(Anticipation:0.51, Trust:0.73)*";"How dare you continue to look away #ClimateEmergency #ClimateAction.." *(Anger:0.32, Disgust:0.21, Surprise:0.63)*". The deniers' tweets show higher levels of anger and disgust, while expectation and trust are more prevalent in the believers' tweets. Even though similar negative feelings of anger and disgust are present in the believers' tweets, they have a softer(more gentle) tone with lower intensity scores (refer Section 3). This relationship between different emotions or similar emotions with different levels of intensity motivated us to use these tasks to support the detection of attitudes.

The main contributions of our work are as follows:***(i.)*** We create a new climate change dataset consisting of tweets with annotations of stance, emotions, and intensity scores *(code and dataset are available here[3])*. ***(ii.)*** To the best of our knowledge, this is the first cross-sectional study to use emotions and their intensity scores to identify the attitude of the tweet. ***(iii.)*** We propose a multi-task system SIMS (IntenSIty-Valued EMotions Help Stance Detection of Climate Change Twitter Data) for stance detection by utilizing emotion recognition and emotion intensity prediction as auxiliary tasks. Our SIMS extracts the words from the tweet text that affect the emotions present in the tweets by using the affect word extractor component. The Emotion Affect Inducer is then responsible for injecting the emotional aspects into the tweet text to capture the overall attitude of the tweet in relation to the associated emotions. The final task-specific and shared representations of the input feature are then fused using the integration module of our proposed approach to determine the correct attitude label of the input tweet. ***(iv.)*** Extensive experiments are conducted on our curated dataset, two publicly available climate change datasets (ClimateICWSM-2023 and ClimateStance-2022), and benchmark stance detection dataset (SemEval-2016). The experimental results show that our proposed SIMS outperforms other computational methods by benefiting from auxiliary tasks and the proposed model architecture. *Please note that we refer to the task of stance detection as SD, emotion recognition as ER, and emotion intensity prediction as EI for the rest of the study.*

[3]https://github.com/apoorva-upadhyaya/Emotion_Intensity_Stance

2 Related Works

Climate Change and Stance Detection (SD). As climate change continues to worsen, social media plays a crucial role in spreading awareness [Jang and Hart, 2015]. Recently, [Upadhyaya *et al.*, 2022b] explores the information-seeking behavior of students about climate change on YouTube. However, discussions on Twitter often get polarized by users' beliefs about climate change, dividing them into two communities of deniers and believers of climate change [Jang and Hart, 2015]. Therefore, we focus on the task of automatically identifying the viewpoint of the tweet towards a target to help identify denier statements as climate deniers often spread misinformation and fake news, leading to issues of climate delay and opposing climate action [Zhou and Shen, 2021]. Earlier work on climate change has largely focused on examining the impact of polarised beliefs on Twitter [Tyagi *et al.*, 2020b], while others identify polarised users based on statements [Tyagi *et al.*, 2020a]. However, some of the more recent works have focused on identifying attitudes from statements toward climate change. [Vaid *et al.*, 2022] introduced the ClimateStance-2022 dataset and offered BERT architectures for stance task. [Upadhyaya *et al.*, 2022a] used sentiment analysis as an auxiliary task for detecting stances in their curated data (ClimateICWSM-2023), but suffered from the drawback of sarcasm and the presence of similar words in tweets from deniers and believers. Moreover, [Kabaghe and Qin, 2020] uses naive bayes and does not have an advanced architecture. Hence, these works motivated us to develop an efficient model that uses BERTweet embeddings and attention, and combines the emotional aspects and their intensity values to better distinguish tweets with different attitudes, even if they have similar sentiments or words. There are several studies that use the SemEval-2016 dataset to detect stance [Reveilhac and Schneider, 2023; Fu *et al.*, 2022; Wang and Wang, 2021]. However, the dataset contains only 29 denier and 335 believer tweets, hence, these works do not focus on understanding the effective characteristics of denier or supporter tweets and also ignore the intensity of similar emotions/feelings in the tweet, which could help in identifying the correct attitude of the tweet despite the presence of similar emotions. Therefore, we propose an approach that uses advanced architecture and combines the emotion and their intensity values to perform the SD task.

Emotion Recognition (ER) and Intensity Prediction (EI). Several climate-specific studies have examined the importance of feelings and emotional aspects in the climate change conversations on social media [Brosch, 2021; Upadhyaya *et al.*, 2023a], proving the importance of the sentiments embedded in tweets for determining the tweeter's attitude towards climate change. In addition, the intensity of emotions has been used in different domains [Zirikly *et al.*, 2019; Abro *et al.*, 2022] and helps us understand how similar emotions can be severe or gentle based on their intensity values, which can also be useful to distinguish between different at-

Category	**Anger**	**Anticipation**	**Disgust**	**Fear**
Deny	**50.05**	12.8	**81.98**	19.81
Believe	25.3	45.5	17.19	**55.98**
Ambiguous	26.3	**46.97**	18.78	49.79
Category	**Joy**	**Sadness**	**Surprise**	**Trust**
Deny	7.41	**39.86**	28.73	8.05
Believe	24.09	36.98	**33.01**	**59.23**
Ambiguous	**39.62**	24.75	27.02	58.03

Table 1: % of emotions present in tweets

titudes toward climate change (refer Section 3 of our study). Hence, these studies have inspired us to explore how emotions and their intensity can further support the stance task.

3 Dataset

We initially use climate change denier and believer query hashtags to collect real-time Twitter data, similar to the existing literature [Tyagi *et al.*, 2020a; Upadhyaya *et al.*, 2022a]. We collected $31,546$ denier and $82,010$ believer English language tweets from 25 July 2021 to 05 Dec 2022 using query hashtags and Tweepy API[4] (after deduplicating the tweets based on tweet text). However, it is suggested that the stance of a tweet may change after removing the query hashtag [Sobhani *et al.*, 2016]. Therefore, we randomly select $10,000$ tweets from collected data and perform manual annotation.

3.1 Data Annotation

Stance Detection (SD). We first remove the query hashtags from the randomly selected set of $10,000$ tweets. The three trained annotators perform the task of manual annotation. In line with the existing literature [Sobhani *et al.*, 2016; Vaid *et al.*, 2022], the annotators tagged viewpoints on climate change for 3 categories: ***(i.)Believe(Favour):*** contain expressions that agree and suggest that climate change is real and happening; ***(ii.)Deny(Against):*** consist of opinions against climate action, climate change, and government policies; ***(iii.)Ambiguous:*** tweets do not contain a clear expression/stance toward climate change. To determine the quality of the annotations, we calculated the agreement between the annotators, as evidenced by the Fleiss-Kappa score [Spitzer *et al.*, 1967] of 0.80, indicating that the annotation and the dataset presented are of considerable quality. We found a total of **5362** believe, **1726** deny, and **2912** ambiguous tweets.

Emotion Recognition (ER) & Intensity Prediction (EI). Previous works have used the NRC Emotion Intensity Lexicon (NRC-EIL) [Mohammad, 2018], which indicates the presence of 8 basic emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust) with their real-valued intensity values [Shoeb and de Melo, 2020; Qamar *et al.*, 2021; Qin and Ronchieri, 2022]. We also use the python library of NRC-EIL[5] to compute the labels for ER and EI tasks. The preprocessed tweet text is fed to the weighted_emotion_scores() function of the NRC-EIL library, which returns a dictionary of emotions and their intensity scores for each tweet (for example, output dictionary: {'anger': 0.698, 'disgust': 0.54, 'fear': 0.31}). For the task ER, we create a list of length 8 and mark all the emotions present as 1, while others that are not predicted by the library are marked as 0 (ER label=[1,0,1,1,0,0,0,0]). Similarly, for the EI task, we create a list of length 8 with values as intensity scores for the corresponding emotions provided by the library and mark the rest as 0 (EI label=[0.698,0,0.54,0.31,0,0,0,0]). To assess the quality of the labels predicted by the NRC-EIL library, three trained annotators manually annotated 1000 randomly selected tweets from our dataset for both ER and EI tasks. It should be noted that the annotators indicate the intensity labels as "high", "moderate" and "low" for their ease of labeling. After careful analysis, we convert the values according to the criteria if the intensity > 0.8 is classified as high, < 0.3 as low and the rest as moderate intensity. However, we keep the intensity values as real values and consider EI as a multiple-output regression task for our study, which corresponds to the real environment. We found a Fleiss-Kappa [Spitzer *et al.*, 1967] score of 0.79 and 0.76 between the semi-supervised and our manual annotations for the ER and EI tasks respectively, indicating that the predicted labels are of considerable quality. The annotations provided by NRC-EIL are therefore taken into account for ER and EI tasks in order to save time and costs.

Data Pre-processing. We first remove the query hashtags from the tweets, as mentioned in Section 3.1, and then remove URLs, punctuation marks, and stopwords. All the text of the tweets is converted to lowercase. We then use NLTK-based[6] TweetTokenizer to tokenize the tweets, followed by NLTK Wordnet Lemmatizer to combine inflected words into root form, and PorterStemmer for stemming.

The percentage of each emotion found in the tweets is shown in Table 1. We report the percentage of tweets of different stances having high/moderate/low-intensity scores of corresponding emotions in Table 2 of Supplementary[3].

4 Proposed Methodology

Problem Statement. *Propose a stance detection approach that combines the tweet text with the emotions-inducing words and utilizes these embedded emotions and their corresponding intensities to further classify the attitude of a climate change tweet into one of the polarized classes (ambiguous/believe/deny).* Our SIMS approach consists of the following components: *Embedding Component, Feature Encoder, Attention, and Classification Layer* (as shown in Figure 1). The input tweet text is initially passed through an affect word extractor that extracts the words influencing the emotions in the tweet. The affect words along with the tweet text are passed through the embedding component containing an emotional affect inducer to generate efficient embeddings to capture the overall representation of the tweet with respect to the emotions present in the tweet. The emotionally affected tweet features are then passed through three separate feature encoders to encode the final input representation specific to each task. The attention component then integrates

[4]http://docs.tweepy.org/en/latest/streaming_how_to.html

[5]https://pypi.org/project/emotion-nrc-affect-lex/

[6]https://www.nltk.org/index.html

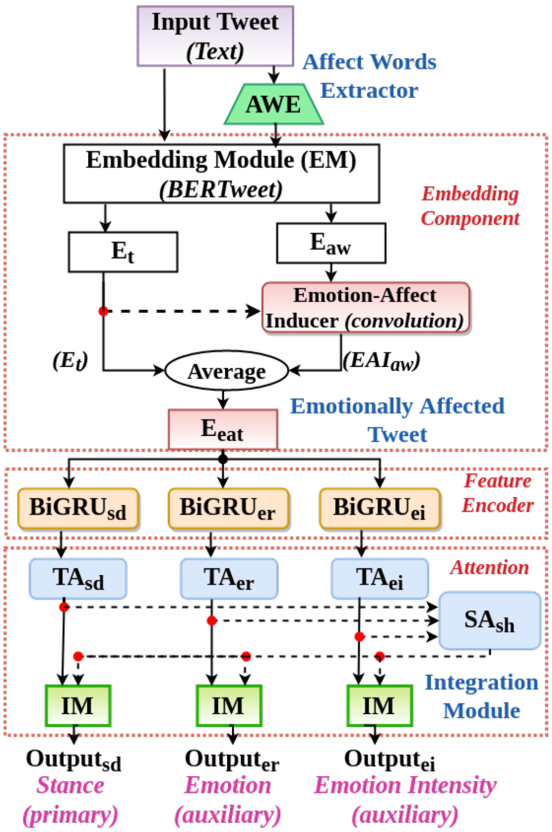

Figure 1: Architectural overview of our proposed SIMS approach

task-specific and shared attention, followed by the classification layer to obtain the output label for each task. We now describe each component in detail.

4.1 Embedding Component

The input tweet text is first passed through an affect word extractor (AWE) to formulate a set of words in the tweet that affect the emotional presence of the tweet. The AWE consists of the function affect_dict() of the NRC-EIL library[5] which returns the keys of the dictionary as the affect words that influence the presence of emotions in the tweet. The embedding component consists of two modules (Embedding Module and Emotion-Affect Inducer):

Embedding Module. The tweet text and affect words are fed to the embedding module (EM), which consists of the pre-trained model BERTweet [Nguyen *et al.*, 2020]. BERTweet has a similar architecture to $BERT_{base}$ but is explicitly trained on tweets, which allows for greater efficiency in identifying the semantics and syntax of words in the text and set of affect words. The tweet text and affect words containing n_t and n_{aw} number of words respectively, where the embedding of each word is fetched from BERTweet with dimension d_e, are then flattened to yield $E_t \in R^{n_t(d_e)}$ and $E_{aw} \in R^{n_{aw}(d_e)}$, respectively (refer Figure 1). The embedded E_t and E_{aw} vectors are then passed through a dense layer of dimension d_f, resulting in $E_t \in R^{d_f}$ and $E_{aw} \in R^{d_f}$ representations for text and affect words.

Emotion-Affect Inducer. This module is responsible for incorporating the effect of emotion into the tweet in order to capture the overall representation of the tweet in terms of the emotions involved. In conjunction with the existing literature demonstrating the importance of discrete linear convolution for modeling the effect of one function on another [Wikipedia, 2001; Bahri *et al.*, 2013; Wu *et al.*, 2021], we also use the numpy operator convolve[7] to obtain the convolution between the embedded text vector (E_t) and the embedded affect words vector (E_{aw}). The equation 1 represents the convolution operation showing the effect of emotional aspects (affect words vector) onto the text representation, where, n is the dimension of $E_{aw}(R^{d_f})$ and m is the dimension of $E_t(R^{d_f})$, resulting in the output of Emotion-Affect Inducer ($EAI_{aw} \in R^{d_f}$) that consists of the tweet text representation with embedded emotional aspects.

$$(E_{aw} * E_t)_n = \sum_{m=-\infty}^{\infty} E_{aw}[m] E_t[n-m] \Rightarrow EAI_{aw} \quad (1)$$

$$E_{eat} = Average(E_t, EAI_{aw}) \quad (2)$$

Finally, we average the tweet text embedding vector (E_t) and the output of the emotion-affect inducer (EAI_{aw}) so as not to miss any stance-specific features of the tweet while retaining the emotional aspects adjoined in the tweet (see equation 2), resulting in the final output of the embedding component, i.e. an emotionally affected representation of the tweet (as in Figure 1), which is then reshaped and passed to the feature encoders ($E_{eat} \in R^{d_f \times 1}$).

4.2 Feature Encoder

The emotionally affected tweet (E_{eat}) obtained from the embedding component is then passed to the three discrete Bi-GRU layers with dimension d_g specific to each task so as to incorporate past and future context information and sequentially encode these long-term semantic dependencies into hidden states ($H \in R^{2d_g \times 1}$). Hence, the feature encoder results in 3 outputs, shown by $BiGRU_{sd}$, $BiGRU_{er}$, and $BiGRU_{ei}$ for SD, ER, and EI tasks respectively (Figure 1).

4.3 Attention

The module initially extracts the task-specific attention and then fuses the shared attention vector with each of the task-specific attention vectors using the *integration module*. In line with the existing literature [Vaswani *et al.*, 2017], in order to extract the important and relevant parts of the input tweet representation, we first create a set of queries, keys, and values by passing the output of the feature encoder to three separate dense layers of dimension (d_a), which results in $Q_{task} \in R^{d_a \times 1}$, $K_{task} \in R^{d_a \times 1}$, and $V_{task} \in R^{d_a \times 1}$, for each task. The task attention vector (TA_{task}) specific to each task is then calculated using the equation 3, where $TA_{task} \in R^{d_a \times 1}$.

$$TA_{task} = softmax(Q_{task} K_{task}^T) V_{task} \quad (3)$$

Here, the three TA_{sd}, TA_{er}, and TA_{ei} vectors are formulated for SD, ER, and EI tasks respectively (shown in Figure 1). In

[7]https://numpy.org/doc/stable/reference/generated/numpy.convolve.html

Features	Precision	Recall	F1-score	Acc.
BERT (Text)	79.42/1.1	75.28/0.6	76.59/1.0	78.34/1.1
BERTweet (Text)	81.55/0.6	77.52/0.3	78.78/1.1	80.19/0.8
BERTweet (Text+AWE) (EM)	83.09/2.0	79.99/1.5	81.87/1.3	82.11/1.5
EM+EAI (emo.affect)	83.31/0.8	81.20/0.6	82.65/0.8	82.37/1.0
EM+EAI+ TA (Attn.)	84.56/1.2	81.79/1.1	**83.02/0.7**	**84.29/1.2**
EM+EAI+ TA+emo & int. i/p	86.24/0.7	83.59/0.8	**85.19/1.0**	**86.47/0.9**

Table 2: Results (Avg./Std.dev.) of the single task stance detection in various combinations

addition, to take advantage of the shared attention features and to use the features that are common to all tasks, we averaged the task-specific attention vectors to obtain the shared attention vector, where SA_{sh} = Average (TA_{sd}, TA_{er},TA_{ei}). The SA_{sh} together with the task-specific attention vector is then fed to Integration Module.

Integration Module. integrates the TA_{task} and SA_{sh} vectors by using the fusion technique of absolute difference and element-wise product that proves to be effective in various previous works [Mou *et al.*, 2015]. The final output of the attention component (IM_{op}) is fetched using equation 4, is then flattened, and passed to the classification layer to provide outputs for each task.

$$\text{IM}_{op} = [TA_{task}; SA_{sh}; TA_{task} - SA_{sh}; TA_{task} \odot SA_{sh}] \quad (4)$$

4.4 Classification Layer

The final representation of the tweet (IM_{op}) of the attention framework is then passed through three output separate channels for stance ($Output_{sd}$), emotion ($Output_{er}$), and emotion intensity ($Output_{ei}$) tasks separately. We compute categorical cross-entropy loss (L_c^{sd}) for SD, binary cross-entropy loss for ER (L_b^{er}), and mean squared error loss for the EI task (L_{mse}^{ei}). The integrated loss function (L) of our SIMS framework is realized in equation 5:

$$L = x * L_c^{sd} + y * L_b^{er} + z * L_{mse}^{ei} \quad (5)$$

where x, y, and z represent the constants between 0 and 1 indicating the per-task loss-share to the overall loss.

5 Experiments

5.1 Dataset

Experiments are conducted on the following datasets: ***(i.)*** Our curated dataset is described in Section 3 in detail; ***(ii.)ClimateICWSM-2023 [Upadhyaya* et al.*, 2022a]:*** consists of climate change tweets of believe $(60, 430)$ and deny $(13, 125)$ stances that are collected using hashtags; ***(iii.)ClimateStance-2022 [Vaid* et al.*, 2022]:*** In this benchmark dataset, $3, 777$ climate change tweets are included that support, oppose, and have ambiguous views regarding climate change prevention. ***(iv.)SemEval-2016 [Mohammad* et al.*, 2016]:*** is a benchmark stance detection dataset used in SemEval-2016 shared task 6.A that covers Atheism, Climate Change is a Real Concern, Feminist Movement, Hillary Clinton, and Abortion as targets with tweets having favor, against, or neutral stances. *The distribution of emotion and intensity scores in all* 3 *publicly available datasets can be found here*[3].

5.2 Implementation Details

We perform stratified 5-fold cross-validation and report accuracy, macro precision, macro recall, and macro F1 scores. Using Sklearn resampling, we oversample the minority classes in training data. We execute experiments on NVIDIA GeForce GTX 1080Ti GPU servers (TDP of 250W) with the carbon efficiency of 0.38 kgCO_2eq/kWh. Carbon footprint is calculated using the Machine Learning Impact calculator [Lacoste *et al.*, 2019]. A cumulative of 10 hours of computation was performed on the hardware, including single and multi-task models training and evaluation on all datasets used in our work, resulting in total emissions to be $\approx$ 0.95 kgCO_2eq . However, the SIMS model required 50 minutes to be trained on our curated dataset, resulting in $\approx$ 0.11 kg CO_2eq. emission. The best parameters for experiments are: Embeddings dimension for *BERT(d_e)*: 768, *Bi-GRU memory cells (d_g)*: 128, *fully connected layer dimension of embedding module (d_f) and attention component (d_a) [with ReLu activation]*: 128, *output neurons/channels*: 3 [softmax activation] (SD), 8 [sigmoid activation] (ER) and 8 [ReLU activation] (EI), *loss*: categorical cross-entropy (L_c^{sd}) for SD, binary cross-entropy loss function for ER (L_b^{er}), and mean squared error (MSE) for EI (L_{mse}^{ei}) tasks; *optimizer*: Adam(0.001). The best parameter values are selected using TPE in the Hyperopt[8] python library that minimizes loss functions. Furthermore, we fine-tune the loss weights for all tasks by using Grid Search from Scikit-learn (SD *(x)*=1, ER *(y)*=0.5, and EI *(z)*=0.3).

5.3 Baselines

We compare our SIMS with the below baselines on our curated dataset. ***Semi-Supervised(Model3) [Reveilhac and Schneider, 2023]***: performs stance detection using the semi-supervised approach with stance, linguistic, entity, and other features on SemEval dataset. ***RoBERTa-Base [Vaid* et al.*, 2022]***: a stance detection framework for climate change tweets (ClimateStance-2022). ***MT-LRM-BERT [Fu* et al.*, 2022]***: Use SemEVAL-2016 and other benchmark datasets to detect stances with a multi-task approach that considers sentiment and opinion as additional tasks. ***SP-MT [Upadhyaya* et al.*, 2022a]***: a novel multi-task framework that performs stance detection with the help of sentiment analysis on the ClimateICWSM-2023 dataset. ***ESD [Vychegzhanin and Kotelnikov, 2021]***: is an optimal ensemble of classifiers and feature set to detect stance using SemEval and other datasets. ***S-MDMT [Wang and Wang, 2021]***: a multi-task multi-domain framework to perform stance detection using SemEVAL-2016 dataset. ***HAN [Wang* et al.*, 2020]***: a hierarchical attention neural model proposed for stance detection. ***MNB [Kabaghe and Qin, 2020]***: Tweets about climate change are classified by multinomial naive bayes into

[8]http://hyperopt.github.io/hyperopt/

Model	Stance + Emotion(SD+ER)			Stance + Emo. Intensity(SD+EI)			Stance + Emotion +Emo. Intensity (SD+ER+EI)		
	Precision	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
EM	86.18/0.75	83.50/0.81	85.03/1.16	85.66/0.25	81.72/0.81	83.89/1.20	87.55/1.07	84.61/2.02	86.18/1.71
EM+EAI	89.09/1.41	86.33/1.05	87.58/0.88	86.62/0.92	84.45/0.71	85.48/1.04	89.49/0.89	87.33/1.05	88.71/0.85
EM+EAI +Concat (TA+SA)	89.60/0.73	88.46/1.17	89.28/1.11	88.01/1.32	86.27/0.56	87.69/1.53	91.35/0.47	89.05/0.59	90.06/0.80
EM+EAI +Integrate (TA+SA)	90.55/1.22	88.78/0.90	90.01/1.39	89.74/0.66	88.59/0.42	89.31/0.82	**92.08/1.01**	**90.36/0.62**	**91.84/0.71 (SIMS)**

Table 3: Results (Avg/St.dev) of Multi-task architectures for stance detection on our climate dataset. SIMS outperforms other variants while meeting statistical significance under t-tests (p <0.05).

Model	Precision	Recall	F1 score
-	**Avg/St.dev**	**Avg/St.dev**	**Avg/St.dev**
SIMS [Proposed]	**92.08/1.01**	**90.36/0.62**	**91.84/0.71**
Semi-Super.	85.06/1.12	83.25/1.24	84.66/1.19
RoBERTa-Base	83.38/1.55	85.24/1.28	84.69/1.89
MT-LRM-BERT	87.12/1.61	88.70/0.99	88.59/1.29
SP-MT	87.95/1.11	90.01/1.80	89.29/1.31
ESD	81.55/1.72	84.39/2.05	83.28/2.31
S-MDMT	86.12/1.02	88.67/0.39	86.91/0.44
HAN	84.61/1.22	84.23/1.78	84.54/1.65
MNB	78.11/0.66	79.51/0.73	78.43/1.33
AT-JSS-LEX	88.66/0.34	87.51/0.69	87.15/0.35
DNN	77.64/1.58	76.38/1.08	77.15/1.18

Table 4: Results of SIMS with baselines on our climate dataset. SIMS outperforms all baselines while meeting statistical significance under t-tests (p <0.05).

Model	Accuracy	F1-score
SIMS [Proposed]	**95.01**	**93.52**
SP-MT	93.95	90.24
LR	81.48	81.00
ESD	89.65	85.11
HAN	89.47	86.00
AT-JSS-LEX	88.02	84.01
MNB	85.44	78.08
DNN	84.61	76.23
SVM-ngram	85.55	66.33

Table 5: Results of our proposed approach on publicly available ClimateICWSM-2023 dataset.

positive, negative, and neutral beliefs. ***AT-JSS-LEX [Li and Caragea, 2019]***: a multi-task framework for stance detection by using sentiment lexicon loss on SemEval dataset. ***DNN [Chen* et al.*, 2019]***: a neural network that classifies users as climate change deniers/believers on Twitter.

6 Results

We report the results with respect to our main task i.e. stance detection (SD) as the current work aims to improve the performance of the main task using ER and EI as auxiliary tasks.

Significance of SIMS Components. Tables 2 and 3 show the performance improvement of the SD task when the proposed components are used. As can be seen in Table 2, using the BERTweet embedding in the single-task stance detection results in a 2.86% increase in the F1 score, as BERTweet trained on tweets only helps to detect the tweet text features more efficiently than BERT_{base} embeddings. In the single-task model variant (Table 2), adding affect words and capturing their emotional aspects in the tweet improves performance by 4.91% in the F1 score, confirming the importance of the Emotion-Affect Inducer (EAI) model component. Furthermore, the introduction of task attention (TA) leads to an average F1 score of 83.02 for the single-task model. Similarly, in our multi-task frameworks (see Table 3), the addition of the Emotion-Affect Inducer (EAI) along with the task (TA) and shared attention vectors (SA) improves model performance in all variants (SD + ER, SD + EI, SD + ER+ EI). Moreover, the use of integrating the task and shared attention vectors over the concatenation operation enhances the performance of SD + ER + EI model variant combination with the F1 value from 88.71 to 91.84, resulting in 3.53% (row 2 & 4 of Table 3) increase rather than 1.52% increase in the F1 when using the concatenate operation (row 2 & 3 of Table 3), thus demonstrating the importance of the fusion mechanism of the integration module (IM).

Effectiveness of Auxiliary Tasks. The performance improvement of the single-task variant framework when using emotions and their intensity values as input features (with 85.19 F1 score in Table 2) motivates us to analyze the effect of ER and EI when used as auxiliary tasks. Table 3 shows that the combination of SD + ER performs slightly better than SD + EI with an average F1 score of 90.01 and 89.31 respectively, indicating that ER task has contributed more than EI for capturing stance features. The better and clearer separation of emotions between the believers' and deniers' tweets contributed to a more efficient SD task and justifies the assignment of higher loss weights to the ER task (refer Table 1). However, the combination of the ER and EI tasks further improves the performance of the SD task, resulting in an accuracy of 93.72, as the presence of similar emotions can be distinguished by their intensity values (see Table 2 of Supplementary), proving the importance of ER and EI for SD.

Comparisons With Baselines. ***(i.)Our curated climate dataset:*** Our SIMS overtakes the other baselines when these methods are trained and tested on our curated dataset, confirming the effectiveness of our approach (refer Table 4 for the results). SIMS outperforms the semi-supervised approach (Model3) and ROBERTa-Base shows that the multi-task setting with the better embedding technique of BERTweet can improve the stance task. The methods SP-MT, MT-LRM-

BERT, and AT-JSS-LEX perform better than the other models because these approaches use sentiment information to detect the attitude of the tweet. However, SIMS outperforms these methods because separating the sentiments into multiple emotions at the more granular level and further distinguishing the presence of similar emotions in different stances according to their intensity values makes our approach more efficient for the SD task. S-MDMT uses target classification as a separate task, while ESD and HAN use different hierarchical and attentional features. However, our SIMS performs better by using the EI and ER as auxiliary tasks and incorporating the emotional aspects into the tweet representation. Our single-task method surpasses the MNB and DNN methods, which shows the importance of an advanced architecture to improve the task SD. During the writing of this paper, we were unaware of any other work which utilized emotion task for identifying the stance of climate change tweets. ***(ii.)ClimateICWSM-2023:*** We use the baselines from the work of [Upadhyaya *et al.*, 2022a], who created this dataset. Based on Table 5, our approach SIMS surpasses other methods, suggesting that the fine-grained separation of positive, negative, and neutral sentiments into 8 categories of emotions can better identify the attitude of the tweet, while their emotion intensity scores and inducing the words influencing these emotions further help to distinguish the presence of similar emotions but with different intensity levels and context among the believe, deny and ambiguous stances. ***(iii.)ClimateStance-2022:*** We use the baseline methods from the [Vaid *et al.*, 2022] work that created the dataset. SIMS performs significantly better than the other models (refer Table 6), proving the importance of capturing emotional aspects and that the use of BERTweet embeddings together with the fusion of task and shared attention can support the SD.***(iii.)SemEval-2016:*** It can be observed from Table 7, SIMS outperforms other models with an overall F1 score 71.24, especially in climate (C), Hillary (H), and abortion (AB) target domains as these domains have clearer separation of emotions and their intensity levels among different stances while performing similarly with S-MDMT for feminism (F) target with F1 score as 63.27. Hence, it proves that our model generalizes well with different targets and domains and can be suited to other topics as well for the task of SD.

Model	**Precision**	**Recall**	**F1 score**	**Acc.**
SIMS [Prop.]	**0.541**	**0.532**	**0.536**	82.05
RoBERTa-Base	0.528	0.502	0.510	82.05
BERT-Base	0.507	0.446	0.464	77.51
BERT-Large	0.530	0.470	0.489	77.78
RoBERTa-Large	0.473	0.507	0.489	**82.54**
DistilBERT	0.497	0.430	0.448	79.37

Table 6: Results of our proposed SIMS approach on publicly available ClimateStance-2022 dataset.

Model	**AT** F_{avg}	**C** F_{avg}	**F** F_{avg}	**H** F_{avg}	**AB** F_{avg}	**Mac** F_{avg}
SIMS	76.10	**71.45**	63.27	**74.05**	**71.34**	**71.24**
Model3	**83.00**	70.00	63.00	67.00	70.00	70.6
MT-LRM -BERT	76.14	53.05	63.12	74.67	70.32	67.46
SP-MT	69.5	63.5	63.2	67.5	70.5	66.84
S-MDMT	69.50	52.49	**63.78**	67.20	67.19	64.03
ESD	66.64	43.82	62.85	67.79	64.94	61.20
HAN	70.53	49.56	57.50	61.23	66.16	61.00
AT-JSS -LEX	69.22	59.18	61.49	68.33	68.41	65.33
SVM-ngram	65.19	42.35	57.46	58.63	66.42	58.01

Table 7: Results of Stance task on SemEval-2016 Dataset with Baselines

6.1 Error Analysis

We identify some scenarios where our SIMS fails to correctly predict the attitude, apart from the skewness of the dataset, which we have tried to address using the resampling technique (*deny:* 17.26%, *ambiguous:* 29.12% and *believe:* 53.62%). ***Incomplete context:*** We note that in some cases where the tweet text is insufficient, the tweet's attitude is labeled as ambiguous. However, the presence of other modalities could help identify the correct attitude, e.g. "The Real Reason We'll Freeze To Death This Winter "https://t.co/xyzabc via @YouTube #climatehoax #climatechange", based on tweet text (without hashtags), ground-truth annotation: ambiguous, SIMS prediction: believe (due to the close proximity between ambiguous and believe stance through similar words such as "real", "freeze", and anticipation emotion), however, the video attached to the tweet provides the correct stance of the tweet, motivating us to focus on the presence of other modalities that could efficiently help with the stance task. ***Implicit/Hidden Stance:*** In some of the tweets (example: "Which begs the question; How many @COP26 attendees will burn FF's? When will COP26 set a leading example by going..#climatechangeisreal..", "How many more countries will you be visiting his year so you pollute the planet further?#climatehoax", it is evident that the authors support climate change but are opposed to government action to combat it. Such tweets are often difficult for the model to identify, so further categorization of stance classes and goals, such as "believer of climate change but against climate action" or similar terms, can further improve SD and could also help authorities to understand viewpoints.

7 Conclusion

In this paper, we curate a novel dataset with annotations for stance, emotion, and emotion intensity labels. Our proposed SIMS induces the emotional aspects associated with the attitude of the tweet and fuses the task-specific and joint feature representations obtained from embedding and attention components using the integration module. Our model outperforms other baselines on climate change datasets by avoiding the drawbacks of previous works that lack an efficient architecture and distinguishes different attitudes even if they consist of similar emotions with different intensity levels. The model's performance on the SemEval dataset further suggests that our framework generalizes well across a variety of domains. Future work on the potential pathways that support the stance detection process will be useful. These include focusing on other modalities to address insufficient context, finer categorization, and other tasks like aspect-based sentiment.

Ethical Statement

We address the United Nations Sustainable Development Goal 13: Climate Action by proposing an approach that helps identify climate change deniers on social media platforms. As climate skeptics on social media platforms such as Twitter spread more and more misinformation and disinformation, it is crucial for government agencies and concerned authorities to identify such misleading content and stop its spread before it becomes harmful to society. Our proposed approach will therefore be beneficial to these organizations when implemented in real-time, as our approach can predict the stance of the tweet based on the textual content as soon as the user posts something. Detailed analysis of the real-time suitability of the model, such as latency and prediction time, is part of our future work depending on the deployment environment. In addition, we are conducting our work with public data from social media. however, to ensure individual privacy, we do not share any personal information. As a result, our publicly available dataset consists only of the tweet IDs and the annotations.

Acknowledgments

This research was partially funded by the SoMeCliCS project under the Volkswagen Stiftung and Niedersächsisches Ministerium für Wissenschaft und Kultur and by the Federal Ministry of Education and Research (BMBF), Germany under the project LeibnizKILabor with grant No. 01DD20003.

References

[Abro *et al.*, 2022] Hassam Uddin Abro, Zafi Sherhan Shah, and Haseeb Abbasi. Analysis of covid-19 effects on well-being - study of reddit posts using natural language processing techniques. In *2022 International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE)*, pages 1–7, 2022.

[Bahri *et al.*, 2013] Mawardi Bahri, Ryuichi Ashino, and Rémi Vaillancourt. Convolution theorems for quaternion fourier transform: properties and applications. In *Abstract and Applied Analysis*, volume 2013. Hindawi, 2013.

[Brosch, 2021] Tobias Brosch. Affect and emotions as drivers of climate change perception and action: a review. *Current Opinion in Behavioral Sciences*, 42:15–21, 2021.

[Chen *et al.*, 2019] Xingyu Chen, Lei Zou, and Bo Zhao. Detecting climate change deniers on twitter using a deep neural network. In *Proceedings of the 2019 11th International Conference on Machine Learning and Computing*, pages 204–210, 2019.

[Falkenberg *et al.*, 2022] Max Falkenberg, Alessandro Galeazzi, Maddalena Torricelli, Niccolò Di Marco, Francesca Larosa, Madalina Sas, Amin Mekacher, Warren Pearce, Fabiana Zollo, Walter Quattrociocchi, et al. Growing polarization around climate change on social media. *Nature Climate Change*, pages 1–8, 2022.

[Fu *et al.*, 2022] Yujie Fu, Xiaoli Li, Yang Li, Suge Wang, Deyu Li, Jian Liao, and Jianxing Zheng. Incorporate opinion-towards for stance detection. *Knowledge-Based Systems*, 246:108657, 2022.

[Jang and Hart, 2015] S Mo Jang and P Sol Hart. Polarized frames on "climate change" and "global warming" across countries and states: Evidence from twitter big data. *Global Environmental Change*, 32:11–17, 2015.

[Kabaghe and Qin, 2020] Chuma Kabaghe and Jason Qin. Classifying tweets based on climate change stance. *Training*, 66(60.9):61, 2020.

[Lacoste *et al.*, 2019] Alexandre Lacoste, Alexandra Luccioni, Victor Schmidt, and Thomas Dandres. Quantifying the carbon emissions of machine learning. *CoRR*, abs/1910.09700, 2019.

[Li and Caragea, 2019] Yingjie Li and Cornelia Caragea. Multi-task stance detection with sentiment and stance lexicons. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6299–6305, 2019.

[Mohammad *et al.*, 2016] Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. Semeval-2016 task 6: Detecting stance in tweets. In *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, pages 31–41, 2016.

[Mohammad, 2018] Saif M. Mohammad. Word affect intensities. In *Proceedings of the 11th Edition of the Language Resources and Evaluation Conference (LREC-2018)*, Miyazaki, Japan, 2018.

[Mou *et al.*, 2015] Lili Mou, Rui Men, Ge Li, Yan Xu, Lu Zhang, Rui Yan, and Zhi Jin. Natural language inference by tree-based convolution and heuristic matching. *arXiv preprint arXiv:1512.08422*, 2015.

[Nguyen *et al.*, 2020] Dat Quoc Nguyen, Thanh Vu, and Anh Tuan Nguyen. BERTweet: A pre-trained language model for English tweets. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Online, October 2020. Association for Computational Linguistics.

[Pörtner *et al.*, 2022] Hans-O Pörtner, Debra C Roberts, Helen Adams, Carolina Adler, Paulina Aldunce, Elham Ali, Rawshan Ara Begum, Richard Betts, Rachel Bezner Kerr, Robbert Biesbroek, et al. *Climate change 2022: Impacts, adaptation and vulnerability*. IPCC Geneva, Switzerland, 2022.

[Qamar *et al.*, 2021] Saira Qamar, Hasan Mujtaba, Hammad Majeed, and Mirza Omer Beg. Relationship identification between conversational agents using emotion analysis. *Cognitive Computation*, 13:673–687, 2021.

[Qin and Ronchieri, 2022] Zhikang Qin and Elisabetta Ronchieri. Exploring pandemics events on twitter by using sentiment analysis and topic modelling. *Applied Sciences*, 12(23):11924, 2022.

[Reveilhac and Schneider, 2023] Maud Reveilhac and Gerold Schneider. Replicable semi-supervised approaches to state-of-the-art stance detection of tweets. *Information Processing Management*, 60(2):103199, 2023.

[Salama and Aboukoura, 2018] Sefat Salama and Khalil Aboukoura. Role of emotions in climate change communication. *Handbook of Climate Change Communication: Vol. 1: Theory of Climate Change Communication*, pages 137–150, 2018.

[Shoeb and de Melo, 2020] Abu Awal Md Shoeb and Gerard de Melo. EmoTag1200: Understanding the association between emojis and emotions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, November 2020.

[Sobhani *et al.*, 2016] Parinaz Sobhani, Saif Mohammad, and Svetlana Kiritchenko. Detecting stance in tweets and analyzing its interaction with sentiment. In *Proceedings of the fifth joint conference on lexical and computational semantics*, pages 159–169, 2016.

[Spitzer *et al.*, 1967] Robert L Spitzer, Jacob Cohen, Joseph L Fleiss, and Jean Endicott. Quantification of agreement in psychiatric diagnosis: A new approach. *Archives of General Psychiatry*, 17(1):83–87, 1967.

[Tyagi *et al.*, 2020a] Aman Tyagi, Matthew Babcock, Kathleen M Carley, and Douglas C Sicker. Polarizing tweets on climate change. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 107–117. Springer, 2020.

[Tyagi *et al.*, 2020b] Aman Tyagi, Joshua Uyheng, and Kathleen M Carley. Affective polarization in online climate change discourse on twitter. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 443–447. IEEE, 2020.

[Upadhyaya and Chandra, 2022] Apoorva Upadhyaya and Joydeep Chandra. Spotting flares: The vital signs of the viral spread of tweets made during communal incidents. *ACM Transactions on the Web*, 16(4):1–28, 2022.

[Upadhyaya *et al.*, 2022a] Apoorva Upadhyaya, Marco Fisichella, and Wolfgang Nejdl. A multi-task model for sentiment aided stance detection of climate change tweets. *arXiv preprint arXiv:2211.03533*, 2022.

[Upadhyaya *et al.*, 2022b] Apoorva Upadhyaya, Catharina Pfeiffer, Oleh Astappiev, Ivana Marenzi, Stefanie Lenzer, Andreas Nehring, and Marco Fisichella. How learnweb can support science education research on climate change in social media. In *12th International Conference on Methodologies and Intelligent Systems for Technology Enhanced Learning*. Springer, 2022.

[Upadhyaya *et al.*, 2023a] Apoorva Upadhyaya, Marco Fisichella, and Wolfgang Nejdl. A multi-task model for emotion and offensive aided stance detection of climate change tweets. In *Proceedings of the ACM Web Conference 2023*, pages 3948–3958, 2023.

[Upadhyaya *et al.*, 2023b] Apoorva Upadhyaya, Marco Fisichella, and Wolfgang Nejdl. Towards sentiment and temporal aided stance detection of climate change tweets. *Information Processing & Management*, 60(4):103325, 2023.

[Vaid *et al.*, 2022] Roopal Vaid, Kartikey Pant, and Manish Shrivastava. Towards fine-grained classification of climate change related social media text. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, pages 434–443, 2022.

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

[Vychegzhanin and Kotelnikov, 2021] Sergey Vychegzhanin and Evgeny Kotelnikov. A new method for stance detection based on feature selection techniques and ensembles of classifiers. *IEEE Access*, 9:134899–134915, 2021.

[Wang and Wang, 2021] Limin Wang and Dexin Wang. Solving stance detection on tweets as multi-domain and multi-task text classification. *IEEE Access*, 9:157780–157789, 2021.

[Wang *et al.*, 2020] Zhongqing Wang, Qingying Sun, Shoushan Li, Qiaoming Zhu, and Guodong Zhou. Neural stance detection with hierarchical linguistic representations. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28:635–645, 2020.

[Wikipedia, 2001] Wikipedia. Convolution. https://en.wikipedia.org/wiki/Convolution, 2001. Accessed: 2023-02-02.

[Wu *et al.*, 2021] Jiaze Wu, Hao Liang, Changsong Ding, Xindi Huang, Jianhua Huang, and Qinghua Peng. Improving the accuracy in classification of blood pressure from photoplethysmography using continuous wavelet transform and deep learning. *International journal of hypertension*, 2021, 2021.

[Zhou and Shen, 2021] Yanmengqian Zhou and Lijiang Shen. Confirmation bias and the persistence of misinformation on climate change. *Communication Research*, page 00936502211028049, 2021.

[Zirikly *et al.*, 2019] Ayah Zirikly, Philip Resnik, Ozlem Uzuner, and Kristy Hollingshead. Clpsych 2019 shared task: Predicting the degree of suicide risk in reddit posts. In *Proceedings of the sixth workshop on computational linguistics and clinical psychology*, pages 24–33, 2019.