Tech Science Press

Predicting Violence-Induced Stress in an Arabic Social Media Forum

Abeer Abdulaziz AlArfaj[1], Nada Ali Hakami[2,*] and Hanan Ahmed Hosni Mahmoud[1]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, 11671, Saudi Arabia
[2]Department of Computer Science, College of Computer Science and Information Technology, Jazan University, Jazan, Saudi Arabia
*Corresponding Author: Nada Ali Hakami. Email: nmhakami@jazanu.edu.sa

Abstract: Social Media such as Facebook plays a substantial role in virtual communities by sharing ideas and ideologies among different populations over time. Social interaction analysis aids in defining people's emotions and aids in assessing public attitudes, towards different issues such as violence against women and children. In this paper, we proposed an Arabic language prediction model to identify the issue of Violence-Induced Stress in social media. We searched for Arabic posts of many countries through Facebook application programming interface (API). We discovered that the stress state of a battered woman is usually related to her friend's stress states on Facebook. We applied a large real database from Facebook platforms to analytically investigate the correlation of violence-induced stress states and the victim interactions on social media. We extracted a set of textual, spatial, and interaction attributes from various features. Therefore, we are proposing a hybrid model–an interaction graph model incorporated in a deep learning neural model to leverage post content and interaction data for violence-induced stress detection. Experiments depict that our proposed hybrid model can enhance the prediction performance by 10% in F1-measure. Also, considering the user interaction information can learn an interesting phenomenon, where, the sparse social interactions of violence-induced stress stressed victims is higher by around 15% percent non-battered users, signifying that the structure of the friends of such victims is less connected than non-stressed users.

Keywords: Arabic language analysis; violence-induced stress detection; hybrid model; deep learning

1 Introduction

Social media involvement has increased greatly in recent times, as it permits them to: display ideas on various matters, to exhibit their emotions about challenges they have daily, and to discuss various issues [1–3]. Facebook is the most significant platform that is usually utilized daily [4–6], most people in Arab countries use Facebook to post and comment on different issues including violence-induced stress against women. Statistics have shown that Saudi Arabia uses Facebook extensively, with 10 million people interacting on the Facebook [3]. researchers usually explore Facebook posts by analyzing them and

classifying these posts using sentiment analysis. Few studies for Arabic language posts have been published due to the difficulty in analyzing Arabic texts [7]. The research involving sentiment analysis of Facebook posts in Arabic is divided into three approaches [5]. The Arabic language can be divided into formal, and Colloquial Arabic that is used daily among people [6].

Sentiment Analysis is usually utilized for data mining and deep learning concerning opinions. Opinion deep learning is a science of analytically classifying thoughts, desires, behaviors, and feelings of people about several types of issues [7]. Sentiment Analysis has been extensively used to cover multiple domains, such as healthcare and social issues. Studied issues can be caused by different types of entities, such as organizations, event-driven issues, and humanities [7]. Sentiment Analysis is a multi-step identification problem that performs classification on the Aspect, document, or Sentence Level. Sentiment Analysis has a first phase of classifying the Aspect, document, or Sentence polarity into three classes: positive class implying positive thoughts, or negative class with bad thoughts or neutral [8]. In this research, we are utilizing Facebook posts as sentence sentiment analysis. Research in sentiment language analysis especially Arabic depicts supervised and unsupervised sentiment analysis [9–11]. Sentiment analysis utilizes lexical and deep learning methodologies [12–15]. Deep learning classification and identification are used for identifying the schism of Arabic language posts [15]. In [16] and [17], the authors discussed coverless information hiding techniques that can help uncover negative thoughts in the underlying posts.

The research gaps in earlier research are the lack of linguistic models to classify violence-induced stress. Limited thoughts of transforming Facebook post language have been employed. Also, validation using language-level ground truth has not been performed. There is no labeled dataset for the Arabic language that can be used to predict violence-induced stress from Facebook posts or any other social media interaction.

Therefore, this paper, we propose a novel technique for the Arabic language Facebook posts taxonomy. the proposed model is divided into two phases. The first is data collection through Facebook API to collect posts and comments by the language of the posts. We also examined the noises in the collected data which contains many unrecognized symbols and characters that are not present as part of the Arabic language. In the second phase, we will build the following two classifiers: Support vector machines (SVM) and k-nearest neighbors (KNN) and then combine them with machine learning methods. We validate our model using F-Measure and then we perform analyses using RapidMiner and the Tableau validation tools [18].

This paper is divided as follows, Section 2, summarizes the related works. The proposed model is depicted in Section 3. While in Section 4, we explain the validation and verification experiments and discuss the results. In Section 5, we present the conclusions of our research.

2 Literature Review

2.1 Facebook Post and Comments Analysis

Facebook is a social media tool for writing posts and posting comments daily. To gain real insight and identify different features from those posts, research methods need to apply classification methods to them. The authors in [19], utilized supervised learning and sentiment analysis of Arabic posts by constructing a sentiment lexicon database, the lexicon database includes 200,000 Arabic lexical terms. they gathered user posts from Facebook and Twitter with a pre-processing step. They achieved accuracy results of 88.5% [19].

Authors in [20] extracted multiple features and classified them to negative, positive, neutral, or mixed feelings. Their feature dataset is divided into four categories as defined. They extracted 20 features about each post and 6 properties for each comment. This model utilized feature extraction propagation and the precision metric was computed for each category. They used more properties about the features themselves such as positive and negative word count, and post-length.

Social media can be utilized to predict violence-induced stress behavior occurrence against and can Improve women's lifestyles and wellbeing. Huge amounts of negative feelings and fear are shared online, especially on the Facebook platform. This available data can help communities to allocate violent behavior and rescue women in need. Research shows that online support groups can lead to benefits for people when they can link to others with similar situations. Their participants can share their feelings about improper situations and the shame that surrounds them [21]. online forums generate a huge amount of data, that includes valuable insights. But, these data have to go through advanced analysis techniques such as natural language analysis and deep learning for obtaining useful information, e.g., for identifying and predicting violent behavior victims. There are many research papers done in predicting such situations in English platforms, which can contribute highly to rescuing victims in an efficient way [22–25]. Authors in [22] used electroencephalogram (EEG) data to identify cases with anxiety through different processing techniques. Authors in [23] classified features such as Bhattacharya and t-tests, utilizing logistic regression and naïve Bayes classifiers. Authors in [24] utilized EEG obtained from forehead sensors to analyze the generated signals. Authors in [25] used a natural language processor to identify depression from Twitter tweets by analyzing lexicon that is used by miserable people. They utilized verbal analysis and term count models. Authors in [26] extracted features using Dirichlet allocation for topic identification and used the N-gram technique for word analysis. Various machine learning techniques are used such as AdaBoost. Unfortunately, little research has been done on Arabic social platforms to detect violence-induced stress behavior against women. Therefore, this research tends to close such gap by presenting analytical models for the prediction of violence-induced stress behavior of from online Arabic posting.

Previous research focused on monitoring various aspects such as mental disorders and other health issues using natural language processing and analysis. It conducted a search for significant features of the languages for the sentiments analysis of the posts. Also, other aspects such as violence-induced stress or suicidal thoughts are examined using text analysis accompanied with data collection techniques. questionnaires, confessions, online forums and annotated posts.

2.1.1 Predicting Violence-Induced Stress Incidents Through Questionnaires

Data is collected from Facebook posts from participants who reply to self-evaluation questionnaires. self-evaluation using Questionnaires is recorded in [27]. Authors in [28] perceived answers anonymously to an anxiety self-evaluation assessment related to violence victims, in addition to their Facebook posts. Violence was identified when victims reported the incidence. The authors in [29] collected the scores of negativity in Facebook posts on predictability scales. Behavioral trends linked to fear, mixed feelings, reaching for others are detected from such data. They proposed a statistical decision tree classifier that detects early signs of the risk of violence incidence with a ranking of the future risk involved. Authors in [30] found numerous aspects in the data collected by victims, such as negative feelings, isolation, and repeated pattern of fear terms. The authors in [31] used behavioral trends on Facebook to predict depression following violence incidence by self-reported victims through their Facebook status. They performed t-tests to distinguish the victims' behaviors in first violence incidence and in the repeated ones. Authors in [23] studied the prediction of post-traumatic anxiety through analyzing Facebook posts. The authors collected huge amount of such posts and analyze the tools of sentiment features extraction, such as ANEW. These data were used in the learning random forest classifiers to identify violence victims from non-victims. Authors in [29–32] collected data from Twitter to predict violence incidence from tweeter. Authors in [33], employed analysis dynamics in physician rating websites during the early wave of the COVID-19 pandemic. Also in [34], sentiment analysis was employed.

Arabic language analyzing is very challenging due to its rich morphological dictionary where a single word can have different meanings. Also, Arabic has a formal and informal version where they seem very

different morphologically and also syntactically [9]. Most Arabic posts at the social platforms are written using informal Arabic language with no defined orthographies. Arabic natural language processing models usually use Arabic formal language [10] with degrading performance for Arabic informal language [11]. Also, many intelligent models are needed to analyze various Arabic text problems [12]. In [13], the authors utilized sentiment analysis solutions [13–16], while other researchers used semantic analysis to predict emotions in Arabic texts [11–18].

Limitations Occur, in Facebook posts content, for violence against women detection. First, posts from battered women are usually short and limited to an average of nine words on Facebook. Also, battered women do not express their victim states directly in posts. Also, battered women with high emotional violence-induced stress have low level activity on Facebook, as stated by recent research [3]. This low activity causes data sparsity problem, which may curb the accuracy of Facebook content violence-induced stress detection process especially in the Arabic language. Fig. 1 depicts a Facebook post that contains only eight a women saying "I wished to go home for the Eid holiday", "أتمنى ان اذهب لبيت أهلي لقضاء أجازة", "العيد". Although there is no mention of violence complain in the post, from the continuation comments done by the user and her Facebook friends, the model can discover that the user is a victim of some kind of domestic violence. Thus, simply trusting only the user post violence-related stress is inadequate. Also, interaction graphs of battered women *vs*. non-battered women are depicted in Fig. 2.

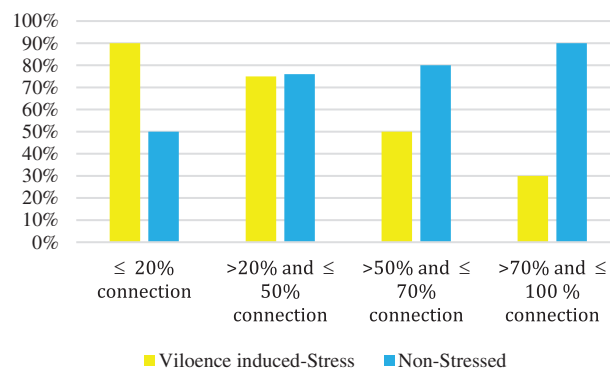**Figure 1:** Example of a Facebook post and interactive comments

Figure 2: Battered women interaction graph *vs*. non-battered women

2.2 Our Proposed Work

The proposed model defined attributes representing violence-induced stress from Facebook post interaction and user attributes. (1) Facebook post-interaction attributes from comments of user's single post, and (2) user attributes from user's collected posts. The post-interaction attributes are linguistic, and social interaction such as likes, shares, and comments collected from a single-post interaction list. The user attributes are a collection of behavior attributes as extracted from a user's post in time T, and interaction attributes computed from Facebook interactions with other users.

To leverage the user attributes as well as post-content data, this paper is proposing a new hybrid methodology of interaction graph model and a deep learning neural network that shows the ability to learn unified dormant features from several modalities, and an interaction graph model is employed for correlation modeling.

The overall phases are:

1. The design of a deep learning convolutional neural network (D-CNN) with cross encoders to produce user attributes from post attributes;
2. Defining an attribute labeled graph to incorporate user interaction attributes, posting attributes, and the learned-content attributes for violence-induced stress victimizing detection.

We tested and validated the proposed model for various attributes on a real dataset from Facebook. Experiments illustrate that by developing the social media interaction attributes, our model can enhance the detection F1-score by 10% over the state-of-art models. This specifies that the selected attributes are suitable cues in facing the data sparsity problem. Also, the introduced model can efficiently incorporate Facebook post content and interaction to improve the violence-induced stress detection accuracy.

The proposed research also employed in-depth experiments on a large dataset from Facebook Arabic posts. Our model performs an analysis of the linear dependency of users' stress status and their interactions on Facebook, to face the problem of (1) interaction content, by studying the differences between battered and non-battered users' interactions content; and (2) interaction structure, by studying the structural diversity, and strong/Non-strong tie. Our research exposes some social occurrences. For instance, our study concludes that the count of sparse connection of violence-induced stress victims is 16% higher than non-victims, representing that the social structure of victim's friends inclines to be less connected.

The contributions of this Research are:

- Proposing a hybrid model incorporating deep learning with interaction graph modeling to influence both content and interaction attributes to improve the detection of violence-induced stress.
- Building an Arabic violence-induced stress posting datasets from the ground-truth dataset from the Facebook platform.
- Presenting an in-depth investigation on a real large dataset and extracted correlations between social media interactions and violence.
- Constructing social structure of violence-induced stress victims.

The rest of this paper is presented as follows: Section 2 gives an overview of related works. Section 3 describes the hybrid model and the training method for the violence-induced stress detection model, Section 4 details the proposed models and presents the data collection in the labeling phase. Section 4 presents the experiments conducted with performance evaluation. conclusions of this research are summarized in Section 5.

3 The Proposed Model

In this section, we are presenting the hybrid model that incorporates deep learning process with interaction graph modeling to influence both content and interaction attributes to improve the detection of violence-induced stress. We propose an Arabic violence-induced stress posting datasets from the ground-truth dataset from the Facebook platform.

3.1 The Proposed Framework

We defined attributes representing violence-induced stress from Facebook post interaction and user attributes. (1) Facebook post-interaction attributes from comments of user's single post, and (2) user

attributes from user's collected posts in a specific period. The post-interaction attributes are linguistic, and social interaction such as likes, shares, and comments collected from a single-post interaction list. The user attributes are a collection of behavior attributes as extracted from a user's post in time T, and interaction attributes computed from Facebook interactions with other users.

To leverage the user attributes as well as post-content data, we are proposing a new hybrid methodology of interaction graph model and a deep learning neural network has the ability to learn unified dormant features from several modalities, and an interaction graph model is employed for correlation modeling.

The overall phases are:

1) The design of a D-CNN with cross encoders to produce user attributes from post attributes;
2) Defining an attribute labeled graph to incorporate user interaction attributes, posting attributes, and the learned content attributes for violence-induced stress victimizing detection.

The framework was tested and validated using the proposed model for various attributes on a real dataset from Facebook. Experiments illustrate that by developing the social media interaction attributes, our model can enhance the detection F1-score by 10% over the state-of-art models. This specifies that the selected attributes are suitable cues in facing the data sparsity problem. Also, the introduced model can efficiently incorporate Facebook post content and interaction to improve the violence-induced stress detection accuracy.

We also employed in-depth experiments on a large dataset from Facebook Arabic posts. Our model performs an analysis of the linear dependency of users' stress status and their interactions on Facebook, and face the problem of (1) interaction content, by studying the differences between battered and non-battered users' interactions content; and (2) interaction structure, by studying the structural diversity, and strong/Non-strong tie. Our research exposes some social occurrences. For instance, we conclude that the count of sparse connection of violence-induced stress victims is 16% higher than non-victims, representing that the social structure of victim's friends inclines to be less connected.

The contributions of this Research are:

- Proposing a hybrid model incorporating deep learning with interaction graph modeling to influence both content and interaction attributes to improve the detection of violence-induced stress.
- We build Arabic violence-induced stress posting datasets from the ground-truth dataset from the Facebook platform.
- We convey an in-depth investigation on a real large dataset and extracted correlations between social media interactions and violence.
- We constructed social structures of violence-induced stress victims.

3.2 Attributes Categorization

To face the problem of violence-induced stress detection, we define two subsets of attributes to compute the dissimilarities between the violence-induced stressed and non-stressed users on Facebook platforms:

1. Post-level attributes of the single post;
2. user attributes extracted from posts in time T.

3.2.1 Post-Level Attributes

Post-level attributes designate the linguistic, spatial content, and the social interaction factors such as likes, shares, and comments of a single post.

For linguistic features, our model uses the common linguistic properties in Arabic sentiment analysis. We adopt ALP [29], an Arabic Language Platform, to analyze and tokenize lexicons, and then map them

into positive, negative, or neutral emotions. The model also mapped emotions emoji into linguistic attributes and also punctuation marks into emotions such as I as "unbelievable", ? into "really",…, etc. Facebook utilizes Unicode as a depiction of emojis [34], which can be mined directly. Linguistic, visual, and interaction attributes are depicted in Tab. 1. Spatial attributes are represented by API from OpenCV to implement image processing and color mapping, e.g., warm color *vs*. cool color, as depicted in Tab. 1. The model adopted a six-color attributes theme for affective image classification. In our research, we did not employ the direct emotional identification process as spatial attributes for multi-dimensional spatial attributes for deep learning. Direct spatial emotional prediction produces only a few dimensions as an attribute.

Table 1: Linguistic, visual, and interactions attributes from Facebook posts

Features	Category	Description
Linguistic features	Emotion phrases or words	Number of optimistic and non-optimistic emotion words
	Emotion emoji	Number of optimistic and non-optimistic emoji
	Punctuations and associated words	!? …
	Grade Adverbs and associated words	Examples; "I am a little sad", " I am strongly afraid" used scores are from 1 to 5
Visual	Color theme	Dominant colors: Red-Blue-Green (R-B-G)
	Contrast	The average value for brightness
	Hue warmth	The ratio of cool Hue
	Hue clearness	Saturation of dull colors less than 0.6
Social Interaction	Facebook Interaction	A number of share, comments, likes, dislikes, etc…..

3.2.2 User Attributes

User attributes are mined from several user posts in a period T. T is defined as one month in the proposed model because violence-induced stress often is caused by aggregate events. Victims usually express their violence-induced stress in a chain of posts. Also, the aforesaid interaction patterns in time T usually include data useful for violence-induced stress detection. Besides, the data in posts are sparse, therefore, the proposed framework incorporates additional data around posts, e.g., interactions through comments with friends. Therefore, properly selected user attributes can deliver a macro-scope of violence-induced stress conditions, and evade missing data and noises.

3.3 The Proposed Model Framework

Problems in the detection of violence-induced stress are as follows:

1. Defining user attributes from user posts and interactions face the lack of modality in the posts
2. leveraging social media interaction structure patterns is not an easy task.

To face these problems, we suggest a new hybrid model by incorporating an interaction graph model with a deep neural network (D-CNN). D-CNN is used for learning unified hidden features from compound modalities, and the interaction graph model (IGM) captures linear dependency correlations. In the following subsections, the model architecture will be presented.

3.3.1 Model Architecture

The architecture of the proposed model is depicted in Fig. 3. The model inputs are post-level attributes, user attributes, and interaction attributes. The solution is depicted as follows:
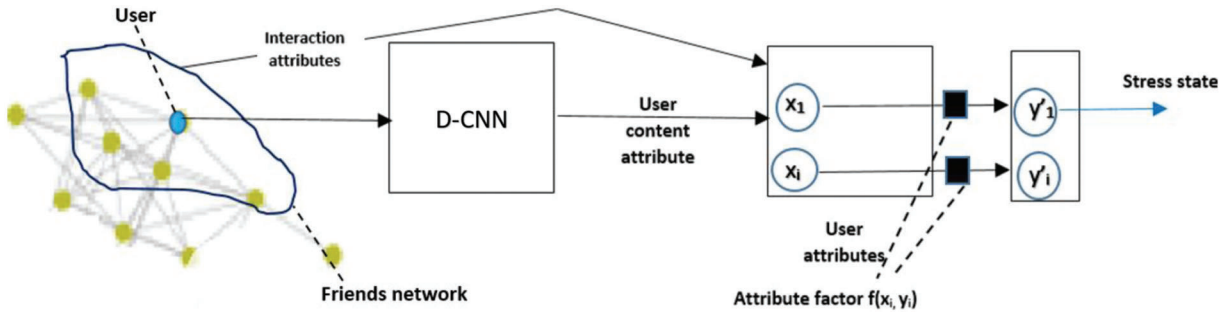

Figure 3: Our proposed model comprises two phases. The first phase is a deep CNN (D-CNN). The second phase is the interaction graph model (IGM). The D-CNN produces user content features using filters and forwards them as inputs to the IGM

1. D-CNN with cross encoders to produce interaction attributes from post-level attributes. The D-CNN is utilized in learning local attributes.
2. A labeled interaction graph includes all aspects of user attributes. The interaction graph model is usually employed in social media modeling. It is used in extracting social dependency for various prediction tasks.

As an example, the user with a blue dot. post attributes are handled via a D-CNN with cross encoders to general user-content- content attributes. The user- attributes are denoted by xi. A user attribute xi has three attributes: user content, posting behavior, and interaction aspect. features of all users go through the same path. In the IGM, all three aspects connect user attributes to the corresponding violence-induced stress state. Social aspects link the violence-induced stress state of various users. Dynamic aspects link violence-induced stress state some time period of time. The output of user violence-induced stress state at time t is yi.

For the user depicted in Fig. 3: attributes from posts are mapped to post attributes as depicted with various modalities in different colors, white color signifies modalities not existent in the post. The post-level attributes are then input to cross encoders that are inserted in the D-CNN. The D-CNN will add attributes from cross encoders into the grouped user attributes by pooling feature maps. The user attributes of a user X at time t are indicated by $u_i^t (i = 1, 2, 3, \ldots)$ as depicted in Fig. 3. The violence-induced stress status of each user X at time t is defined by $V_i^t (i = 1, 2, 3, \ldots)$. The user attributes and the violence-induced stress states are associated with an attribute aspect, while violence-induced stress states of various users are associated with social aspects. Violence-induced stress states of a single user at contiguous times are associated by dynamic aspects. We describe the graph by computing several aspects, we can then develop all users' violence-induced stress states over time.

3.3.2 Learning Model

Combining user attributes induces the problem of Missing modality. To solve such problem, our model uses a cross-encoder to capture the representation of every single post with multiple modalities. The text, social and spatial attributes of a post are denoted by A_{tx}, A_s, A_{sp}, the cross encoder (CE) is defined as follows:

$$CE = f\left(w_{tx}A_{tx} + w_{s}A_{s} + w_{sp}A_{sp} + b\right) \qquad (1)$$

$$\left(v_{tx} + v_{s} + v_{sp}\right) = 1/(1 + exp(-w_{D}I)) \qquad (2)$$

where, I is the modality-invariant, w_{tx}, w_{s}, w_{sp} and b are encoder parameters, whereas, w_{D} is the decoder parameter. $\left(v_{tx} + v_{s} + v_{sp}\right)$ are the reconstructed computed inputs.

The basic modality of the cross encoder is to construct missing training modalities and to learn cross-data correlation such as that negative text emotions correlate with cool hue in images. The training of the cross encoder used a cropped subset of data $\left(v_{tx}, v_{s}, v_{sp}\right)$ that requires utilize tile all the three modalities.

The gradient descent algorithm trains the cross encoder, the energy function ψ is defined in terms of the parameters of the cross encoder as:

$$\psi = \frac{1}{2}\sum_{M}^{n} \| v_{M} - A_{M} \|^{2} + \frac{\xi}{2}\sum_{M}^{n} \| v_{M} \|^{2} - \| A_{M} \|^{2} \qquad (3)$$

The post attributes, which are extracted from a user's post in time T, generate a time series representing a user as a time series of posts, we employ a large learning D-CNN [32]. The D-CNN inter-connections than standard CNN layers. It emphasizes learning static local features from image pixels and audio time series. user content features are learned from a set of posts in a time series to denote stress statuses over a period T forming a 1-D series.

Cross encoder map patches with missing aspects into modality-invariant aspects. They are employed as filters in the D-CNN over the post series to form one feature map. Thus the hidden user content features can be produced from the post-level attributes of a single post.

Pooling is used to encapsulate feature maps into fewer instances. Though users might have an unlike number of posts in different time period T, the time over which the posts are collected is the same. We pool each feature map into one pooled feature. Since we engage pooling over time T rather than a constant number of posts, we will apply the mean-over-time pooling criteria (M-O-T). M-O-T is computed by adding the activations, since the post instances are collected over the same T.

3.3.3 Correlations Between Post Content and Interactions

Time-dependent correlation is difficult to be classified utilizing classifiers such as SVM, we employ a labeled interaction graph model, to represent social interactions and posts for learning and classifying user-level violence-induced stress states.

Learning hidden correlations between Facebook post content and friend's interactions is difficult Time-linear dependency is tough to be modeled utilizing SVM classifiers. Therefore, we utilize an interaction graph model (IGM), to incorporate interactions and Facebook content for learning the classification of user-level violence-induced stress states.

We employ an objective function β that optimizes the conditional probability of violence-induced stress states Z given a series of attributes, as depicted in Algorithm 1.

Algorithm 1: Learning by Interaction Graph

Input: *time-varying attribute series with* violence-induced stress states *on user nodes in the graph G, learning rate* R;

Output: *Feature parameter* F = {*f1, f2, f3*} *and* violence-induced stress *vector* Z;

    1. *Initialize* Z *using a random number generator;*

    2. *Initialize the parameter* F;

    3. **Do until** *convergence;*

        {

        Compute the gradients for the three parameters; $\nabla f1$, $\nabla f2$, $\nabla f3$;

        Compute f1= f1+R× ∇ f1;

        Compute *f2= f2+R× ∇ f2;*

        Compute *f3= f3+R× ∇ f3;*

        }

The learning process is a predictive estimate model to estimate the parameter F = {*f1, f2, f3*} from the labeled items and to optimize the objective β such that $F^{*} = \arg\max_{P} \varnothing(F)$

For maximizing β, we employ the gradient decent model. We compute the decent gradients for each parameter in our β function as follows:

$$\frac{d\varnothing}{df1} = E\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right) - E_{prob.(f1)}(Z/G)\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right) \tag{4}$$

$$\frac{d\varnothing}{df2} = E\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right) - E_{prob.(f2)}(Z/G)\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right) \tag{5}$$

$$\frac{d\varnothing}{df3} = E\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right) - E_{prob.(f3)}(Z/G)\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right) \tag{6}$$

where, $E\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right)$ is the mean of the attributes summation over Y and G with in the training set,

and $E_{prob.(p1)}(Z/G)\left(\sum_{j=1}^{J}\sum_{i=1}^{I}g(x_j^i,\ z_j^i)\right)$ is the mean of the attribute summation aspects computed by the estimated formula.

The violence-induced stress *vector* Z is finally calculated as follows:

$$Z^{*} = \arg\max\ [\![\varnothing(Z/G,F)]\!] \tag{7}$$

4 Experimental Results

In this section, we are introducing the efficacy and impact of our hybrid deep learning model on violence-induced stress detection.

4.1 Data Collection

We performed data collection to evaluate and test the proposed model, we collect two databases using labeling methods, which are depicted as follows: Database DATASET1. In this research, we have built a ground truth database with correct labels from large Facebook data. The data which was collected from Facebook is usually enormous. We engaged an automatic sentence pattern tagging to excerpt labels from crawled data over Facebook. We crawled 200 million posts data via Facebook REST APIs from January 2016 to January 2020. Facebook users post textual content, images, and engage in interactions (share, like, comment). the user-produced contents, and relationship, also encloses plentiful information for an investigation that can inspect emotions and stresses.

The next step was to recognize the violence-induced stress state from posts over a period of time (in our research we choose one month). Facing the massive number of posted images, we employ tags and users' comments for computerized image labeling [29]. This is performed by searching for posts including patterns such as "I feel so sad, how can he do that" and "I trusted him so much", which indicate the presence of stress. The time periods enclosing such phrases are labeled as "violence-induced stress" periods. Likewise, we detect "non-stress" periods from posts that include phrases such as "I feel relaxed this week". These patterns have displayed high precision *vs*. user state labels annotated by online surveys in social media.

we collected over 1200 months of posts that are considered as stressed, and over 1000 months of posts labelled as non-stressed users' posts. There are 260,000 posts from 5500 users in total. We utilize this database for analysis and testing in the experimental studies, which is denoted by Dataset1 as depicted in Tab. 2.

Table 2: Overview of the Facebook-Violence-induced stress Dataset1 and data distribution

	Violence-induced stress	Non-Stress	Total
Number of posts	139,000	121,000	260,000
Number of users	2900	2600	5500
Number of time periods	1200 months	1000 months	2200 months
Average number of posts/month	116 post/month	121/month	118 post/month
Average number of month/user	0.41	0.38	0.4
Average interacting user/month	5.21	7.21	

Another collected dataset denoted as Dataset2 in this paper. We confirmed our model reliability by comparing our results to the ground truth labeling of Dataset2 as depicted in Tab. 3. Dataset2 is collected from the users who were surveyed and answered about the score of a violence-induced stress scale. The violence-induced stress scale is considered as stress when the scale is higher than 80, otherwise it is labelled as non-stress. We collected the scores answered by the users, and employed the scale scores as labels denoting the ground truth.

4.2 Experiments

In this research we set the experiments by training our proposed supervised using Dataset1. Testing is then performed using Dataset2 because it is the ground truth. The validation of the proposed model is verified by using a ground truth data with labels. In the experiments, we use 8-fold validation, with 20 random runs. We compared our model with several classification models as described in Tab. 4. We utilize the Scikit-learning model for comparison. In the experiment evaluation for the proposed model, we consider both

effectiveness and efficiency. The classification effectiveness and performance are measured using the accuracy, recall, precision, and F1-Measure. While efficiency is evaluated by computing the CPU time of the training phase. The experiments are executed on an x64 based processor with 4.9 GHz Intel Core and 16 GB memory.

Table 3: Details of Dataset2 (ground truth) including data distribution

	Violence-induced stress	Non-Stress	Total
Number of posts	1459	1845	3304
Number of users	98	112	210
Number of time periods	98 months	112 months	210 months
Average number of posts/month	14.9 post/month	16.5/month	15.7 post/month

Table 4: Other models description for comparison

Model	Description
Logistic classification Regression model [19]	Training a logistic model and then classify cases in the test subset.
Support Vector Machine classifier [5]	Binary classification with radial base function kernel.
Random Forest [23]	Decision trees with random forests of attributes
Gradient Decision Tree [13]	Training is performed on a gradient decision tree with feature extraction
Deep Neural Network with autoencoders [29]	User stress classification with a CNN network with cross encoders

4.3 Experimental Results

To evaluate the performance of our proposed model, we bear several tests utilizing diverse models using Facebook's Dataset1. In the various experiments, we utilized the three attributes depicted in Section 2, namely: user interaction, posting behavior and content attributes produced from the post-level attributes by the proposed D-CNN. Tab. 5 describes the experimental results. Our model achieves higher results than the other models depicted before in Tab. 4. The results verified that our model can efficiently influence the Facebook interaction and attributes for violence-induced stress detection. Compared with the model in [29], which employs user-level stress classification using social data, our model enhances the detection accuracy by 10% on the F1-measure scoring metric. These results exhibit the feasibility of the detection of many conditions using data from social interactions. our model can considerably increase the performance by using social interaction in addition to other data. We implemented also t-tests with p-value is less than 0.01, demonstrating the improvement in our model is statistically significant.
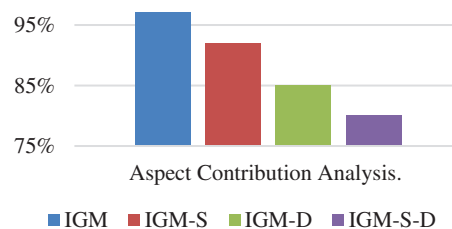
To compare the efficiency of our model *vs*. the other models, we utilize the training CPU execution time for each model. The results are depicted in Tab. 5. All models have good performance with respect to the training CPU time of the different models which varies from tens of seconds to tens of minutes. Our model has a slightly slower training speed but achieves higher performance compared to the other models.

Table 5: Comparison of results of different models

Model	Accuracy	Recall	Precision	F1-measure	CPU execution time for training
Logistic classification Regression model [19]	75.5%	82.3%	75.8%	83.11%	42 s
Support Vector Machine classifier [5]	72.6%	79.2%	72.5%	80.43%	10 min
Random Forest [23]	77.02%	83.8%	79.22%	85.12%	69.5 s
Gradient Decision Tree [13]	81.45%	86.5%	83.66%	87.22%	268.12 s
Deep Neural Network with auto encoders [29]	83.21%	87.00%	84.78%	89.28%	26 min
Our proposed model IGM + D-CNN	92.52%	96.41%	92.3%	94.31%	20 min

The following factors were tested in all compared models and they are as follows:

a) Aspect Contribution Analysis. The selection of aspects is essential to the effectiveness of the Interaction Graph Model. We considered three aspects in our model, i.e., attribute aspect, social aspect, and dynamic aspect. The impact of these aspects in our proposed model, we studied the model performance with different permutations of aspects, as depicted in Fig. 4. in the beginning, we utilized the three aspects together, denoted by Interaction graph model (IGM), then we eliminate the subsequent aspects in order: social aspect, dynamic aspect and then both aspects, denoted as IGM-S, IGM-D, and IGM-S-D. We realized that the model with the least performance is related to the sole inclusion of the attribute aspect. Conversely, including the attribute aspect and any of the social or dynamic aspect yields an improvement of the model performance. We can denote that social or dynamic aspects are highly effective for stress detection. Especially, the social aspect considerably enhances the model accuracy to an average of 91%. The highest model performance is obtained when utilizing the three aspects together.

Aspect Contribution Analysis.

■ IGM ■ IGM-S ■ IGM-D ■ IGM-S-D

Figure 4: Aspect contribution analysis

b) Training scalability Analysis. To assess the data size factor of our model, we trained the learning module with different data sizes, and compare the F1-Score of model prediction performance. In this experiments, we utilize the three aspects as model input. Fig. 5 depicts the prediction performance trend with different size subsets of the whole data. It is obvious that when training the model with small portion (10%), our proposed model fails to reach an acceptable prediction performance. The prediction performance gets better by increasing the size of the training data, reaching 90% accuracy with 40% portion of the whole data, increased to 92% compared when utilizing 50% of the data. Furthermore, the prediction performance increases with the increasing size of the training data, which proves the verify the model scalability with respect to dataset size.
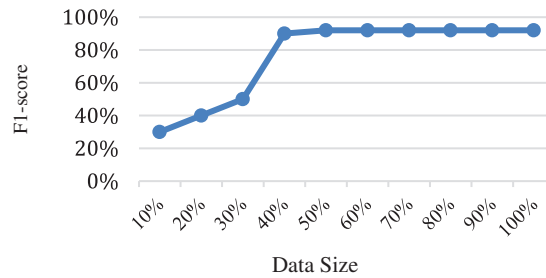
Figure 5: Training scalability analysis

c) Model convergence. We also studied the model convergence for the IGM algorithm as depicted in Fig. 6, where the F1-score increases with the iterations count. the IGM model converges within 2,200 iterations, which is sufficiently fast to perform effective model training on large datasets.
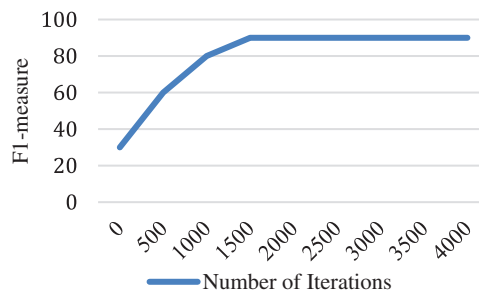

Figure 6: F1-score *vs*. iterations count

d) Impact of D-CNN Size. We have to note that the D-CNN size is an acute issue in setting up the learning model. Thin CNNs produce trivial network that cannot detect correlation in the training data. However, very deep CNN yield a complex model which is hard to tune and can cause over-fitting. To choose a suitable D-CNN for prediction, we evaluated the D-CNN architecture with various count of layers. Fig. 7 displays the experiments. It is noted that 2 to 4 layers do not obtain satisfactory performance. five layers attains a peak performance, while six to seven layer-model does not achieve any better result. This is because the D-CNN becomes too large to be tuned within acceptable training time.
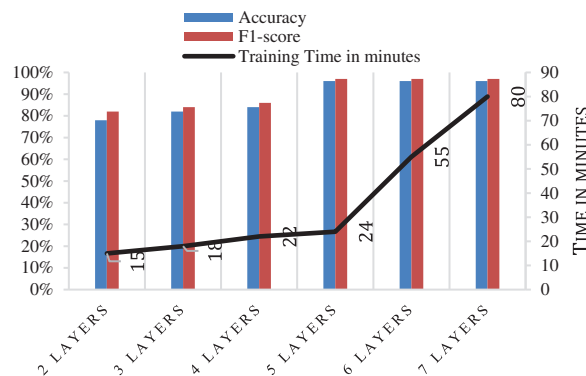

Figure 7: Impact of D-CNN Size

Experiments denoting attribute contribution and its impact are depicted in Fig. 8 we utilized post-level attributes and user attributes from single post as well as user interactions in a month period. To assess the attribute contribution and compare our model performance *vs*. other models in the literature by utilizing different attributes combination as input. The attributes are classified into three classes: post-level and user attributes as well as interaction behavior denoted as PL, UA, and IB respectively. We compare the prediction accuracy of the proposed D-CNN+IGM with SVM model and CNN model with auto-encoders, with all the probable combinations of the three groups of attributes.
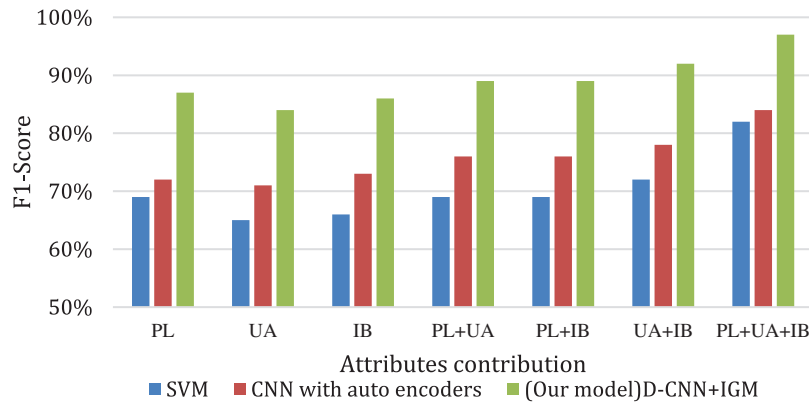

Figure 8: Results of different attribute combinations

From the comparison results, we find that the compared models attain the best performance when employing the three group of attributes. When employing only the post-level attributes, the F1-score of our model decreases to 87% which is acceptable. While for the SVM classifier, the F1-score decreases to 69% percent, which is very bad for a binary prediction. This result validates the impact of feature combination of CNN.

5 Conclusion

Violence-induced stress is a threatening psychological health problem. Also, violence can be life threatening if not stopped. Most violence victims can be confined and have only social media to express their emotions. Most people use social media platforms to post their activities and emotions. These activities facilitate applying online social media data for violence-induced stress detection. In this paper, we proposed a supervised deep learning model for detecting stress states from users' Facebook data, analyzing posts' content and user interaction data. Sentiments are analyzed by collecting Facebook posts and comments over several months about violence-induced stress against women especially. We used the Facebook tag system to get the matching topic. Facebook posts are then labeled. We utilized real Facebook posts data as a training basis, we investigated the correlation between stress states and their interaction activities. To influence both post content and interaction data of users' posts, we presented a hybrid model which incorporates the Interaction graph model (IGM) with a deep CNN. In this research, we learned several phenomena of stress. We establish that the sparse connection degree of stressed victims is more prominent than that of non-stressed users. stressed users' friends hold less interaction than non-stressed users. Also, Experiments depict that our proposed hybrid model can enhance the prediction performance by 10% in F1-measure.

Limitation of this research is the feature extraction method is limited and should be expanded to incorporate unsupervised learning, also it should be global and limited to one language.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] K. Smith, "Mental health: A world of depression," *Nature News*, vol. 15, no. 7, pp. 210–230, 2019.

[2] H. Yazdavar, H. Alolimat and A. Sheth, "Semi-supervised approach to monitoring clinical depressive symptoms in social media," in *Proc. of IEEE/ACM Int. Conf. on Advances in Social Networks Analysis and Mining*, New York, NY, USA, pp. 1191–1198, 2020.

[3] R. Subhani, W. Mumtaz and A. Malik, "Machine learning framework for the detection of mental stress at multiple levels," *IEEE Access*, vol. 5, pp. 19545–19556, 2020.

[4] M. Tadesse, H. Lin and L. Yang, "Detection of depression-related posts in reddit social media forum," *IEEE Access*, vol. 7, pp. 44883–44893, 2019.

[5] F. Elgohary, T. Sultan and M. Dosoky, "A computational approach for analysing and detecting emotions in arabic text," *International Journal of Engineering Research and Applications*, vol. 3, no. 1, pp. 100–107, 2019.

[6] N. Abdulla and S. Alrifai, "Towards improving the lexicon-based approach for Arabic sentiment analysis," *International Journal of Information Technology and Web Engineering*, vol. 9, no. 3, pp. 55–71, 2019.

[7] M. Mageed, M. Diab and S. Kubler, "Samar: Subjectivity and sentiment analysis for Arabic social media," *Computer Speech & Language*, vol. 28, no. 1, pp. 20–37, 2019.

[8] M. Abdullah, M. Alasawa and S. Mahrous, "Emotions extraction from Arabic tweets," *International Journal of Computers and Applications*, vol. 7, no. 2, pp. 1–15, 2021.

[9] M. DeChoudhury, M. Gamon and E. Horvitz, "Predicting depression via social media," in *Proc. Seventh Int. AAAI Conf. on Weblogs and Social Media*, Cairo, Egypt, pp. 44–49, 2019.

[10] M. DeChoudhury, S. Counts and A. Hoff, "Characterizing and predicting postpartum depression from shared Facebook data," in *Proc. of the 20th ACM Conf. on Computer Supported Cooperative Work & Social Computing*, NY, USA, pp. 626–638, 2019.

[11] S. Tsugawa, Y. Kikuchi and H. Ohsaki, "Recognizing depression from twitter activity," in *Proc. of the 33rd Annual ACM Conf. on Human Factors in Computing Systems*, Paris, France, pp. 3196–3217, 2015.

[12] D. Blei, A. Nagi and M. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, no. 1, pp. 993–1022, 2019.

[13] H. Schwartz, J. Eichstaedt and L. Ungar, "Towards assessing changes in degree of depression through Facebook," in *Proc. of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, Alexandria, GA, USA, pp. 121–125, 2019.

[14] G. Coppersmith, M. Dredi and M. Mitchell, "Clpsych 2015 shared task: Depression and PTSD on twitter," in *Proc. of the Computation and Linguistic*, Cairo, Egypt, pp. 31–39, 2018.

[15] T. Pedersen and M. Alan, "Screening Twitter users for depression and PTSD with lexical decision lists," in *Proc. of the 4th Workshop on Computational Psychology*, Riyadh, Saudi Arabia, pp. 46–53, 2018.

[16] X. Zhang, W. Zhang, W. Sun, X. Sun and S. Jha, "A robust 3-D medical watermarking based on wavelet transform for data protection," *Computer Systems Science & Engineering*, vol. 41, no. 3, pp. 1043–1056, 2022.

[17] X. Zhang, X. Sun, W. Sun and S. K. Jha, "Robust reversible audio watermarking scheme for telemedicine and privacy protection," *Computers Materials & Continua*, vol. 71, no. 2, pp. 3035–3050, 2022.

[18] Y. Sun and D. Wu, "A relief based feature extraction algorithm," in *Proc. of the SIAM Int. Conf. on Data Mining*, London, England, pp. 195–218, 2018.

[19] C. Lin, C. Chuang and Z. Cao, "Forehead EEG in support of future feasible personal healthcare solutions: Sleep management, headache prevention, and depression treatment," *IEEE Access*, vol. 5, pp. 10612–10621, 2020.

[20] I. Barandiaran, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 234–244, 2018.

[21] H. Peng, C. Xia and X. Li, "Multivariate pattern analysis of EEG-based functional connectivity: A study on the identification of depression," *IEEE Access*, vol. 7, pp. 92630–92641, 2019.

[22] T. Kasteren, G. Englebienne and B. Krose, "An activity monitoring system for elderly care using generative and discriminative models," *Personal and Ubiquitous Computing*, vol. 19, no. 6, pp. 489–498, 2020.

[23] S. Bagroy, P. Kumaraguru and M. Choudhury, "A social media based index of mental well-being in college campuses," in *Proc. of the CHI Conf. on Human Factors in Computing Systems*, Pekin, China, pp. 1634–1646, 2020.

[24] E. Kiciman and A. Summar, "Discovering shifts to suicidal ideation from mental health content in social media," in *Proc. of the CCS Conf. of Computing Systems*, Athens, Greece, pp. 2098–2110, 2019.

[25] M. Kern, G. Park and L. Ungar, "Gaining insights from social media language: Methodologies and challenges," *Psychological Methods*, vol. 21, no. 4, pp. 507–525, 2020.

[26] P. Cavazosehg and L. Bierut, "A content analysis of depression-related tweets," *Computers in Human Behaviour*, vol. 54, no. 3, pp. 351–357, 2019.

[27] D. Mowery, C. Bryan and M. Conway, "Towards developing an annotation scheme for depressive disorder symptoms: A preliminary study using twitter data," in *Proc. of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, Lafayette, LA, USA, pp. 89–98, 2015.

[28] L. Baer, D. Jacobs and J. Reizes, "Development of a brief screening instrument: The hands," *Psychotherapy and Psychosomatics*, vol. 69, no. 1, pp. 35–41, 2020.

[29] M. Saad, "The impact of text pre-processing and term weighting on text classification," *Text Pre-processing*, vol. 29, no. 3, pp. 351–361, 2020.

[30] B. Lakshmi and S. Parthasarathy, "Human action recognition using median background and max pool convolution with nearest neighbour," *International Journal of Ambient Computing and Intelligence*, vol. 10, no. 2, pp. 34–47, 2019.

[31] R. Wang and M. Wang, "Web text categorization based on statistical merging algorithm in big data environment," *International Journal of Ambient Computing and Intelligence*, vol. 10, no. 3, pp. 20–32, 2019.

[32] M. Ali and A. Mahmoud, "Adam deep learning with SOM for human sentiment classification," *International Journal of Ambient Computing and Intelligence*, vol. 10, no. 3, pp. 92–116, 2019.

[33] S. Adnan, Y. Xiangbin and A. Hafez, "Mining patient opinion to evaluate the service quality in healthcare: A deep-learning approach," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 1, pp. 2925–2942, 2019.

[34] S. Adnan, Y. Xiangbin and J. Wang, "Mining topic and sentiment dynamics in physician rating websites during the early wave of the COVID-19 pandemic: Machine learning approach," *International Journal of Medical Informatics*, vol. 9, no. 2, pp. 112–123, 2021.