Am I No Good? Towards Detecting Perceived Burdensomeness and Thwarted Belongingness from Suicide Notes

Soumitra Ghosh[1] , **Asif Ekbal**[1] , **Pushpak Bhattacharyya**[2]

[1]Department of Computer Science and Engineering, IIT Patna, India
[2]Department of Computer Science and Engineering, IIT Bombay, India
ghosh.soumitra2@gmail.com, asif@iitp.ac.in, pb@cse.iitb.ac.in

Abstract

The World Health Organization (WHO) has emphasized the importance of significantly accelerating suicide prevention efforts to fulfill the United Nations' Sustainable Development Goal (SDG) objective of 2030. In this paper, we present an end-to-end multitask system to address a novel task of detection of two interpersonal risk factors of suicide, *Perceived Burdensomeness (PB)* and *Thwarted Belongingness (TB)* from suicide notes. We also introduce a manually translated code-mixed suicide notes corpus, *CoMCEASE-v2.0*, based on the benchmark CEASE-v2.0 dataset, annotated with temporal orientation, *PB* and *TB* labels. We exploit the temporal orientation and emotion information in the suicide notes to boost overall performance. For comprehensive evaluation of our proposed method, we compare it to several state-of-the-art approaches on the existing CEASE-v2.0 dataset and the newly announced CoMCEASE-v2.0 dataset. Empirical evaluation suggests that temporal and emotional information can substantially improve the detection of *PB* and *TB*.

1 Introduction

According to the World Health Organization (WHO), more than 700,000 people, or one in every 100, committed suicide in 2019. In 2019, the global suicide rate per 100,000 people was 9.0, with 11.2 in the Africa region, 10.5 in Europe, and 10.2 in Southeast Asia. The Eastern Mediterranean area had the lowest rate (6.4). While global suicide rates fell by 36% between 2000 and 2019, the Americas region witnessed a 17% increase[1]. The worldwide catastrophe produced by the COVID-19 pandemic has amplified risk factors for suicidal behaviour, making suicide prevention even more vital now.

According to the interpersonal theory of suicide [Joiner and others, 2005], suicidal desire arises when individuals experience persistent emotions of *perceived burdensomeness (PB)* and *thwarted belongingness (TB)*. The feeling of being a burden on friends, family, and/or society, as well as the potentially dangerous belief that one's death is worth more than one's life, is referred to as *PB*. *TB*, on the other hand, is also a risk factor for developing the impulse to commit suicide if someone feels isolated from friends, family, or other important social networks. Identifying latent vulnerability factors that raise the likelihood of suicide conduct might help determine preventive efforts.

Suicidal behavior is a challenging problem to investigate, requiring huge samples because of the low baseline rates of suicide attempts and deaths in the general population. Most importantly, individuals dying by suicide hinder the employment of any methods for psychological assessments by researchers. In such a case, a suicide note might be a valuable asset in attempting to assess an individual's specific personality status and mind rationale [Pestian and Matykiewicz, 2008].

- **Suicide note excerpt (Hinglish):** अगर मेरी गल्ती थी भी तो मेरी wife & daughter को सुसाइड क्यों करवाया गया , torture करके ।

- **English translation:** *Even if it was my fault then why was my wife & daughter made to commit suicide, by torture.*

The usage of code-mixed forms of communication is common in multilingual and multicultural societies worldwide, notably in India, Bangladesh, Pakistan, Singapore, and even the United States of America, the United Kingdom, etc. The above excerpt is from a suicide note written in Hindi-English code-mixed language.

Specific time perspective orientations (TPOs), such as having an aversive attitude towards the past (Negative Past) and having a hopeless, nihilistic attitude towards life (Fatalistic Present), are associated with suicide ideation. [Shahnaz *et al.*, 2019] found large and significant differences in TPOs between individuals with a history of suicide ideation and non-suicidal participants. Emotion is also recognised to have an important influence in predicting the author's mental state in a suicide note. The association of various emotions of different polarities with varying time perspectives may help learn the underlying emotional feeling conveyed in the content of suicide notes (SNs). Recent advances in computational approaches may be used to better understand various ideas in psychology research, which might aid in the early detection, evaluation, management, and follow-up of those experiencing suicidal thoughts and behaviors.

---

[1]https://news.un.org/en/story/2021/06/1094212

Most existing automated systems need the availability of datasets for training prediction models, and relevant datasets are scarce. Despite the fact that many multilingual speakers in densely populated countries such as India, Pakistan, and Bangladesh use English-Hindi code-mixed language, just a few research [Gupta *et al.*, 2018; Srivastava and Singh, 2020] have attempted to work on this issue. Furthermore, the existing resources [Ghosh *et al.*, 2020; Ghosh *et al.*, 2021] are in English only and annotated at the sentence level, limiting their value in developing automated systems on code-mixed languages and addressing tasks at the document level. During the peak of the COVID-19 epidemic, a large increase in telehealth usage was driven more by persons seeking mental health services than care for physical conditions[2]. Artificial Intelligence (AI)-based automated systems can be incorporated into digital interventions, particularly web and smartphone apps, to enhance user experience and optimise personalised mental health care, more so, in developing countries with high population and minimal healthcare facilities. This motivated us to devise an approach for utilizing the sentence-level information inherent in existing datasets and addressing correlated tasks at the document level by leveraging the underlying connection between them. Also, such an approach can be built on top of current state-of-the-art transformer-based pre-trained models to account for the data scarcity problem and produce good results.

To this end, the current study presents an automated method to simultaneously detect the presence of two interpersonal states of individuals, *PB* and *TB*, from their SNs, given the temporal orientation information and the emotional states of the note content. We introduce, *CoMCEASE-v2.0*, a manually translated code-mixed suicide notes corpus based on the benchmark CEASE-v2.0 dataset and annotated with temporal orientation, *PB*, and *TB* labels. We compare our proposed technique to various state-of-the-art approaches using the existing CEASE-v2.0 dataset and the recently announced CoMCEASE-v2.0 dataset to provide a complete evaluation. The note-level personality annotations with reference note-ids can be accessed from https://www.iitp.ac.in/~ai-nlp-ml/resources.html#CoMCEASE.

The main contributions are summarized below:

- This is the first study towards detection of *perceived burdensomeness* and *thwarted belongingness* from suicide notes using computational methods.

- We introduce a good quality manually translated code-mixed corpus of SNs based on the benchmark CEASE-v2.0 English dataset with manual annotations for *TPO*, *PB* and *TB*.

- We propose a Temporal and Emotion-assisted Multitask Framework (*TEMF*) to jointly detect *PB* and *TB* from SNs.

2   Related Work

Several studies have performed a content analysis of SNs and identified distinguishing features that differentiate SNs from other types of texts [Handelman and Lester, 2007; Pestian *et al.*, 2012]. Suicidal thoughts has been linked to the concepts of *PB* and *TB* in various research. The frequency of occurrence of *PB* and *TB* in SNs was investigated in [Gunn III *et al.*, 2012]. Gender differences in the determinants and correlates of *PB*, *TB*, and acquired suicidal capacity were investigated in [Donker *et al.*, 2014].

Temporal perspective was found to have a significant effect in predicting suicidal thoughts in teenagers [Laghi *et al.*, 2009]. Notable changes were observed in the time perspective of depressed patients when compared to non-depressed people [Lefèvre *et al.*, 2019]. Analyzing the emotional balance in SNs, [Teixeira *et al.*, 2021] discovered links between the concepts and emotional states of persons who committed suicide by recreating the knowledge structure of such notes.

The bulk of suicide note research have concentrated on identifying suicide notes [Handelman and Lester, 2007], with a few focusing on related risk factors for suicide [Ghosh *et al.*, 2021]. Furthermore, massive pre-trained transformer-based language models have taken NLP by storm in recent years, delivering state-of-the-art performance on many downstream tasks. To this end, our study introduces a manually translated code-mixed suicide notes corpus with reliable annotations for *TPO*, *PB*, and *TB*, and use computational techniques to detect two crucial interpersonal risk factors, *PB* and *TB*, using an effective multitask network.

3   Dataset

We consider the benchmark CEASE-v2.0 dataset [Ghosh *et al.*, 2021] for this study. This is the only available emotion annotated corpus of suicide notes publicly available for research purposes to the best of our knowledge. The dataset consists of 4932 sentences from over 350 suicide notes annotated with 15 fine-grained emotions at the sentence level. We used the sentence-to-note reference identifiers (IDs)[3] to establish the order of sentences in the dataset, as the original resource was released at sentence-level, all shuffled to prevent reconstruction and preserve anonymity. We perform three main tasks on this existing dataset: (A). Manually annotate for *PB* and *TB* (at the note-level), (B). Manually annotate for *TPO* (at the sentence-level), and, (C). Produce Hindi-English (Hi-En) code-mixed translations for the CEASE-v2.0 English dataset (at the sentence level).

3.1   Annotator Details

Each task was performed by three annotators (A1, A2, and A3), two doctoral degree holders, and one graduate student, sufficiently acquainted with labeling tasks and well-versed with the concepts of *PB*, *TB* and *TPO*. The annotators are familiar with the nuances of Hindi-English code-mixed communication. All of them are native Hindi speakers, well-proficient in English, and also use Hi-En code-mixed language in daily communications.

---

[2]https://www.frontiersin.org/research-topics/25568/artificial-intelligence-in-mental-health

[3]provided by [Ghosh *et al.*, 2021] on request to use it for research purpose solely

3.2 Annotations for *PB* and *TB*

Each suicide note is labeled *PB/Not-PB* and *TB/Not-TB* to indicate the existence of *PB* and *TB*. The final annotations were acquired by using a majority vote on the labels provided by the three annotators. We obtain a Fleiss-Kappa [Spitzer *et al.*, 1967] agreement of 0.791 and 0.76 for the *PB* and *TB* tasks respectively, which indicates substantial agreement among the annotators.

3.3 Annotation for TPO

We perform the annotations for the temporal orientation at the sentence level as a single note often contains references to various time perspectives. Consider the suicide note excerpt below:

> *The rest of my life would only be a burden for others. I am unable to do anything because of poor health.*

We can observe that the first sentence relates to a *future* event while the following sentence relates to a *present* event. We follow the connotations of *past*, *present* or *future* as described by [Kamila *et al.*, 2018]. Any event(s) that has started and ended is labelled as *past* whereas an ongoing event is marked as *present* and event(s) that are yet to happen are labelled as *future*. We obtain a Fleiss-Kappa agreement of 0.745 which indicates that the annotations are of substantial quality.

3.4 Code-Mixed Manual Translation of Benchmark CEASE-v2.0 Dataset

We construct code-mixed (Hindi-English) translations of each English sentence in the CEASE-v2.0 dataset while preserving *fluency (F)* (syntax) and *adequacy (A)* (semantics). The resulting dataset is known as *CoMCEASE-v2.0*. We use the matrix language frame (MLF) theory [Joshi, 1982] to create the translations, which states that a code-mixed text has a dominant language or matrix language (here, Hindi) and an inserted language or embedded language (here, English). The insertions could be words or bigger elements that adhere to the matrix language's grammatical structure. Based on previous works [Gupta *et al.*, 2018] on code-mixed text creation, we identified three key scopes where the words or components from the original English sentence can be maintained in the translated code-mixed text. These are: Named Entities (NEs) of types 'Person', 'Location' and 'Organization', noun phrases, and adjective words. However, we need not consider the NEs as the original CEASE-v2.0 dataset comes with anonymized forms of all NEs (as also mentioned in [Ghosh *et al.*, 2021]).

Each annotator (say A1) was asked to mark every translated sentence of the next annotator (A2) for both *F* and *A* with an ordinal value from a scale of 1-5[4]. We obtain high average *F* and *A* scores of 4.31 and 4.65, respectively, which indicates that the translations are of good quality. A couple of sample translations from the *CoMCEASE-v2.0* dataset are shown below:

---

[4]**Fluency** - 5: *Flawless*, 4: *Good*, 3: *Non-native*, 2: *Disfluent*, 1: *Incomprehensible*; **Adequacy** - 5: *All*, 4: *Most*, 3: *Much*, 2: *Little*, 1: *None*

1. **Original (En):** *Do not feel sorry for me.*
   Translation (Hi-En): मेरे लिए sorry feel मत करो।

2. **Original (En):** *I was immediately taken in a wheelchair to the psychiatrist's office.*
   Translation (Hi-En): मुझे तुरंत wheelchair से psychiatrist के office ले जाया गया।

HE	A1	A2	A3	Average
F	4.25	4.28	4.40	4.31
A	4.45	4.80	4.71	4.65

Table 1: Translation quality scores from human evaluation

3.5 Dataset statistics

Out of the 364 suicide notes in CEASE-v2.0 dataset which we annotated for *PB* and *TB*, 27% of the notes are marked as either *PB* (10.44%), *TB* (11.81%) or both (4.67%). Total of 55 notes are labelled as *PB* and 60 notes as *TB*. We consider 4885 sentences from the 4932 phrases in the CEASE-v2.0 dataset for the temporality annotations and our experiments. The remaining 47 sentences were not considered due to the following reasons: non-matching mapping IDs (as provided by the authors in [Ghosh *et al.*, 2021]), confusion among all annotators (while annotating for TPO), single-word sentences. The dataset is distributed over the various temporal categories as follows: *past* - 1508 sentences (30.87%), *present* - 2035 sentences (41.66%) and *future* - 1342 sentences (27.47%). The average sentence length in the *CoMCEASE-v2.0* dataset is 16.73, whereas it is 14.96 in the original dataset.

4 Methodology

We present a detailed discussion on the proposed *TEMF* system in the following sub-sections. We depict the overall architecture of the approach in Figure 1.

4.1 Task Definition

Given a suicide note (S^i) with sentences annotated with temporal orientation and emotion labels, the task objective is to detect the presence or absence (1 or 0) of *PB* and *TB*. Let D denotes the dataset consisting of the N suicide notes, then $D = (S^1, S^2, ..., S^N)$. A note S^i is a collection of sentences such that $S^i = (s_1^i, s_2^i, ..., s_n^i)$, where n denotes the number of sentences in the ith note. The emotion labels (E^i) and temporality labels (T^i) corresponding to each sentence in S^i is denoted as $E^i = (e_1^i, e_2^i, ..., e_n^i)$ and $T^i = (t_1^i, t_2^i, ..., t_n^i)$, respectively. The proposed approach work towards maximising the value of the following function:

$$\underset{\theta}{argmax}(\Pi_{i=0}^{N} P(y_{PB}^i, y_{TB}^i | S^i, E^i, T^i; \theta)) \quad (1)$$

where $y \, \epsilon \, \{1, 0\}$ and θ denotes the model parameters to be optimized.

4.2 Temporal and Emotion-assisted Multi-task Framework (TEMF)

This section discusses our proposed system, which comprises the six main components: a). Input Embeddings, b). Document Encoder, c). Sentence Encoder, d). Attention Module
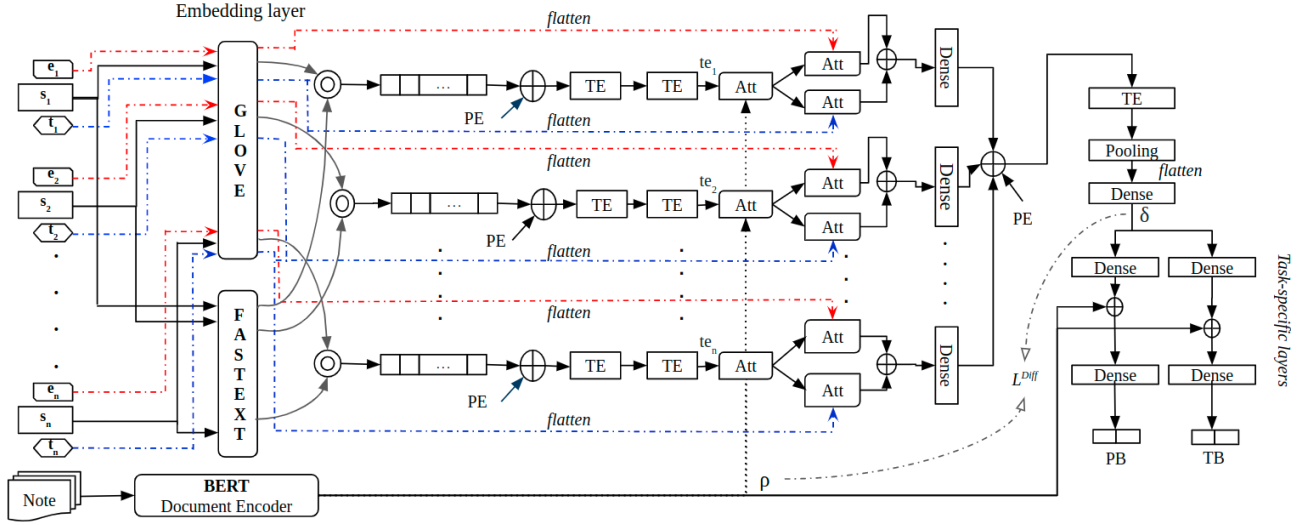
Figure 1: Architecture of the Temporal and Emotion-assisted Multitask Framework (TEMF). PE: positional encodings; TE: transformer encoder; Att: additive attention.

e). High-level Document Abstraction, f). Task-specific Dense layers and Output Layers.

Input Embeddings: Since our focus is on working with the Hinglish dataset, we consider the pre-trained GloVe[5] (English) and fastText[6] (Hindi) word embeddings to capture the semantics of the words of the input sentences. We embed the emotion class and temporality class labels with the word vectors ($\mathbb{R}^{1 \times D}$) from GloVe embeddings as all the class labels are in English. D is the dimension of the embedding (in our case, 300).

The sentence embedding representations from GloVe and fastText are projected to a single vector space by finding their mathematical mean, as inspired by the work in [Wang *et al.*, 2019]. Each s^i is embedded as $W_{emb}^i = \{w_1, w_2, ..., w_c\}$, $W \in c \times d$, where c is the length of s^i.

Document Encoder: We use the BERT (*base*) [Devlin *et al.*, 2019] pre-trained model to encode the information of a suicide note at the document level. The representation from the $[CLS]$ token of the BERT encoder's last encoder layer encodes the whole document's contextual information.

Sentence Encoder: At the sentence level, we employ the transformer encoder (TE) [Vaswani *et al.*, 2017] to capture the local semantic information as well as long-term dependencies. First, the positional encodings (PE^i) and the word embeddings (W_{emb}^i) are added to generate the input embeddings ($IE^i = W_{emb}^i + PE^i$). This allows the model to use word-order information, including relative and absolute positional data. Each transformer module performs a multi-head self-attention (MHA) operation on the input, followed by a fully connected point-wise feed forward network (FFN) to produce knowledge representation, passing the embedding representation of each sentence IE^i through two successive

layers of transformer encoder.

$$R^{(l)^i} = \text{MHA}(IE^i, IE^i, IE^i) \qquad (2)$$

$$S^{(l)^i} = \text{FFN}(R^{(l)^i}) \qquad (3)$$

where l is the number of layers in the network.

Attention Module: We infuse the document-level contextual information from BERT (ρ) to the sentence-level transformer outputs (te_i) by performing additive-attention (Att) [Bahdanau *et al.*, 2015] to make each sentence aware of the global context. The attention-mechanism can be realized by the following equations:

$$\gamma = W_3^T \tanh(W_1 \rho + W_2 te_i^c) \qquad (4)$$

$$\alpha_i = \frac{exp(\gamma(\rho te_i^c))}{\sum_{j=1}^{c} exp(\gamma(\rho te_j^c))} \qquad (5)$$

$$\phi_i = \sum_{t=1}^{c} \alpha_i te_i^c \qquad (6)$$

where W_1, W_2, W_3 are the learnable weight matrices, $tanh$ is a non-linear function.

We further apply two independent additive attention operations on each context-aware sentence output, one with the temporal embedding as the query vector and the other with the emotion class embedding as the query vector. These attention mechanisms enable to focus on the relevant words in the sentence concerning each sentence's temporal and emotional information. The temporal and emotion-aware representations of each sentence are concatenated and passed through a sentence-specific dense layer.

High-level Document Abstraction: The dense representations give an abstract representation of the input sentences, which are concatenated ($\oplus$) together and passed through another TE to learn the temporal and emotion-aware global context information. The sentence-level position embedding is also added with the input to the TE to use the sentence-order

---

[5]https://nlp.stanford.edu/data/glove.840B.300d.zip
[6]https://fasttext.cc/docs/en/crawl-vectors.html

information. Hierarchical transformer encoders have shown promising output while capturing document-level contextual information. We apply a max-pooling operation on the transformer output (δ) to capture the essential latent semantic information across the input.

Task-specific Dense layers and Output Layers: The max-pooled output (Δ) is passed to two task-specific dense layers that capture the intermediate features particular to each task, followed by two task-specific fully connected layers with *softmax* activation function that acts as the output classification layers. The BERT document representation is added to the task-specific intermediate representations to enhance each feature set.

Calculation of loss: We train the model through a unified loss function as shown below:

$$L = \alpha * L^{PB} + \beta * L^{TB} + L^{Diff} \quad (7)$$

Where L^{PB} and L^{TB} are the categorical cross-entropy losses, and α and β are the loss weights for the two tasks.

Alongside the cross-entropy losses (L^T), we calculate a differential loss to minimize the mean squared difference (L^{Diff}) between the output representations of BERT and TE to build an adequate shared feature space. The intuition is to reduce the vulnerability mainly due to two reasons: (1) the variation in different encoding methods and (2) handling two document representations (one generated from the whole document and one generated from constituent sentences.

$$L^T = -\frac{1}{N}\sum_{i=1}^{N} t_i log O^T \text{ and } \mathrm{L}^{Diff} = \sum_{b=1}^{n}(\delta_i - \rho_i)^2 \quad (8)$$

where $t_i \ \epsilon \ \{0, 1\}$ is the ground truth label of each task and T denotes the two tasks.

5 Experiments and Results

In this section, we discuss the experiments performed and the results and analysis.

5.1 Baselines

We evaluate the efficacy of our proposed method on the CEASE-v2.0 dataset considering the following state-of-the-art systems as baselines: Convolutional Neural Network [Kim, 2014], Convolutional Neural Network+Context Long Short Term Memory (CNN+cLSTM) [Poria *et al.*, 2017], BERT [Devlin *et al.*, 2019] and CMSEKI [Ghosh *et al.*, 2021]. In CNN+cLSTM system, CNN is used for feature extraction at utterance level followed by a cLSTM to learn context-aware utterance representations. The CMSEKI system was introduced in the work that presented CEASE-v2.0 dataset, addressing detection of depression, sentiment and emotion, using common-sense knowledge. We adapted the CMSEKI system to address our *PB* and *TB* tasks. For evaluation on the *CoMCEASE-v2.0* dataset, we consider the following state-of-the-art systems as baselines: (1) K-Max Pooling CNN (k-max CNN) [Wang *et al.*, 2019] that uses meta-embeddings formed from GloVe and fastText, (2) multilingual BERT (mBERT) [Devlin *et al.*, 2019] pre-trained model.

Models	F1^{PB} (%)	F1^{TB} (%)
CNNST [Kim, 2014]	46.6	50.4
CNN+cLSTMST [Poria *et al.*, 2017]	45.4	45.8
BERT$_{Base}$ST [Devlin *et al.*, 2019]	51.16	52.02
CMSEKI^{MT} [Ghosh *et al.*, 2021]	49.68	53.56
TEMF^{MT} *(proposed)*	**54.02**	**56.64**
TMF^{MT} *(ablation 1)*	53.33	52.43
EMF^{MT} *(ablation 2)*	51.63	48

Table 2: Results on the CEASE-v2.0 dataset. Here, ST and MT denote single-task and multi-task system, respectively.

5.2 Implementation Details

We used 10-fold cross-validation on both datasets and the macro-F1 measure to evaluate all models, as the data is highly skewed over the various classes. For our experiments, we chose the average note length of 13 as the document length. The sequence lengths were set to the average sentence lengths of 15 and 17 for the CEASE-v2.0 and *CoMCEASE-v2.0* datasets, respectively. For each TE, we used 5 self-attention heads with an embedding dimension of 300 and a feed-forward dimension of 600. All models were trained using a batch size of 4 (to maximise GPU use) and 6 epochs. We chose a learning rate of 2e-5 for the *TEMF^{MT}* system. The Grid Search approach was used to fine-tune parameters. The best model on the validation set for each fold were saved for testing. We used the BERT$_{Base}$ (English and multilingual) pre-trained models[7]. All the dense layers (except the output dense, which uses softmax activation) used ReLU activation. The task-specific fully-connected and output layers have 128 and 2 neurons each for each task. We trained all the models with Adam optimizer using backpropagation. The implementations were done on an NVIDIA GeForce RTX 2080 Ti GPU. To account for the non-determinism of different TensorFlow GPU operations, we report F1 scores averaged across the five 10-fold cross-validation runs.

5.3 Results and Analysis

We observe from Table 2 that the proposed *TEMF^{MT}* system outperforms all the baselines commendably with improvements of 2.86% (on *PB* Task) and 3.08% (on *TB* task) from the next best performing models BERT$_{Base}$ST and CMSEKI^{MT}, respectively, on the CEASE-v2.0 dataset. The *TEMF^{MT}* system also obtains highest F1-scores of 51.27% and 55.21% on the CoM-CEASE-v2.0 dataset, with improvements of 2.34% (on *PB* Task) and 4.12% (on *TB* task) from the next best performing model mBERT$_{Base}$ST. The *PB* task yielded higher scores for all models than the *TB* task, which might be due to a slightly higher number of *TB* occurrences in the dataset than *PB*.
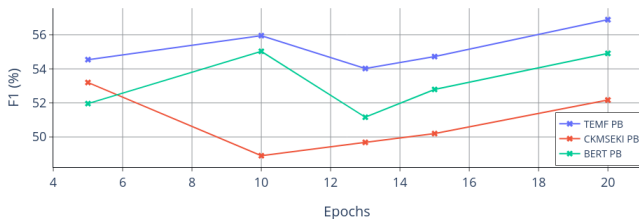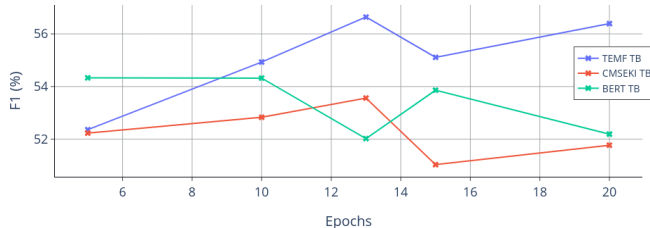
Varying Context Length: Figure 2 and 3 depicts the F1 results of our *TEMF* system with comparison to the best performing baselines on different context lengths (5, 10, 13, 15, 20). We observe an upward trend towards the improvement

---

[7]https://github.com/google-research/bert

Models	F1^{PB} (%)	F1^{TB} (%)
k-max CNN [Wang *et al.*, 2019]	47.43	51.06
mBERT$_{Base}$ST [Devlin *et al.*, 2019]	48.93	51.09
TEMF^{MT} *(proposed)*	**51.27**	**55.21**
TMF^{MT} *(ablation 1)*	50.03	47.01
EMF^{MT} *(ablation 2)*	49.37	47.02

Table 3: Results on *CoMCEASE-v2.0* dataset

of *PB* and *TB* scores by our *TEMF^{MT}* system, outperforming the CMSEKI^{MT} and BERT^{MT} systems, as contextual information is increased. Furthermore, on the TB task, unlike the BERT^{MT} and CMSEKI^{MT} systems where performance deteriorates with the increase in context length, the proposed *TEMF^{MT}* model improves substantially.

Figure 2: Graphical depiction of results of *TEMF* on varying context length of suicide notes.

Figure 3: Graphical depiction of results of *TEMF* on varying context length of suicide notes.

Comparison with Prior Works: We observe from Tables 2 and 3 that our proposed *TEMF^{MT}* system outperforms the considered state-of-the-art baselines on both the CEASE-v2.0 and CoM-CEASE-v.0 datasets. The best performing baselines on the *PB* and *TB* tasks are BERT$_{Base}$ST and CMSEKI^{MT}, respectively for the CEASE-v2.0 dataset and mBERT$_{Base}$ST for the CoMCEASE-v2.0 dataset. Moreover, the poor performance of the various baselines on both datasets show how critical it is for present context modelling approaches to perceive interpersonal content from texts.

Ablation Study: We modify TEMF^{MT} by removing the temporal orientation labels and emotion labels in the input, one at a time, to develop EMF^{MT} and TMF^{MT}, respectively. Tables 2 and 3 show that eliminating either one of the time or emotional information from the input has a significant detrimental influence on the overall performance of the model for both the datasets. Furthermore, the results show that removing the temporal information from the input has a more significant impact than removing the emotion data.

Error Analysis: We identify some circumstances in which our proposed technique fails to accurately categorize *PB* and *TB* instances.

1. **Inconsistent grammar rules:** Writing style in suicide notes is often informal and does not follow the standard rules of sentence structure. This makes it difficult for our model to recognize syntactic information correctly. For example, *Kept me in a locked shower stall for 1 hour. Next sent me burnt food. Giant bugs crawling over my hands. NO FUN!!*

2. **Insufficient context:** Lack of knowledge of the context makes the task of detection for *PB* and *TB* challenging. For example, *They tried to get me, I got them first!*

3. **Sarcastic undertone:** The proposed method finds it hard to capture sarcasm, and the problem becomes even critical with lack of sufficient context. For example, *It's amazing to look at my grade sheet.*

6   Conclusion

In this work, we proposed a Temporal and Emotion-aware Multitask Framework (TEMF) for joint detection of two interpersonal risk factors of suicide: *perceived burdensomeness* and *thwarted belongingness*, from suicide notes. We introduced a manually translated code-mixed Hindi-English corpus (*CoMCEASE-v2.0*) of suicide notes with manual annotations for *TPO*, *PB* and *TB*. We empirically observe that the infusion of temporal and emotion information has a positive impact on improving the performances of the *PB* and *TB* detection tasks from suicide notes. Also, the proposed method enables utilizing the existing sentence-level CEASE-v2.0 dataset to address note-level tasks.

Future study into the potential pathways that increase the ability to commit suicide would be beneficial as well, such as duration, intensity, and frequency for *PB*, *TB* and acquired capability for self-injury.

Ethical Consideration

Our resource creation utilizes publicly available CEASE-v2.0 [Ghosh *et al.*, 2021] benchmark suicide notes dataset. We followed the data usage restrictions and did not violate any copyright issues. This study was also evaluated and approved by our Institutional Review Board (IRB). We shall make the code and data available for research purposes (on acceptance), through appropriate data agreement procedure.

Acknowledgement

References

[Bahdanau *et al.*, 2015] Dzmitry Bahdanau, Kyung Hyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In *ICLR*, 2015.

[Devlin *et al.*, 2019] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, 2019.

[Donker *et al.*, 2014] Tara Donker, Philip J Batterham, Kimberly A Van Orden, and Helen Christensen. Gender-differences in risk factors for suicidal behaviour identified by perceived burdensomeness, thwarted belongingness and acquired capability: cross-sectional analysis from a longitudinal cohort study. *BMC psychology*, 2(1):1–10, 2014.

[Ghosh *et al.*, 2020] Soumitra Ghosh, Asif Ekbal, and Pushpak Bhattacharyya. Cease, a corpus of emotion annotated suicide notes in english. In *LREC*, 2020.

[Ghosh *et al.*, 2021] Soumitra Ghosh, Asif Ekbal, and Pushpak Bhattacharyya. A multitask framework to detect depression, sentiment and multi-label emotion from suicide notes. *Cognitive Computation*, pages 1–20, 2021.

[Gunn III *et al.*, 2012] John F Gunn III, David Lester, Janet Haines, and Christopher L Williams. Thwarted belongingness and perceived burdensomeness in suicide notes. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 33(3):178, 2012.

[Gupta *et al.*, 2018] Deepak Gupta, Pabitra Lenka, Asif Ekbal, and Pushpak Bhattacharyya. Uncovering code-mixed challenges: A framework for linguistically driven question generation and neural based question answering. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, 2018.

[Handelman and Lester, 2007] Lori D Handelman and David Lester. The content of suicide notes from attempters and completers. *Crisis*, 28(2):102–104, 2007.

[Joiner and others, 2005] Thomas E Joiner et al. *Why people die by suicide*. Harvard University Press, 2005.

[Joshi, 1982] Aravind Joshi. Processing of sentences with intra-sentential code-switching. In *Coling*, 1982.

[Kamila *et al.*, 2018] Sabyasachi Kamila, Mohammed Hasanuzzaman, Asif Ekbal, Pushpak Bhattacharyya, and Andy Way. Fine-grained temporal orientation and its relationship with psycho-demographic correlates. In *NAACL-HLT*, 2018.

[Kim, 2014] Yoon Kim. Convolutional neural networks for sentence classification. In *EMNLP*, 2014.

[Laghi *et al.*, 2009] Fiorenzo Laghi, Roberto Baiocco, Maria D'Alessio, and Grazia Gurrieri. Suicidal ideation and time perspective in high school students. *European Psychiatry*, 24(1):41–46, 2009.

[Lefèvre *et al.*, 2019] Héline Kaya Lefèvre, Christine Mirabel-Sarron, Aurélie Docteur, Virginie Leclerc, Alexandra Laszcz, Philip Gorwood, and Catherine Bungener. Time perspective differences between depressed patients and non-depressed participants, and their relationships with depressive and anxiety symptoms. *Journal of Affective Disorders*, 246:320–326, 2019.

[Pestian and Matykiewicz, 2008] JP Pestian and P Matykiewicz. Classification of suicide notes using natural language processing. *Proceedings of ACL Bio NLP*, 967, 2008.

[Pestian *et al.*, 2012] John P Pestian, Pawel Matykiewicz, Michelle Linn-Gust, Brett South, Ozlem Uzuner, Jan Wiebe, K Bretonnel Cohen, John Hurdle, and Christopher Brew. Sentiment analysis of suicide notes: A shared task. *Biomedical informatics insights*, 5:BII–S9042, 2012.

[Poria *et al.*, 2017] Soujanya Poria, Erik Cambria, Devamanyu Hazarika, Navonil Majumder, Amir Zadeh, and Louis-Philippe Morency. Context-dependent sentiment analysis in user-generated videos. In *ACL*, pages 873–883, 2017.

[Shahnaz *et al.*, 2019] Arezoo Shahnaz, Boaz Y Saffer, and E David Klonsky. Examining time perspective orientation in suicide ideation and suicide attempts. *Journal of Social and Clinical Psychology*, 38(8):627–646, 2019.

[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.

[Srivastava and Singh, 2020] Vivek Srivastava and Mayank Singh. Phinc: A parallel hinglish social media code-mixed corpus for machine translation. In *Proceedings of W-NUT*, 2020.

[Teixeira *et al.*, 2021] Andreia Sofia Teixeira, Szymon Talaga, Trevor James Swanson, and Massimo Stella. Revealing semantic and emotional structure of suicide notes with cognitive network science. *Scientific reports*, 11(1):1–15, 2021.

[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. *Advances in neural information processing systems*, 30, 2017.

[Wang *et al.*, 2019] Bin Wang, Xiaobing Zhou, and Xuejie Zhang. Ynuwb at semeval-2019 task 6: K-max pooling cnn with average meta-embedding for identifying offensive language. In *SEMEVAL*, 2019.