This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

A Knowledge-Based Recommendation System that includes Sentiment Analysis and Deep Learning

Renata L. Rosa, Gisele M. Schwartz, Wilson V. Ruggiero, and Demóstenes Z. Rodríguez, *Senior Member, IEEE*

***Abstract*—Online social networks (OSN) provide relevant information on users' opinion about different themes. Thus, applications, such as monitoring and recommendation systems (RS) can collect and analyze this data. This paper presents a Knowledge-Based Recommendation System (KBRS), which includes an emotional health monitoring system to detect users with potential psychological disturbances, specifically, depression and stress. Depending on the monitoring results, the KBRS, based on ontologies and sentiment analysis, is activated to send happy, calm, relaxing, or motivational messages to users with psychological disturbances. Also, the solution includes a mechanism to send warning messages to authorized persons, in case a depression disturbance is detected by the monitoring system. The detection of sentences with depressive and stressful content is performed through a Convolutional Neural Network (CNN) and a Bi-directional Long Short-Term Memory (BLSTM) - Recurrent Neural Networks (RNN); the proposed method reached an accuracy of 0.89 and 0.90 to detect depressed and stressed users, respectively. Experimental results show that the proposed KBRS reached a rating of 94% of very satisfied users, as opposed to 69% reached by a RS without the use of neither a sentiment metric nor ontologies. Additionally, subjective test results demonstrated that the proposed solution consumes low memory, processing, and energy from current mobile electronic devices.**

***Index Terms*—Sentiment analysis, knowledge personalization and customization, recommendation system, social networks, deep learning.**

I. Introduction

The number of active online social network (OSN) users has grown considerably, and some studies indicate there will be 2.95 billion users by the end of 2020 [1]. This high number of users, on OSN, is mainly due to the increase of the number of mobile devices, such as smartphones and tablets, connected to the Internet. Currently, OSN have become a rich and universal means of opinion expression, feelings, and they reflect the bad habits or wellness practices of each user. In recent years, the analysis of the messages posted on OSN have been used by many applications [2], [3] in the industry of health care informatics.

The sentiments and emotions, expressed on the messages posted on OSN, provide clues to different aspects of the behavior of users; for instance, sentences containing words with negative meaning may indicate sadness, stress, or dissatisfaction [4]. Conversely, it can be inferred that if a person is in a positive mood state, this person can be more self-confident and emotionally stable [5]. Users have different behaviors on OSN, if the sentiment intensity value of posted sentences remain at low levels, or if it frequently changes from high to low levels and vice versa, these facts can indicate some emotional disturbance, such as depression or stress events [6]. Hancock, Gee, Ciaccio et al. [7] and Liu [8] observed that users write short sentences when they are experiencing a period of depression. Also, these users use the first person pronoun in their sentences and suffer from chronic insomnia. Therefore, their behavior can be reflected in the sentences posted on OSN. The presence of certain words in the sentences can be monitored and analyzed to identify users at a high risk of attempting suicide and an appropriate intervention can take place [9].

Depression is one of the most prevalent mental disorders in all regions and cultures around the world [10]. Unfortunately, depression recognition rate remains low. Most of the studies about health systems [11]–[13] use sensor devices to detect mental disorders. In [14], the proposed trained classifier, which is trained using electroencephalogram signals, is able to detect stress with an average accuracy of 80.45% using 4-fold cross validation. In [15], authors use heart rate variability data to propose a classification model that considers different stress levels, baseline, mild stress and severe stress, reaching accuracy values of 74%, 81%, 82%, respectively.

There is a scarce number of studies that use textual information from OSN data to detect physiological disorders. Xue et al. [16] use different machine learning (ML) classifiers to perform emotion classification focused on psychological disorders from micro-blogs, reaching an average accuracy of 80%. In [17], the proposed model to detect stress based on the information of Twitter activities reached an accuracy of 69%. In [18], authors study the causes of postpartum depression using OSN information. ML algorithms are also used in studies about mood monitoring systems analyzing messages from OSN [19], reaching an accuracy of 57%. Ma and Hovy [20] introduce a network architecture to analyze sentences meaning through character-level representations by using a combination of Long Short-Term Memory (LSTM), a Convolutional Neural Network (CNN) and Conditional Random Field (CRF). Lample et al. [21] combine Recurrent Neural Networks (RNNs) with CRFs to obtain the best results on Named Entity Recognition (NER) datasets. A bi-directional LSTM (BLSTM), an improved version of the LSTM, is also used for labeling tasks.

Renata L. Rosa and Demostenes Z. Rodríguez are with the Federal University of Lavras, Brazil. CEP: 37200-000.

Gisele M. Schwartz is with Universidade Estadual Paulista Júlio de Mesquita Filho, Biosciences Institute of Rio Claro, Brazil. CEP: 13506-900.

Wilson V. Ruggiero is with the Polytechnic School of the University of São Paulo, Brazil. CEP: 05508-010.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

The deep learning approach has been explored in several areas [22], such as personality analysis [23], age group classification on OSN [24], sentiment analysis [25], among others. However, this approach is not widely explored in psychopathology studies. In this context, our study intends to test the performance of deep learning algorithms in scenarios of depressed, stressed, and non-depressed and non-stressed users' detection.

A Recommendation System (RS) application can be used as a method to enhance the user's emotional health, improving the person's mood in case of negative emotional states [26]. RS based on ontology is being used for health purposes [27], presenting reliable results from diseases treatment plans.

In this context, the main goal of this work is to introduce an RS that uses an approach named Knowledge-Based Recommendation System (KBRS), which aggregates an ontology collection for health scenarios, named Nuadu [28], which is not addressed in other RSs designed to improve emotional health. The proposed KBRS also includes the sentiment analysis approach and an emotional health monitoring system. The monitoring system filters sentences from an OSN that allows to identify potential users with depression or stress conditions. To accomplish this task, an objective method based on an BLSTM-RNN is used to detect potential psychological disorders, along an CNN. Later, a KBRS is activated to send happy, calm, relaxing, or motivational messages to these users. These messages have different intensity levels depending on the sentiment intensity of the sentences posted on an OSN, which is determined by an enhanced sentiment analysis metric, eSM2. This proposed metric is based on a word-dictionary, considering the Portuguese language, and enhanced with additional information such as user's profile data, user's geographic location, and the theme of the sentence. Furthermore, in the cases of depression detection, the solution sends warning messages to authorized people who are previously registered in the system. According to the subjective test results, users reported high satisfaction with the KBRS, improving their emotional states. Tests were also performed with a traditional RS without the Nuadu ontology and the eSM2 for comparing it with the proposed KBRS. Furthermore, subjective tests reported that the application running on the user mobile electronic device had low complexity and low-power consumption.

In short, this paper proposes:

- An innovative solution to monitor and to detect potential users with emotional disturbances, based on the classification of sentences with depressive or stressed content. A CNN is used for character-level representation and BLSTM-RNN for the disorder entity recognition.
- An improved RS that incorporates a personalized sentiment metric, named eSM2, and a health ontology approach, specifically the Nuadu ontology.
- An enhanced sentiment metric. It is studied and validated that the sentiment metric performance is improved by incorporating the user's profile information, the geographic location, and the theme of the sentence.
- An application installed on a mobile device, which is easy to be used, and it consumes low memory, processing, and energy.

The remainder of this paper is structured as follows: Section II presents the related work about RS, emotional health monitoring system, machine learning and deep learning approach, and sentiment and affective analysis; Section III presents the methodology of the proposed RS using a depression and stress detection by machine learning using CNN, BLSTM-RNN model and the eSM2 sentiment metric for mood assessment. In Section IV, the experimental results are presented. Section V presents the discussions; finally, Section VI presents the conclusions and outlines the future work.

II. Related Work

A. Sentiment and Affective Analysis

The sentiment analysis helps industries to formulate marketing strategies, support after-sale services [29], develop health monitoring system, RS [3], among other services.

Sentiment analysis can be performed by: (i) machine learning [30]; (ii) lexicon-based technique using a word-dictionary of textual information or corpus-based approach, in which the polarity value is computed based on the occurrences of the terms in the corpus; (iii) a hybrid technique, which combines machine learning and word-dictionary approaches.

The machine learning approach needs a large number of data to obtain reliable results from sentiments; for instance, Chen et al. [31] performs the machine learning approach with a neural network model using BiLSTM-CRF and CNN using 14,492 sentences in the training phase.

The lexicon-based technique uses an intensity scale with emotional words; examples of word-dictionaries are WordNet, Sentimeter-Br2 [32], [33] and eSM [3]. This approach is used in this research to perform sentiment analysis.

Sentimeter-Br2 is a word-dictionary with its respective sentiment intensity (positive or negative words), considering n-grams, verbal tenses and adverbs. The sentiment intensity value of an S-sentence, using the Sentimeter-Br2, is calculated by (1):

$$Sentimeter\_Br2(S) = \frac{SU + SB + ST}{k + p + q + r} \quad (1)$$

where: Sentimeter-Br2(S): result of the global sentiment intensity of the S-sentence; k is related to the sentence tense, k = 1, if the sentence has a verb in the past participle; and k = 0 if the sentence is in another tense or the sentence does not have a verb; p is the total number of unigrams in the F-sentence, with the exception of words with no sentimental intensity value (stopwords); q is the total number of bigrams; r is the total number of trigram; ST represents the sentiment score of a trigram; SB is the sentiment score of a bigram and SU represents the sentiment score of an unigram.

The eSM is an enhanced sentiment metric based on Sentimeter-Br2 that additionally considers age, gender and educational level, which are obtained from the user profile. The sentiment intensity determination of an S-sentence using eSM is given by (2). This relation was obtained from subjective test results, in which each person posted sentences on social networks. Then, these sentences were scored by both the same person who posted the sentences and the eSM relation. In the

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

eSM formulation can be observed the relation between the sentiment intensity and the user profile characteristics.

$$eSM(S) = Sentimeter\_Br2(S) * C * exp(a_1 * A_1 ... \\ + a_n * A_n + g_1 * M + g_2 * F \\ + e_1 * G + e_2 * nG) \quad (2)$$

where: C represents a scale constant; a1...an represents binary factors related to age ranges, if one of them is equal to one, the others are zeros; A1...An are the weight factors of each age range, considering four ranges; g1 and g2 are binary factors related to the gender; M and F are the weight factors of gender, man or woman, respectively; e1 e e2 represents binary factors related to educational level (higher education or not); G and nG are the weight factors of educational level, higher education or not, respectively.

In [34], the authors show that teenagers behave differently in blogs, observing some peculiarities in the writing style. In our study, the parameters are extended to be used in a sentiment metric; specifically, the geographic location of the user and the theme of the sentence captured on an OSN.

Sentiment analysis is related to positive, negative or neutral classification. The affective analysis is not limited to three-sentiment polarities because it considers the different emotions, such as sadness and anger; although they have the same sentiment polarity (negative), the emotions are totally different. The emoticons, icons for representing emotions [35], and expressions such as “LOL” (laughing out loud) represent a sentiment and affective meaning, which are commonly found in sentences extracted from OSN. Thus, affective analysis is used for measuring the person’s emotion.

In the sentiment and affective analysis, there are some points to be explored, such as if the user profile influences the sentiment metric performance, which characteristics must be considered and how to perform the association between the user’s profile and the sentiment metric. For instance, the sentiment intensity determined by a metric can change depending on the gender [36].

It is worth noting that, currently, scarce studies about lexicon-based metrics take into account profile parameters. In this research, our proposed sentiment metric, eSM2 complements the eSM by considering the user’ geographic location and the theme of the sentence.

B. Recommendation System

RS predicts useful items for the user, considering what the user may be interested in. For this prediction, some data are extracted, for example, user’s profile, user’s preferences and past behavior [37].

There are commonly three RS approaches: content-based, collaborative filtering and hybrid-based. The content-based approach works with the description of an item and the profile of the user’s preference; the suggestion of items is based on what the user already liked. The collaborative filtering analyzes the user’s behavior and preferences and explores similar preferences among people [38]. The hybrid approach combines both methods.

Traditional RSs are developed based only on words searched by the user [39] on the Internet, but many of these words could be searched too far back, and the search would not therefore represent current information.

Sentiment analysis began to be explored in RS [40] to suggest more updated contents and based on the person’s mood. The semantic technique is based on some knowledge base defined as ontology, considering the user’s opinions to complete the lack of information, through inferences [41]. As stated before, a KBRS uses a knowledge base and offers many benefits [41], one of which is the treatment of the cold-start problem.

This paper proposes a KBRS that incorporates a sentiment metric, which is the difference from the system aforementioned. For performance comparison purposes, tests were performed using both a content-based traditional RS and the proposed KBRS.

C. Emotional Health Monitoring System

Emotional disorder problems need continuous monitoring; more specifically, depression and stress disorders. Commonly, due to the unpredictable behavior of these disorders, the mood assessment is captured by traditional standard procedures through rating scales and questionnaires [42]. Also, there are solutions to treat human body signals, such as, PSYCHE [43], which is a sentiment analysis solution, based on portable sensing devices. The voice signal has also been studied for emotional assessment [44], along with attitudinal indicators, such as sleep quality, galvanic skin response, activity, and gesture.

Studies in psychiatry [45]–[47] found that linguistic styles can indicate depression disorders. Nguyen et al. [48] use the lexicon approach and the most used words in a group of depressive users are detected. Anger, according to [6], [49] indicates a negative emotion, and it can represent stress disorders. These studies perform the textual and linguistic features analysis, but they are not explored in a health monitoring system. The advantage of the textual analysis is that this technique does not need special and specific equipment; thus, it is a cheaper solution.

The mental health of a person can be reflected on his/her mood. There are applications to improve the user’s optimism [50], which log self-reported mood, in which the mood charts have been recommended by psychiatrists and therapists, and clients monitor their own mental health. Projects for depression detection have been implemented by Wang et al. [51] reaching an accuracy around 80%, through a model based on sentiment analysis in micro-blogs. However, the research by Wang et al. [51] and [52] focus only on depression monitoring and an RS is not implemented.

As stated before, OSNs have an enormous volume of data that can be used for monitoring the users’ mental health. A system to provide continuous emotional monitoring at home, job, or school/university is a hard task; therefore, OSN is a way of capturing the users’ emotions in these environments. One of the great challenges of related works is to propose a solution that gets a high accuracy for the emotional disorder detection. In this work, it is intended to confront this problem.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

D. Machine Learning and Deep Learning Approach

The ML is a useful technique to conduct classification [53], statistical analysis, feature selection, and data normalization. ML classifies problems in general, including affective analysis and the mood detection problem. This technique uses algorithms to perform a supervised or unsupervised method.

The word2vec [54] is a popular approach for capturing words to send to an ML algorithm. It allows modeling words as vectors, and it is based on the skip-gram and Continuous Bag-of-Words (CBOW) model to compute the distributed representations of words. The CBOW model predicts a word in a context and the skip-gram finds the context given by a word. The work described in Vo and Zhang [55] presents a sentiment-driven and a standard embedding associated with a variety of pooling functions, to extract the sentiment of Twitter comments.

ML is frequently used to perform emotion classification [56], which are based on six different classes, representing the 'Ekman' model of emotion [57]: anger, disgust, fear, joy, sadness, and surprise. Stress and depressive messages are detected by negative sentences that contain one or more of these emotion classes.

The SVM algorithm has been widely used for emotion classification tasks, with good generalization properties [58]. SVM is used for classification with several feature selection techniques in the depression detection context [59]. The SMO algorithm is used to train SVM; SMO was the most accurate for predicting depression among senior citizens [60].

Random Forest and Naïve Bayes algorithms are also used for sentiment analysis, for detecting psychological disturbances on OSN; the results present an accuracy of about 0.8 [16].

Deep learning algorithms have been used to extract sentiment features in conjunction with semantic features [25].

Recent studies using deep CNNs [61], [62] presented significant performance improvement in natural language processing (NLP) tasks. Collobert et al. [61] used words embedding into a CNN to solve NLP problems, such as part-of-speech (POS) tagging and semantic labeling.

Problems with NER, which involves noun phrases identification in classes, such as the solution used for the Vietnamese language [63] uses a deep learning method composed of the combination of BLSTM, CNN, and CRF model, reaching a F1-score of 88.59%. BLSTM works with labeling tasks considering past and future contexts.

In this work, CNNs perform the character-level representations to feed a BLSTM based on RNN to perform the entity representations. Another BLSTM-RNN performs the relation between entities, which is stacked on the first BLSTM-RNN. The main goal of the study is to improve the classification accuracy of sentences posted by users that present depression and stress disorders.

III. Methodology

This section presents the methodology to build the proposed KBRS solution. The first step in this research was to conduct subjective tests to help determine the model of the proposed KBRS. Later, each component of the KBRS is explained.

A. Subjective Tests

Subjective tests were performed in a laboratory environment to determine the eSM2 metric parameters. The tests helped to establish user's preferences regarding the kind of messages they would like to receive, and to evaluate the KBRS resources consumed in the electronic device. Finally, a remote method was used to validate the performance of the monitoring system and the RS.

Assessors were selected to answer questions about their emotional state and write sentences on the social network Facebook. Some of the users were diagnosed with problems of acute stress and mild to moderate depression level, according to their clinical history.

The sentences were remotely extracted from an OSN with negative and positive nouns, adjectives, and verbs. The sentences were analyzed by the machine learning algorithms, including the CNN, BLSTM-RNN model for classification of sentences with depressive, stress, and non-depressive and non-stress content. It is important to highlight that the assessors were instructed to write sentences on the OSN, if he/she was feeling motivated to do so, simulating a real situation of writing posts in their daily routine.

The tests with the assessors to be used into the sentiment metric were performed in two phases, face-to-face and remote subjective tests. The face-to-face data collection was conducted in a laboratory environment to find which parameters of the user's profile could affect the sentiment intensity value of a sentence. The collection task also helped to evaluate the performance of the monitoring system. The remote subjective method was performed to validate the proposed sentiment metric. Thus, the initial task in a laboratory helped to prove the initial hypothesis, and the remote method validated it.

The data collection process was performed by Portuguese native-speaking assessors. The 146 assessors, comprising 74 men and 72 women with ages ranging from 18 to 43 years old, with different profiles, such as region of birth (north, south, and southeast of Brazil), educational level, among other characteristics, presented in Table I. The ages were separated into three groups (A1, A2, and A3) and the sentences were separated into three themes (entertainment, work or study, and family).

A web interface questionnaire was presented to assessors with a question about which words could identify their emotional state, in stressful conditions or at moments of depression. Furthermore, another question was presented about what kind of message the user would prefer in depression and stress conditions (happy, calm, relaxing or motivational messages); out of which the person could choose one or two kinds of messages. The kinds of messages chosen by the users are used in the RS. Additionally, users indicate the number of messages and the period of day when they would like to receive the messages.

The questionnaire helped to know what person's characteristics could affect a sentiment metric. Later, the assessors wrote sentences on OSN, which were captured by a script. The sentences were classified by both the person who wrote the sentences and by the Sentimeter-Br2 metric; each sentence was classified on a sentiment scale from -5 to +5.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

TABLE I
PROFILES ANALYZED ON SOCIAL NETWORK AND RESPECTIVE VALUES

Parameter	Values
Gender	man, woman
Age	(ranges) A1: 18 – 26; A2:27 -35; and A3: 36 - 43
Themes	entertainment, work/studies and family
Geographic location	North, south and southeast
Educational Level	higher education, bachelor or equivalent, MSc or PhD

In the second phase, the assessors' messages were remotely monitored over a 5-week test period; all the sentences captured were also analyzed by both the Sentimeter-Br2 sentiment metric and by the assessor who posted the sentences. At the end of the 5-week period, the assessors came back to the laboratory to test the performance of the KBRS in a standard mobile phone. In total, 27,308 sentences were extracted from the OSN and evaluated by the sentiment metric. A sentiment correction factor based on the user's profile was modeled using the subjective test results. This modeled correction factor can be applied to traditional sentiment metrics.

The update of the the Sentimeter-Br2 dictionary, considering new slang and expressions, was performed by specialists, who added new words and their respective scores.

B. The Proposed Knowledge-Based Recommendation System

The KBRS contains the emotional health monitoring system, which uses the deep learning model and the sentiment metric named eSM2. A high-level view of the proposed KBRS architecture is shown in Fig. 1.

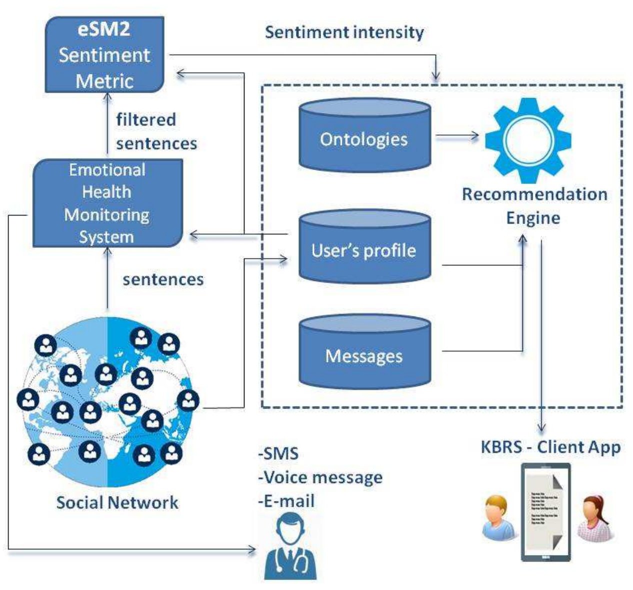

Fig. 1. Architecture of the proposed KBRS, which considers sentences extracted from social networks.

The sentences are extracted from an OSN, as shown in Fig. 1. The emotional health monitoring system identifies which sentences present a stress or depression content using machine learning algorithms and the emotion of the sentence content. The monitoring system is able to send warning message to people that are previously registered. Later, the selected sentences are analyzed by the sentiment metric (eSM2) and the sentiment intensity is used as input of the recommendation engine. The KBRS server establishes a communication with the KBRS client application, in which the user receives a specific message according to his/her profile, ontology aspects, and the sentiment value calculated from his/her sentences extracted from OSNs.

The system consists of the following components:

- User profile and user data: database built from the data captured from OSNs.
- Messages: there is a database with 360 messages, 90 messages for each kind (relaxing, motivational, happy, or calm messages) to be suggested to the user by the recommendation engine. The users can previously choose one or two kinds of messages when they undergo a period of stress or depression. The messages were written by 3 Specialists in psychology and validated by 3 other Specialists.
- Depression or stress detection by machine learning: the sentences are extracted from OSN and they are filtered by machine learning to detect depression or stress conditions. It is implemented in the emotional health monitoring system.
- Sentiment analysis by eSM2: the sentences are filtered and scored by the eSM2 sentiment metric, from -5 to +5. This range was tested and validated in previous studies [3], [64]. The sentiment intensity of the sentence will determine the message intensity. There are three levels of messages: extreme, intermediate and lower. If the monitoring system detects that the user is very/ deeply stressed, a very positive message is sent to him or her. The sentiment intensity range and the respective message intensity levels are presented in Table II; the message intensity levels were determined according to the users'opinions. Examples of very positive messages include intensity adverbs, such as much, very, strongly, among others. Also, the kind of message corresponds to the preference defined in the tests in the first phase.
- Ontologies: mechanism of classes, objects, and relationship used for the recommendation engine. The ontologies are expressed by the Ontology Web Language (OWL). The data of each class can be extracted from OSN. The Nuadu ontology collection is used for health scenarios. In this paper, the following Nuadu's classes were used:

  - Personal ontology: describes the personal information stored for each individual, including gender, work, studies, and other preferences.

  - Activity ontology: describes information about the person's activities. The entries may indicate changes in the routine of the person.

  - Sleep ontology: describes users' routines and schedules.

  - Risk ontology: describes information about smoking and alcohol consumption; these habits reflect bad

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

habits showing a tendency to greater stress and disease development.

- Context ontology: describes the environment of the person (home, study, work, or travel). This information is important because it can explain a period of missing activity entries or change in sleeping times.

- Recommendation engine: mechanism responsible for generating a list of recommendations.

In the proposed system, users personal information and context information on Facebook is used. However, users do not always post this related information on their Facebook account. In case users do not post personal information, a standard information is used, such as sleep routine of 8 hours, no unhealthy habits, no preferences about work or study. It is important to note that in our tests only 5% of the users do not post this information.

A traditional RS is also implemented, in which only the words searched by a person on OSN are used to feed the system, forming a content-based RS. For the sake of simplicity, the traditional content-based RS will not be explained in this section.

TABLE II
SENTIMENT INTENSITY RANGE AND RESPECTIVE MESSAGE INTENSITY LEVEL

Sentiment Intensity Range	Message Intensity Level
-5.0 to -3.0	An extreme level of positive message
-2.9 to -0.1	An intermediate level of positive message
0.0 to + 5.0	A lower level of positive message

C. Emotional Health Monitoring System by Machine Learning using the BLSTM-RNN model

In the emotional health monitoring system, user's sentences are extracted from an OSN and filtered through the machine learning approach. In case a stress or depression disorder is identified the KBRS is activated.

In the solution, the main characteristics of depression or stress are short texts, negative emotions, low values of sentiment intensity and use of the first person pronoun. In case depressive sentences are detected by machine learning, warning messages are sent by different options (voice message or e-mail). These messages are only sent to authorized people that were previously registered in the system.

The machine learning approach was used to identify sentences with depressive, stress, and non-depressive and non-stress content extracted from an OSN. These sentences were filtered, in which expressions and emoticons were available in the models for detecting stress and depression expressions, such as “hate my life”, “feeling sad”, “I am stressed”, among others. Positive emotional sentences were also filtered to improve the classification of depression versus non-depression sentences, thereby decreasing the false positive for depression detection. In order to distinguish depression and stress conditions, the emotion of the sentence content is considered. Thus, sentences related to anger, disgust and surprise emotions are associated to the stress [6], [49]; and depression is associated to fear and sadness emotions [48].

The dataset used to classify stress, depression, and non-stress and non-depression expressions, in the training phase, was built using sentences written by 146 assessors on an OSN. In total, 27,308 labeled Facebook messages were used, of which 23.700% and 26.197% of the messages correspond to depression and stress sentences, respectively. And, 50.103% messages are related to non-stress and non-depression sentences. It is worth noting that in the testing phase, additional 146 assessors participated.

The CNN with BLSTM-RNN model, SMO, Random Forest and Naïve Bayesian classification were used in this work. Studies [65] present good results using approaches such as Hidden Markov Model (HMM) and Gaussian Mixture Model (GMM). However, preliminary tests of our work presented low values of accuracy and F1-score using this approach. Then, they were not used in further tests.

This research used the Theano library [66] for deep learning architecture implementation and other algorithms. In all the experiments, 10-fold cross-validation was used to test and to evaluate the accuracy of the stress or depression detection by machine learning techniques. The classification was performed in a binary attribute, as depressed/non-depressed and stresses/non-stressed sentences.

In the deep learning architecture, the CNN computes the character-level representation with characters serving as inputs. The convolutional kernel of the CNN performs the convolutions for the characters of the words; for each convolution i, the kernel output ko is performed, as shown in (3).

$$ko_i = htaf(M_i r_{ci} + b_i) \tag{3}$$

where parameter M_i is the parameter matrix and b_i is the learned bias vector; htaf represents the hyperbolic tangent activation function and r_{ci} represents the character-level representation of word i.

In HTAF, each network layer presents a bias node that is connected to all other nodes. The HTAF is used for the hidden and output layers for calculating the backpropagated error signal of the deep learning architecture.

The Softmax layer is responsible for finding the probability P of the relation labels, according to (4). The hidden activation layer h is derived from the input characters.

$$P = softmax(M_i h + b_i) \tag{4}$$

Another test was performed using the SVM classifier instead of the Softmax function for performance comparison purposes.

Fig. 2 presents the neural network model, which used the output vectors of BLSTM to feed the HTAF layer; the character-level representation, along with word embedding vector, served to feed the BLSTM-RNNs. The output vectors of BLSTM are sent to the DE layer to choose the label sequence; the DE represents the disease extraction (stress or depression) by the Softmax output layer.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

The hidden states *ha, hb* account for capturing information in next steps in direct and reverse directions. The LSTM output performs the bottom-up (↑ h) and top-down (↓ h) computation, the bottom-up captures the information in the current step and in the previous steps in the neural network model and the top-down computation calculates the information using the reversed inputs. The direct and reverse directions are performed until reach the disease extraction. An example of sentence captured is "'imposition causes stress"', in which the subject of the sentence (imposition) represents a1 and the object of the sentence (stress) represents b1; the y parameter represents the input vector in a LSTM unit, in which y1 is the first word, y2 is the second word and y3 is the third word. The DE layer generates the label stress/non stress in the example of Figure 2.

The tests used a batch size of 10, momentum equal to 0.8, a learning rate of 0.01, 50 epochs and dropout rate of 0.5. These values were chosen after experiments with different values until the best results were reached.

The F1-score, accuracy and Precision Recall Area Under the Curve (PR AUC)[67] were used as performance metrics to compare the results obtained by different machine learning algorithms. F1-score is the harmonic mean between recall and precision. The PR-AUC is commonly used in cases of unbalanced data.

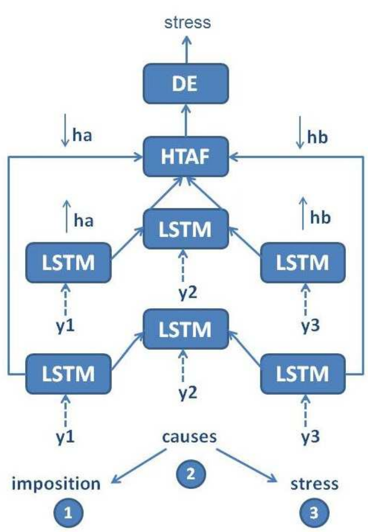

Fig. 2. The mechanism of the BLSTM-RNN method for classifying the relation of the sentence "'imposition causes stress"'.

D. eSM2 Sentiment Metric

The correction factor used herein depends on the age range, gender, theme, educational level, and geographic location of the person, besides the theme of the sentence. These data were obtained from 146 volunteer assessors. They could also provide this information on the application interface. Therefore, the user had the right to make their data available or not, in accordance to the privacy policies. Also, 29 assessors were identified with a depressive profile, and only these assessors were requested to send contacts of authorized people who will receive warning messages.

The correction factor to be applied to sentiment analyses based on a user profile is necessary because a sentence can be scored with different sentiment values, depending on the user's characteristic and theme. The parameters age, gender, educational level, and geographic location were considered because they were identified in preliminary tests as the most impacting factors on the global score obtained by a sentiment metric. Also, these factors can be extracted from OSN, using the programming languages PHP and AJAX. The theme is identified by an automatic script based on keywords.

The novel proposal of sentiment intensity metric, named eSM2, considering an S-sentence, is introduced in (5). The eSM2 is based on Sentimeter-Br2 and a correction factor that uses different parameters related to user profile. A second and a third order polynomial function were also tested, but exponential function obtained the best performance.

$$\begin{aligned} eSM2(S) = Sentimeter\_Br2(S) * C * exp(a_1 * A_1 ... \\ +a_n * A_n + g_1 * M + g_2 * F + e_1 * E_1 \\ +... + e_n * E_n + t_1 * T_1 + ... + t_n * T_n \\ +l_1 * L_2 + ... + l_n * L_n) \end{aligned} \tag{5}$$

where:

- C is a scale constant;
- a_1... a_n are binary factors related to age ranges; if one of them is equal to one, the others are zeros. A_1...A_n are the weight factors of each age range; this work considered three ranges;
- g_1 and g_2 are binary factors related to the gender; if one of them is equal to one, the other is zero, and M and F are the weight factors of gender, man or woman, respectively;
- t_1...t_n are binary factors related to themes, if one of them is equal to one, the others are zeros. T_1...T_n are the weight factors of each theme. This work considered three themes (entertainment, work/studies, and family);
- l_1...l_n are binary factors related to geographic location; if one of them is equal to one, the others are zeros. L_1...L_n are the weight factors of each geographic location, this work considered three locations in Brazil (north, south, and southeast);
- e_1...e_n are binary factors related to educational level; if one of them is equal to one, the others are zeros; E_1...E_n are the weight factors of each educational level. This work considered three educational levels (higher education, bachelor or equivalent, master or superior), these levels were chosen according to UNESCO [68] educational levels.

Each assessor evaluated the sentiment value of 20 sentences in the laboratory to measure the performance results of the eSM2. With this information, each evaluated sentence represents an equation with 15 unknown variables; thus, in total, there are 2920 equations and an overdetermined system is obtained. Firstly, to solve this problem, each equation is linearized, then the least square method, specifically, a pseudo inverse method is used.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

The following statistical functions were considered for performance assessment of (5), the root mean square error (RMSE), the maximum and average error.

E. The KBRS client-server architecture

The KBRS client and server applications are described as follows.

An application is installed on the mobile device using the client-server architecture through a web application in Hypertext Preprocessor (PHP), JavaScript Object Notation (JSON) and HTML5.

As stated before, the first step towards using the emotional health monitoring system and the KBRS is to access the application interface running on the user device and to grant the necessary permissions for the application to extract the user's profile information.

On the other hand, a Web server stores several information such as the user profile information and the messages databases. Also, different tasks are performed in the Web server, such as the recommendation engine, the sentiment analysis metric, the emotion estimation of each sentence analyzed. It is important to note that the user's device is not overloaded because this device is solely used to receive messages.

The sentences and user profile are extracted from the Facebook social network by an automatic script, periodically. The depression or stress detection is performed by machine learning, and the filtered sentences are scored by the eSM2 sentiment metric to find the sentiment intensity.

Four types of messages, which can vary in sentiment intensity level and in ontology use, are sent to the users' application, considering the kind of messages chosen by each user. In the first type of message, the KBRS uses sentiment analysis with the eSM2 metric; in the second one, the KBRS uses a sentiment metric and no ontology; the third message does not contain the sentiment metric but it contains the ontology. The last message does not contain sentiment metric and ontology. If the user does not write a sentence on an OSN, a default positive message is sent to the user's application client-side.

IV. EXPERIMENTAL RESULTS

This section describes the experiments results regarding the performance evaluation of the emotional health monitoring system, the definition of the eSM2 sentiment metric parameters and the performance evaluation of the proposed KBRS.

A. Performance Evaluation of the Emotional Health Monitoring System

As stated before, $27,308$ labeled messages were used in the training phase. Preliminary test results showed that using 75% of these messages presented similar classification accuracy to using 100% of them; thus, it is demonstrated that the data used is sufficient to train the deep learning model. In this phase, the CNN BLSTM-RNN using SoftMax obtained the best performance considering all the performance assessment parameters.

In the experiments performed in the testing phase, additional 146 assessors participated in subjective tests to evaluate the proposed model. A total number of $25,192$ sentences were extracted from an OSN along 5 weeks.

Table III show the performance of the machine learning algorithms used in the testing phase. Specifically, Table III presents the F1-Score, accuracy and PR-AUC of depressed, stressed, and non-depressed and non-stressed sentences classification.

The CNN BLSTM-RNN using SoftMax had the best performance, according to Table III. The results presented an accuracy of 0.89, 0.90 and 0.93 to detect depressed, stressed, and non-stressed and non-depressed sentences, respectively.

The Softmax presented the best results because it is optimized for back-propagation network [69].

In order to evaluate the classifier methodology performance, participants were monitored, and after each day during the test period, they classified their mood with an emotion icon (emoticon).

In case of users with depressive profile, 100% of the participants considered useful to send a notification message to authorized persons. These messages were sent by e-mail and short-message services.

It is important to note that the results obtained in the present work reached a stress and depression classification accuracy higher than 88% overcoming the results obtained by similar works [11]–[17], [19].

B. Definition of the eSM2 Sentiment Metric Parameters

Fig. 3 presents the average weight values of each parameter considered in the eSM2 model and obtained from the experimental test results.

The results show that gender, age range, geographic location, and theme parameters are the most influential factors in a sentiment metric. It is important to note that the application extract user characteristics only if the user allows it. Also, the application does not extract characteristics, such as religion and race; therefore, ethical standards are not violated.

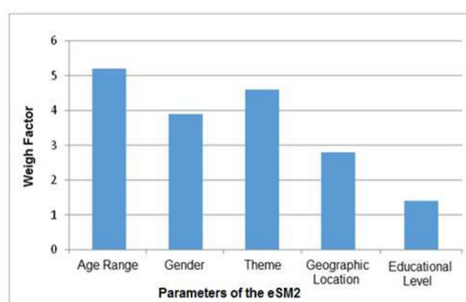

Fig. 3. Weight factors of the parameters used in the eSM2 sentiment metric obtained from subjective test results.

Table IV presents the performance assessment of eSM2 considering RMSE, Maximum Error and Average Error in relation to the sentiment intensity scored by the assessors in subjective face-to-face tests.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

TABLE III
F1-SCORE, ACCURACY AND PR-AUC VALUES FOR DEPRESSION, STRESS AND NON-DEPRESSION/NON-STRESSED SENTENCES DETECTION OBTAINED BY MACHINE LEARNING ALGORITHMS

Machine Learning Technique	F1-Score/Accuracy/PR-AUC for depressed users	F1-Score/Accuracy/PR-AUC for stressed users	F1-Score/Accuracy/PR-AUC for non-depressed and non-stressed sentences
CNN BLSTM-RNN using SoftMax	0.92/0.89/0.90	0.95/0.90/0.93	0.95/0.93/0.94
CNN BLSTM-RNN using SVM	0.90/0.87/0.88	0.91/0.88/0.89	0.92/0.88/0.91
SMO	0.83/0.84/0.82	0.84/0.85/0.83	0.85/0.85/0.83
Random Forest	0.79/0.81/0.78	0.80/0.82/0.79	0.82/0.83/0.80
Naïve Bayes	0.76/0.69/0.75	0.78/0.79/0.77	0.79/0.80/0.72

TABLE IV
PERFORMANCE ASSESSMENT OF ESM2 CONSIDERING RMSE, MAXIMUM ERROR AND AVERAGE ERROR IN RELATION TO THE SENTIMENT INTENSITY SCORED BY THE ASSESSORS

Parameter	eSM2
RMSE	0.21
Maximum Error	0.34
Average Error	0.27

C. Performance Evaluation of the Proposed Recommendation System

The application of KBRS at the end-user device was tested in the laboratory environment, using a mobile device with a Wireless interface (Wi-Fi). The application was built to be performed in an equipment with a 1400Mhz 32-bit Quad-Core processor, 1GB RAM memory, in which only 1 core of the processor is used to execute KBRS. A 5-point scale average value of the parameters was used from 1 to 5, in which value 5 represents the best results and value 1 represents the worst punctuation.

Table V presents the results of the evaluation of the recommendation message latency, the energy consumption by KBRS (considering ontology and eSM2) and the apparent network resource consumption, which tests were performed in the exclusive and standardized equipment used by the assessors in the laboratory.

TABLE V
PERCEIVED VALUE OF CONSUMED RESOURCES IN THE ELECTRONIC DEVICE AND RESPECTIVE EVALUATION IN ACCORDANCE WITH A 5-POINT SCALE

Performance Regarding the Parameters	Perceived Average Value
Latency of the KBRS in general	4.2
Energy consumption of the KBRS application	4.2
Apparent network resource consumption	4.6

Ergonomic aspects were also analyzed by the assessors, who answered the following questions, using a scale from 1 to 6, in which 1 means very unsatisfied, 5 very satisfied and 6 means no opinion.

- How would you rate the time needed to get recommendations?
- How would you rate the appearance and interface?
- How would you rate the variety of the proposed messages?
- How would you rate the general usability?

The results of this analysis indicated that 92% of the assessors liked the variety of the proposed messages and the interface of the application running on the user device. Finally, 89% of the assessors classified the general usability of KBRS with the maximum score.

As stated before, in the data collection process, the users chose one or two kinds of messages (happy, calm, relaxing, or motivational messages) that they would prefer to receive in depression or stress conditions. These messages are used by KBRS. Additionally, users indicate the number of messages and the period of day when they would like to receive the messages.

Table VI presents the results considering the KBRS, which is based on ontology, whereas the traditional content-based RS, named 'no ontology' is used with the eSM2 and without the eSM2. The results from KBRS using the eSM2 metric and ontology reached 94% of satisfied users.

The results of Table VI are related to the users' satisfaction level with message suggestions received on the user's device. The answer options comply to a scale based on adjectives, which are: very good, good, neutral, poor, and very poor.

TABLE VI
PERCENTAGE OF PERCEIVED VALUE OF FOUR KINDS OF RECOMMENDATION SYSTEM CONSIDERING AND NOT CONSIDERING THE ESM2 SENTIMENT METRIC AND CONSIDERING AND NOT CONSIDERING THE ONTOLOGY IN THE RECOMMENDATION SYSTEM

	Messages with eSM2 and ontology	Messages with eSM2 and no ontology	Messages without eSM2 and with ontology	Messages without eSM2 and no ontology
Very good	94%	89%	81%	69%
Good	3%	8%	8%	4%
Neutral	3%	3%	11%	2%
Poor	0%	0%	0%	21%
Very poor	0%	0%	0%	4%

In order to compare the eSM and eSM2 performances, experiments that considered the eSM were also performed. Table VII presents the results of eSM metric, considering the same test scenarios presented in Table VI. The results from KBRS using the eSM metric and ontology reached 89% of satisfied users.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

TABLE VII
PERCENTAGE OF PERCEIVED VALUE OF FOUR KIND OF RECOMMENDATION SYSTEM CONSIDERING AND NOT CONSIDERING THE ESM SENTIMENT METRIC AND CONSIDERING AND NOT CONSIDERING THE ONTOLOGY IN THE RECOMMENDATION SYSTEM

	Messages with eSM and ontology	Messages with eSM and no ontology	Messages without eSM and with ontology	Messages without eSM and no ontology
Very good	89%	80%	74%	63%
Good	5%	4%	12%	2%
Neutral	2%	3%	4%	1%
Poor	2%	8%	8%	18%
Very poor	2%	5%	2%	16%

V. CONCLUSIONS AND FUTURE WORK

In order to improve the KBRS performance, the eSM2 was modeled, considering user profile parameters, geographical location and the theme of the sentence to identify the sentiment intensity of a message. These two parameters are not considered in current sentiment metrics. The performance assessment of eSM and eSM2 metrics was performed, and results obtained by the eSM2 were superior in the perceptual evaluation of the RS. This fact demonstrated the relevance of using additional user profile parameters to improve the sentiment metric performance.

Also, the ontology concept was used in the proposed KBRS. It is important to note that the correction factor proposed in eSM2, based on the user's profile, can be applied to other sentiment metrics.

Currently, there are few works that use OSN data to detect stress conditions. The solution for monitoring the depressed or stressed condition in OSN users, using the CNN for character-level representation, and the BLSTM-RNN for the disorder entity recognition, presented an accuracy for depression and stress detection of 0.89 and 0.90, respectively. These accuracy values are higher than the results obtained in related works.

In the performance assessment tests, the proposed KBRS was compared to another KBRS that does not consider a sentiment metric and ontology. Results demonstrate that the proposed KBRS overcomes the RS without sentiment metric and ontology, reaching 94% and 69% of very satisfied users, respectively. According to the users, an RS that does not consider ontology and a sentiment metric performs a very poor suggestion, using a more generic and not personalized content. The best result obtained by the KBRS proves the effectiveness of the use of ontology and especially the use of a personalized sentiment analysis instead of a general sentiment analysis. In general, the recommended messages sent to the appropriated users improved their emotional state, and this is the most significant contribution of this research.

Furthermore, the proposed KBRS uses an application at end-user devices which is not based on a complex programming language, thus consuming fewer resources from current electronic devices. Also, the client device interface is easy to use.

Nevertheless, considering the satisfactory results obtained by the proposed KBRS, in future works, it can be applied in other services, such as customer complaint systems and user help systems to detect abrupt changes of customers' emotion.

REFERENCES

[1] I.-R. Glavan, A. Mirica, and B. Firtescu, "The use of social media for communication." *Official Statistics at European Level. Romanian Statistical Review*, vol. 4, pp. 37–48, Dec. 2016.

[2] M. Al-Qurishi, M. S. Hossain, M. Alrubaian, S. M. M. Rahman, and A. Alamri, "Leveraging analysis of user behavior to identify malicious activities in large-scale social networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 799–813, Feb 2018.

[3] R. L. Rosa, D. Z. Rodríguez, and G. Bressan, "Music recommendation system based on user's sentiments extracted from social networks," *IEEE Transactions on Consumer Electronics*, vol. 61, no. 3, pp. 359–367, Oct 2015.

[4] R. Rosa, D. Rodr, G. Schwartz, I. de Campos Ribeiro, G. Bressan *et al.*, "Monitoring system for potential users with depression using sentiment analysis," in *2016 IEEE International Conference on Consumer Electronics (ICCE)*. Sao Paulo, Brazil: IEEE, Jan 2016, pp. 381–382.

[5] I. B. Weiner and R. L. Greene, "Handbook of personality assessment," in *John Wiley and Sons*, N.J, EUA, 2008.

[6] H. Lin, J. Jia, J. Qiu, Y. Zhang, G. Shen, L. Xie, J. Tang, L. Feng, and T. S. Chua, "Detecting stress based on social interactions in social networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1820–1833, Sept 2017.

[7] J. T. Hancock, K. Gee, K. Ciaccio, and J. M.-H. Lin, "I'm sad you're sad: Emotional contagion in cmc," in *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, 2008, pp. 295–298.

[8] B. Liu, "Many facets of sentiment analysis, a practical guide to sentiment analysis," *Springer International Publishing*, pp. 11–39, Jan 2017.

[9] Y. P. Huang, T. Goh, and C. L. Liew, "Hunting suicide notes in web 2.0 - preliminary findings," in *Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007)*, Dec 2007, pp. 517–521.

[10] W. H. Organization, "World health statistics 2016: Monitoring health for the sdgs sustainable development goals, world health statistics annual," *World Health Organization*, p. 161, 2016.

[11] Y. Zhang, C. Xu, H. Li, K. Yang, J. Zhou, and X. Lin, "Healthdep: An efficient and secure deduplication scheme for cloud-assisted ehealth systems," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2018.

[12] G. Sannino, I. D. Falco, and G. D. Pietro, "A continuous non-invasive arterial pressure (cnap) approach for health 4.0 systems," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2018.

[13] H. Thapliyal, V. Khalus, and C. Labrado, "Stress detection and management: A survey of wearable smart health devices," *IEEE Consumer Electronics Magazine*, vol. 6, no. 4, pp. 64–69, Oct 2017.

[14] A. E. U. Berbano, H. N. V. Pengson, C. G. V. Razon, K. C. G. Tungcul, and S. V. Prado, "Classification of stress into emotional, mental, physical and no stress using electroencephalogram signal analysis," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Sept 2017, pp. 11–14.

[15] J. Ham, D. Cho, J. Oh, and B. Lee, "Discrimination of multiple stress levels in virtual reality environments using heart rate variability," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2017, pp. 3989–3992.

[16] Y. Xue, Q. Li, L. Jin, L. Feng, D. A. Clifton, and G. D. Clifford, "Detecting adolescent psychological pressures from micro-blog," in *Health Information Science*. Springer International Publishing, 2014, pp. 83–94.

[17] S. Tsugawa, Y. Kikuchi, F. Kishino, K. Nakajima, Y. Itoh, and H. Ohsaki, "Recognizing depression from twitter activity," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 3187–3196.

[18] M. De Choudhury, S. Counts, E. J. Horvitz, and A. Hoff, "Characterizing and predicting postpartum depression from shared facebook data," in *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work; Social Computing*, 2014, pp. 626–638.

[19] R. Rodrigues, R. das Dores, C. Camilo-Junior, and C. Rosa, "Sentihealth-cancer: A sentiment analysis tool to help detecting mood of patients in online social networks," *International Journal of Medical Informatics*, vol. 1, no. 85, pp. 80–95, 2016.

[20] X. Ma and E. Hovy, "End-to-end sequence labeling via bi-directional lstm-cnns-crf," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlin, Germany: Association for Computational Linguistics, August 2016, pp. 1064–1074.

[21] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural architectures for named entity recognition," in *Proceedings of the NAACL*, San Diego, EUA, May 2016.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

[22] M. Khodayar, O. Kaynak, and M. E. Khodayar, "Rough deep neural architecture for short-term wind speed forecasting," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 2770–2779, Dec 2017.

[23] N. Majumder, S. Poria, A. Gelbukh, and E. Cambria, "Deep learning-based document modeling for personality detection from text," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 74–79, Mar 2017.

[24] R. G. Guimarães, R. L. Rosa, D. D. Gaetano, D. Z. Rodríguez, and G. Bressan, "Age groups classification in social network using deep learning," *IEEE Access*, vol. 5, pp. 10 805–10 816, 2017.

[25] O. Araque, I. Corcuera-Platas, J. F. Sánchez-Rada, and C. A. Iglesias, "Enhancing deep learning sentiment analysis with ensemble techniques in social applications," *Expert Systems with Applications*, vol. 77, pp. 236 – 246, 2017.

[26] Y. Chen, M. L.-J. Yann, H. Davoudi, J. Choi, A. An, and Z. Mei, "Contrast pattern based collaborative behavior recommendation for life improvement," Jun 2017, pp. 106–118.

[27] H. Hu, A. Elkus, and L. Kerschberg, "A personal health recommender system incorporating personal health records, modular ontologies, and crowd-sourced data," in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Aug 2016, pp. 1027–1033.

[28] A. Sachinopoulou, J. Leppanen, H. Kaijanranta, and J. Lahteenmaki, "Ontology-based approach for managing personal health and wellness information," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2007, pp. 1802–1805.

[29] J. Tang, Y. Zhang, J. Sun, J. Rao, W. Yu, Y. Chen, and A. C. M. Fong, "Quantitative study of individual emotional states in social networks," *IEEE Transactions on Affective Computing*, vol. 3, no. 2, pp. 132–144, April 2012.

[30] S. Rendle, "Factorization machines with libfm," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, pp. 57:1–57:22, May 2012.

[31] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using bilstm-crf and CNN," *Expert Syst. Appl.*, vol. 72, pp. 221–230, 2017.

[32] R. L. Rosa, D. Z. Rodríguez, and G. Bressan, "Music recommendation system based on user's sentiments extracted from social networks," in *2015 IEEE International Conference on Consumer Electronics (ICCE)*, Jan 2015, pp. 383–384.

[33] R. L. Rosa, D. Z. Rodríguez, and G. Bressan, "Sentimeter-br: A new social web analysis metric to discover consumers' sentiment," in *IEEE International Symposium on Consumer Electronics (ISCE)*, Jun 2013, pp. 153–154.

[34] D. A. Huffaker and S. L. Calvert, "Gender, identity, and language use in teenage blogs," *Journal of Computer-Mediated Communication*, vol. 10, no. 2, 2005.

[35] A. Fraisse and Paroubek, "Twitter as a comparable corpus to build multilingual affective lexicons," in *Workshop on Building and Using Comparable Corpora*, Reykjavik, Iceland, May 2014, pp. 26–31.

[36] M. Thelwall, D. Wilkinson, and S. Uppal, "Data mining emotion in social network communication: Gender differences in myspace," *Journal of the American Society for Information Science and Technology*, vol. 61, no. 1, pp. 190–199, 2010.

[37] H. Garcia-Molina, G. Koutrika, and A. Parameswaran, "Information seeking: Convergence of search, recommendations, and advertising," *Commun. ACM*, vol. 54, no. 11, pp. 121–130, Nov. 2011.

[38] S. Łazaruk, J. Dzikowski, M. Kaczmarek, and W. Abramowicz, "Semantic web recommendation application," in *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)*, Sept 2012, pp. 1055–1062.

[39] I. Cantador, P. Castells, and A. Bellogin, "An enhanced semantic layer for hybrid recommender systems: Application to news recommendation," *Int. J. Semantic Web Inf. Syst.*, vol. 7, no. 1, pp. 44–78, 2011.

[40] A. C. M. Fong, B. Zhou, S. C. Hui, G. Y. Hong, and T. A. Do, "Web content recommender system based on consumer behavior modeling," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 2, pp. 962–969, May 2011.

[41] E. P. J. M. M. del Castillo; J. A. Delgado-López, "Analysis of the state of the topic," in *Hipertext.net*, no. 6, Aug 2008, pp. 3602–3605.

[42] H.-J. E. Harmon-Jones C, Bastian B, "The discrete emotions questionnaire: A new tool for measuring state self-reported emotions," in *PLoS ONE*, vol. 11, no. 8, 2016, pp. 295–298.

[43] R. Paradiso, A. M. Bianchi, K. Lau, and E. P. Scilingo, "Psyche: Personalised monitoring systems for care in mental health," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Aug 2010, pp. 3602–3605.

[44] K. R. Scherer, J. Sundberg, L. Tamarit, and G. L. Salomao, "Comparing the acoustic expression of emotion in the speaking and the singing voice," *Computer Speech and Language*, vol. 29, no. 1, pp. 218 – 235, 2015.

[45] R. Ramirez-Esparza, C. Chung, E. Kacewicz, and J. Pennebaker, *The psychology of word use in depression forums in english and in spanish: Testing two text analytic approaches*, 2008, pp. 102–108.

[46] S. Rude, E.-M. Gortner, and J. Pennebaker, "Language use of depressed and depression-vulnerable college students," *Cogn. Emotion*, vol. 18, no. 8, pp. 1121–1133, 2004.

[47] S. W. Stirman and J. Pennebaker, "Word use in the poetry of suicidal and nonsuicidal poets," *Psychosomatic Med*, vol. 63, no. 4, pp. 517–522, 2001.

[48] T. Nguyen, D. Phung, B. Dao, S. Venkatesh, and M. Berk, "Affective and content analysis of online depression communities," *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 217–226, July 2014.

[49] R. Fan, J. Zhao, Y. Chen, and K. Xu, "Anger is more influential than joy: Sentiment correlation in weibo," *CoRR*, vol. 1309.2402, 2013.

[50] E. Kanjo, L. Al-Husain, and A. Chamberlain, "Emotions in context: examining pervasive affective sensing systems, applications, and analyses," *Personal and Ubiquitous Computing*, vol. 19, no. 7, pp. 1197–1212, Oct 2015.

[51] X. Wang, C. Zhang, Y. Ji, L. Sun, L. Wu, and Z. Bao, "A depression detection model based on sentiment analysis in micro-blog social network," in *Trends and Applications in Knowledge Discovery and Data Mining*, J. Li, L. Cao, C. Wang, K. C. Tan, B. Liu, J. Pei, and V. S. Tseng, Eds., Berlin, Heidelberg, 2013, pp. 201–213.

[52] M. R. Morales and R. Levitan, "Speech vs. text: A comparative analysis of features for depression detection systems," in *2016 IEEE Spoken Language Technology Workshop (SLT)*, Dec 2016, pp. 136–143.

[53] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine learning for predictive maintenance: A multiple classifier approach," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 812–820, June 2015.

[54] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *CoRR*, pp. 1301–1310, Mar 2013.

[55] D.-T. Vo and Y. Zhang, "Target-dependent twitter sentiment classification with rich automatic features," in *Proceedings of the 24th International Conference on Artificial Intelligence*, 2015, pp. 1347–1353.

[56] Muljono, N. A. S. Winarsih, and C. Supriyanto, "Evaluation of classification methods for indonesian text emotion detection," in *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, Aug 2016, pp. 130–133.

[57] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, pp. 169–200, 1992.

[58] B. Schuller, A. Batliner, S. Steidl, and D. Seppi, "Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge," *Speech Commun.*, vol. 53, no. 9-10, pp. 1062–1087, Nov. 2011.

[59] S. Alghowinem, R. Goecke, M. Wagner, J. Epps, M. Hyett, G. Parker, and M. Breakspear, "Multimodal depression detection:fusion analysis of paralinguistic, head pose and eye gaze behaviors," *IEEE Transactions on Affective Computing*, vol. PP, no. 99, pp. 1–1, 2017.

[60] I. Bhakta and A. Sau, "Prediction of depression among senior citizens using machine learning classifiers," *International Journal of Computer Applications*, vol. 144, no. 7, pp. 11–16, 2016.

[61] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *J. Mach. Learn. Res.*, vol. 12, pp. 2493–2537, Nov. 2011.

[62] S. Poria, E. Cambria, and A. F. Gelbukh, "Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis," in *EMNLP*, 2015.

[63] T.-H. Pham and P. Le-Hong, "End-to-end recurrent neural network models for vietnamese named entity recognition: Word-level vs. character-level," in *Proceedings of The 15th International Conference of the Pacific Association for Computational Linguistics*, 2017.

[64] F. Arup Nielsen, "A new anew: Evaluation of a word list for sentiment analysis in microblogs," *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts*, pp. 93–98, Sep 2011.

[65] C. Quan and F. Ren, "Weighted high-order hidden markov models for compound emotions recognition in text," *Information Sciences*, vol. 329, pp. 581 – 596, 2016.

[66] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. WardeFarley, and Y. Bengio, "Theano: a CPU and GPU math expression compiler," *Proceedings of the Python for scientific computing conference*, vol. 4, no. 3, 2010.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2867174, IEEE Transactions on Industrial Informatics

[67] Y. Tang, Y. Q. Zhang, N. V. Chawla, and S. Krasser, "Svms modeling for highly imbalanced classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 1, pp. 281–288, Feb 2009.

[68] UNESCO, "International standard classification of education," *ISCED*, 2006.

[69] Y. L. Cun, B. Boser, J. S. Denker, R. E. Howard, W. Habbard, L. D. Jackel, and D. Henderson, "Advances in neural information processing systems 2," D. S. Touretzky, Ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990, ch. Handwritten Digit Recognition with a Back-propagation Network, pp. 396–404.

1551-3203 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.