



Towards emotion recognition from contextual information using machine learning

Martín G. Salido Ortega¹ · Luis-Felipe Rodríguez¹ · J. Octavio Gutierrez-Garcia²

Received: 4 March 2019 / Accepted: 5 September 2019 / Published online: 13 September 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Emotions influence cognitive processes that underlie human behavior. Whereas experiencing negative emotions may lead to develop psychological disorders, experiencing positive emotions may improve creative thinking and promote cooperative behavior. The importance of human emotions has led to the development of automatic emotion recognition systems based on analysis of speech waveforms, facial expressions, and physiological signals as well as text data mining. However, emotions are associated with a context (in which emotions are actually experienced), hence, this work focuses on emotion recognition from contextual information. In this paper, we present a study aimed to assess the feasibility of automatically recognizing emotions from individuals' contexts. In this study, 32 participants provided information using a mobile application about their emotions and the context (e.g., companions, activities, and locations) in which these emotions were experienced. We used machine learning techniques to build *individual* models, *general* models, and *gender-specific* models to automatically recognize emotions of participants. The empirical results show that individuals' emotions are highly related to their context and that automatic recognition of emotions in real-world situations is feasible by using contextual data.

Keywords Emotion recognition · Context-aware applications · Smart devices · Affective computing

1 Introduction

Emotions play a key role in the development of rational behavior in humans (Ekman and Davidson 1994). Evidence from human behavior studies on neuroscience and psychology indicates that the experience and expression of emotions are essential for individual's survival (Fredrickson 2001; Ekman and Davidson 1994). According to Bechara et al. (2000) and Loewenstein and Lerner (2003), emotions influence cognitive processes such as decision-making. In addition, studies focused on analyzing the neural substrates of cognitive and affective processes have demonstrated that

emotions influence cognitive processes underlying human behavior, e.g., perception, attention, planning, and memory, see (Ahn and Picard 2005; Bechara 2004; Phelps 2006). Particularly, emotions such as anxiety, stress, and anger may have a negative influence on individuals' behavior and well-being [although some studies provide evidence that some of these emotions may also have positive effects, see (Tagar et al. 2011)]. In general, these types of negative emotions increase the likelihood of contracting diseases and suffering heart attacks, in addition, these negative emotions may stimulate the development of psychological disorders such as depression (Rodríguez et al. 2011; Costa and Macedo 2012). Moreover, negative emotions narrow individuals' action repertoire and attentional focus and may lead to social behaviors related to aggression and violence (Fredrickson 2001). In contrast, it has been demonstrated that positive emotions (e.g., happiness, contentment, and joy) broaden individuals' action repertoire, expand attentional focus, and improve memory systems and creative thinking (Breazeal 2003; Fredrickson 2001). Regarding social interactions, positive emotions help to create and sustain social relationships and promote cooperative behavior. Thus, whether positive or

✉ Luis-Felipe Rodríguez
luis.rodriguez@itson.edu.mx

Martín G. Salido Ortega
marting_salido@hotmail.com

J. Octavio Gutierrez-Garcia
octavio.gutierrez@itam.mx

¹ Instituto Tecnológico de Sonora, 85000 Cd. Obregón, Sonora, Mexico

² ITAM, Río Hondo 1, 01080 Ciudad de México, Mexico

negative, the emotions we experience influence our decision making, social interactions, and in general, our wellbeing.

The recognition of individuals' emotions becomes useful to understand the factors that trigger positive and negative emotions and the consequent behavior. Particularly, recognizing individuals' negative emotions may help to devise potential intervention strategies that lead the individual from negative to more positive emotions. Individuals' emotions can be recognized from physiological signals (Seo et al. 2019), facial expressions (Wingenbach et al. 2018), keystroke dynamics (Kołakowska 2018), intonation of voice (Panda et al. 2019), among other factors (Brave and Nass 2007). The affective computing literature includes several works that propose mechanisms, tools, and models designed to recognize individuals' emotions based on these traditional emotion detection modalities (Calvo and D'Mello 2010; El Ayadi et al. 2011; Fasel and Luetttin 2003; Grünerbl et al. 2015; Deng et al. 2017; Yin et al. 2017; Giatsoglou et al. 2017; Wegrzyn et al. 2017; Kanjo et al. 2015). In particular, some studies have focused on the recognition of individuals' emotions by analyzing data collected through smart devices, including smartwatches, smartbands, and smartphones (Politou et al. 2017; Lee et al. 2012; LiKamWa et al. 2013; Kanjo et al. 2015).

Ptaszynski et al. (2010) emphasize the importance of the individual's context in the recognition of emotions by arguing that emotions cannot be perceived in real-world settings independently from the context in which emotions are experienced. In particular, the recognition of individuals' emotions in real situations, beyond controlled experiments, demands novel methods that take into account contextual data that is feasible to be collected by using, for example, people's smart devices. The collection of contextual data avoids obtrusive techniques used in traditional emotion detection methods, which usually require special equipment such as skin conductance and electrocardiogram sensors (Calvo and D'Mello 2010; Nalepa et al. 2019). These traditional emotion detection methods also face challenges associated with precision on face expression recognition, voice analysis, accurate measurements of physiological data, among others. These challenges increase when the emotion recognition is carried out in real-world environments as signals may be contaminated by environmental noise. Particularly, the use of contextual data for emotion recognition enables the implementation of techniques that rely on the collection of user-generated data from smart devices and the analysis of such data using a machine learning approach, which has proven useful to infer human behaviors, see for instance (LiKamWa et al. 2013).

Contextual information includes the location and activity of individuals as well as the people and devices the individuals interact with. Also, aspects such as the time of day and other physical elements including environmental

noise and temperature are considered to define individuals' context (Bellavista et al. 2012; Alegre et al. 2016). Experiments have shown that contextual elements perceived by humans using diverse sensory modalities (e.g., visual and olfactory) influence emotions (Croy et al. 2011; Royet et al. 2000). According to Ortony et al. (1990), individuals' emotions are influenced by environmental events and actions from other people. Ortony et al. (1990) also suggest that the characteristics of perceived objects impact on individuals' emotions. Further evidence shows that some emotion elicitors are associated with social and non-social environmental stimuli such as the approach of a strange versus the approach of a family member (Lewis 2008). Evidence also suggests that the experience of emotions is related to individuals' location such as their workplace and home (Ashkanasy and Daus 2002; Wharton and Erickson 1993; Sandstrom et al. 2017). In real-world applications most of these contextual data becomes useful for the recognition of emotions since these can be collected through sensors integrated in common smart devices such as smartphones.

In this paper, we present a study to evaluate the feasibility of automatically recognizing individuals' emotions from contextual information. We developed a mobile application that allowed the participants of the study to record experienced emotions each time their affective states changed. Regarding the contextual information, participants collected the following data: activity (e.g., studying or resting), thermal sensation (e.g., warm or cold), physical affliction (e.g., tired or hungry), location (e.g., university or home), company (e.g., classmates or family member), and the date and time when the participant recorded such contextual information and the experienced emotions. We used a machine learning approach to analyze the collected data and built (1) *individual* models (from data of the participant that provided the largest number of records), (2) *general* models (from data of all the participants), and (3) *gender-specific* models (grouping data by males and females) to validate the feasibility of automatically determining whether an individual experiences positive or negative emotions in a given context.

This paper is organized as follows. Section 2 analyzes related work. Section 3 presents a description of the study, the data collected, the participants, and the mobile application developed to collect individuals' emotions and contextual data. Section 4 describes the types and characteristics of the machine learning techniques used. Section 5 presents the results of the experiments carried out and Sect. 6 provides a discussion of the results and limitations of this work. Finally, Sect. 7 presents some concluding remarks and future research directions.

2 Related work

Data sources of automatic emotion recognition mechanisms range from speech waveforms (Deng et al. 2017), through physiological signals (Yin et al. 2017) and text data mining (Giatsoglou et al. 2017), to facial expressions (Wegrzyn et al. 2017). However, the present work focuses on emotion recognition from contextual information, hence the related work is focused on mechanisms that recognize emotions mostly from the situations in which emotions are experienced.

Among the first works to approach emotion recognition from contextual information is the work by Conati (2002) who proposed a probabilistic model based on dynamic decision networks to recognize emotions in the context of educational games. Conati's dynamic decision network takes into account the causes and effects of emotions as well as their temporal evolution in conjunction with players' bodily expressions, e.g., eyebrow position. Whereas Conati's model is supported by the OCC emotional model (proposed by Ortony et al. (1990)), her affective model has to be adapted to very specific scenarios where even the goals of the participants have to be defined beforehand. It should be noted that Conati's dynamic decision network was evaluated using a sample scenario without numeric performance metrics.

Oh et al. (2010) utilized contextual information extracted from the use of two mobile applications, namely a phone book and a map browser to create emotional profiles of thirteen students using Bayesian networks. Oh et al. established user's context using date, time, call logs, text message logs, location, user's occupation and gender as well as weather information obtained from web services. They built two Bayesian networks, one to infer user's activity and another one to infer user's emotion. It is worth mentioning that the Bayesian network built to infer emotions takes as input the inferred user's activity. In spite of indicating that (in the evaluation of their approach) the users manually recorded their activities and emotions, they presented empirical results only on activity recognition and results about emotion recognition were left aside.

Lee et al. (2012) utilized smartphones built-in sensors to collect data on users' typing behavior, touchscreen usage patterns, device movements, ambient light conditions, location, time, weather and a user discomfort index based on temperature. In addition, the user's context was complemented by analyzing messages posted by users in a social network every time they experienced an emotion. As in Oh et al. (2010), Lee et al. used a Bayesian network to infer emotions using information from both the social network client and the smartphone's sensors. The emotion recognition system (of Lee et al.) continuously receives

feedback from the users regarding whether the recognized emotion is correct or not in order to re-train the Bayesian network. It should be mentioned that the evaluation of the emotion recognition system involved data of only one participant that used the system for two weeks.

As in Oh et al. (2010) and Lee et al. (2012), Kim and Choi (2011) collected data using smartphone's sensors such as gyroscope and accelerometer to recognize individual's emotions. However, their approach differs from (Oh et al. 2010) and (Lee et al. 2012) in that the user's profile is a combination of sensor data and a set of usage patterns of applications installed on the smartphone. Kim and Choi utilized simple moving average to obtain the temporal usage patterns. It should be noted that Kim and Choi presented only the design of the emotion recognition system and no empirical evaluation was conducted.

Another approach that makes use of contextual information and a mobile data collection mechanism is the work by LiKamWa et al. (2013) who focused on analyzing social interactions and daily activities to infer daily average mood. Users' social interaction was profiled using phone calls, text messages, and emails. As in Kim and Choi (2011), users' daily activities were profiled using application usage in addition to frequently visited locations and browser history. LiKamWa et al. conducted a study involving 32 participants that were instructed to record their mood four times a day for two months. The mood recognition system was supported by a least-squares multiple linear regression. LiKamWa et al. created individual models for each participant, a single general model for all the participants, and a hybrid model combining data from one individual and data from all the other individuals. Using their models, they were capable of recognizing daily average moods with an accuracy of 66% in the case of the general model and an accuracy of 93% in the case of the individual models. However, it should be noted that LiKamWa et al. validated their models using Leave-One-Out Cross-Validation (LOOCV), nevertheless, models validated using LOOCV suffer from overfitting (Liu et al. 2007).

Approaching emotion recognition from a different perspective, Zhang et al. (2010) developed an emotion recognition system taking into account (1) the correlation of the current emotion of an individual with friends' emotions, (2) the correlation of the current emotion with previous ones, and (3) the influence of the individual's environment and activities. Zhang et al. analyzed data from two social networks, namely MSN and LiveJournal, each dataset containing records of 30 and 469,707 users, respectively. Zhang et al. analyzed the temporal correlations using a dynamic factor graph correctly predicting 62% of users' emotions, which were classified as *positive*, *negative*, and *neutral*.

Soleimaninejadian et al. (2018) present an approach for automatic mood detection based on data collected from wearable devices and smartphones. In particular, the

collected data includes users' biometrics (e.g., hearth rate), physical activities (e.g., step count and time spent commuting), sleep quality (e.g., sleep duration), calorie intake, and environment (e.g., weather). Although according to Soleimaninejadian et al., all these data was collected automatically, users had to manually label the data each day with a predefined emotion label: happy, content, anxious, and depressed. Soleimaninejadian et al. analyzed data from five volunteers as well as data from a dataset corresponding to records of one more user. It should be noted that the 6 participants were from four different countries. The data from the five volunteers was collected in a study that lasted 25 days. In particular, Soleimaninejadian et al. aimed to predict and detect valence and arousal mood's dimensions. The accuracy of their models ranged from 71.92 to 89.03% using support vector machines and C4.5 decision trees.

The work by Zhang et al. (2018) presents an approach for the detection of compound emotions, which refer to the simultaneous experience of more than one emotion. In particular, Zhang et al. focused on detecting what they call the *compound emotion state* of the user, which is composed of a combination of six basic emotions (proposed by Ekman (2000)): happiness, sadness, anger, surprise, fear, and disgust. The user's context consisted of four types of data: environment (e.g., noise measurement and location), contact (e.g., social connections and behaviors), application usage (e.g., duration and type of applications used), and physical activity (e.g., step count and sleep duration). The contextual data was collected (every 5 minutes) from smartphone's sensors such as the GPS, microphone, accelerometer, gyroscope, and light sensors. Zhang et al. conducted a study for 1 month that involved 30 students who reported three times a day (in periods of at least 5 hours from record to record) the emotions they experienced in the past few hours. Zhang et al. also proposed a machine learning algorithm based on a factor graph model for automatic detection of compound emotions. In particular, this algorithm was used to build individual models for each participant. The performance results of their proposed machine learning algorithm were compared with three classification algorithms (namely, decision trees, support vector machines, and a logistic regression algorithm) outperforming them with respect to accuracy (among other performance metrics). It is important to note that the proposal by Zhang et al. was designed to build individual models rather than general or gender-specific models.

Zualkernan et al. (2017) present a system for user emotion recognition using contextual data collected from smartphones: (1) data from the smarthphone's accelerometer and (2) data about the typing behavior of the user (e.g., speed, number of backspaces, and time between letters). Zualkernan et al. designed a soft-keyboard to replace the default keyboard implemented in smartphones in order to collect the data about the typing behavior as well as to report their

emotions (namely, neutral, angry, happy, and sad). Users were allowed to indicate their experienced emotion each time they used the provided soft-keyboard. The study conducted by Zualkernan et al. involved 3 participants that provided 307 records in one month. These records were analyzed using a naive Bayes algorithm, a J48 decision tree, a lazy instance-based learning algorithm, a multi-response linear regression, and a support vector machine. According to their results, the J48 decision tree was the best algorithm for emotion recognition with over 90% accuracy and precision.

In the same way as (Oh et al. 2010; Lee et al. 2012; Kim and Choi 2011; LiKamWa et al. 2013; Soleimaninejadian et al. 2018; Zhang et al. 2018; Zualkernan et al. 2017), we make use of a mobile data collection system. In addition, in order to build our emotion recognition models, we collected commonly used contextual data such as location (as in Oh et al. (2010)), time (as in Lee et al. (2012)), and people with whom participants interact with (as in LiKamWa et al. (2013)).

Regarding the type of context utilized to automatically recognize emotions, we found two types of contexts: *real-world* and *digital-world*. A *real-world* context is based on day-to-day individuals' physical situations in which emotions are actually experienced. A *digital-world* context is based on situations derived from how individuals interact with computer devices, smartphones, and/or social networks. Unlike (Conati 2002; Kim and Choi 2011; Zhang et al. 2010; Zualkernan et al. 2017) that make use of a digital-world context based on (1) the interaction of individuals with a videogame, (2) usage patterns of applications, (3) the interaction in social networks, and (4) the user's typing behaviors, respectively, we make use of a real-world context taking into account individuals' locations, activities, companions, among others. In this manner, potential causes of negative or positive emotions may be inferred from the actual context in which individuals carry out their daily activities. Nevertheless, the causes of either negative or positive emotions are out of the scope of this work.

With respect to strong assumptions, in order to recognize the emotions of an individual, Zhang et al. (2010) assumed knowledge about (1) the emotions of the individual's friends and (2) a history of the individual's previous emotions. In contrast, we make no assumption about the emotional state of surrounding people and require no history of previous emotions. Moreover, it is worth mentioning that some emotion recognition systems require collecting data for a given time window in order to recognize emotions, see (Conati 2002; Oh et al. 2010; Lee et al. 2012; Kim and Choi 2011; LiKamWa et al. 2013). This is because some of their features are based on temporal patterns, for instance, Kim and Choi (2011) make use of usage patterns of mobile applications. In contrast, we recognize emotions based on an atemporal context.

It is acknowledged that as in LiKamWa et al. (2013), we created *general* models and *individual* models involving data on all the participants and data on a single participant, respectively. However, we also built *gender-specific* machine learning models for emotion classification. Furthermore, similar to Zhang et al. (2010), we classified emotions as *positive* and *negative*, however, we do not take into account *neutral* emotions because we are interested in determining the contexts that may influence individuals to express *negative* or *positive* emotions. Finally, whereas the emotion recognition system of LiKamWa et al. aims to predict daily average mood, our system aims to predict emotions for every time an individual switches from a context to another. Table 1 includes a summary of relevant characteristics as well as the assumptions made by the related work and our present work.

3 Method

This section provides details of (1) the participants, (2) the data collected about participants' emotions and context, (3) the data validation process, (4) the materials used to collect such data, and (5) the procedure followed by the participants.

3.1 Participants

We recruited 32 undergraduate engineering students to participate in the experiment. Twenty six were male and six were female. We invited engineering students by convenience because this study was conceived and carried out at a technological university. Therefore, their age ranged from 18 to 22 years. The only restriction for users to participate was that they should have a smartphone to install the mobile application developed for data collection (see Sect. 3.4 for details of such application). No other restriction was considered for the selection of participants. All participants signed a *consent form* prior to the experiment in order to get involved in this study voluntarily. The *consent form* stated (1) the purpose of the study, (2) the type of data to be collected, and that (3) no information that could disclose their personal identity will be released.

We motivated the participants to use the data collection application by providing them with movie theater gift cards with a value of 100 Mexican pesos (approximately 5.2 US dollars at the time this paper was written). The participants were told that they could obtain a gift card if they were among the two participants who provide the largest number of records on the first weekend of the study, another two at the middle of the study, two more in a raffle, and another gift card if they were among the two participants who provide the largest number of records from the middle of the study to its completion.

3.2 Collected data

The collected and analyzed data includes (1) type of emotion experienced by participants, (2) type of activity associated with experienced emotions, (3) thermal sensation of the location where the emotions are experienced, (4) whether there is a type of physical affliction, (5) location where the emotions are experienced, (6) people with whom participants interact with, and (7) part of the day (e.g., morning) when the emotions are experienced. These data categories were taken into account in the study based on both (1) a guided brainstorming session that involved a subset of the participants and (2) emotion-related literature. Details about the data categories (and the justification for their inclusion) are as follows:

- *Emotion* The emotions recorded by the participants are *happy, sad, angry, upset, scared, surprised, bored, nervous, and embarrassed*. It should be noted that these emotions were identified in the brainstorming session and are also commonly discussed and analyzed in related literature on emotions, see (Nass et al. 2006; Farmer and Sundberg 1986; Kreibig 2010; Tangney et al. 1992). It is acknowledged that other emotions could have been included in this study (e.g., *enthusiastic*), however, such emotions may be related to the ones included (e.g., *enthusiastic* and *happy* may be closely related). Moreover, we asked feedback from the participants and no mention to additional emotions (other than the ones included) was made.
- *Activity* The participants were asked to indicate the activity from a predefined set of common activity types: *working, studying, eating, attending class, exercising, resting, playing videogames, and walking*. However, due to the numerous activity types that exist, the participants were also allowed to manually record any other activity type. We collected data on participants' activities because some research efforts have concluded that emotions are affected by the activities the individuals perform (Pekrun et al. 2017; Junot et al. 2017).
- *Thermal sensation* The thermal sensations recorded by the participants are *hot, warm, cold and humid*. It should be noted that there is evidence indicating that particular conditions of the atmosphere, the temperature for instance, are related to affective states, see (Keller et al. 2005; Kim et al. 2017; Cabanac 1981). For this reason, the participants were instructed to record the thermal sensation. We used the thermal sensation instead of the temperature because participants' perception about the temperature may vary among them. Whereas there are other thermal sensations, the selected thermal sensations were sufficient to express

Table 1 Related work comparison

Authors' pre-sent research	Conati (2002)	Oh et al. (2010)	Lee et al. (2012)	Kim and Choi (2011)	LiKamWa et al. (2013)	Zhang et al. (2010)	Soleimani-jadian et al. (2018)	Zhang et al. (2018)	Zuallkerman et al. (2017)
Recognized emotion categories	Reproach, shame, and joy	Bored, contented, excited, happiness, nervous, normal, relaxed, sadness, and upset	Happiness, surprise, anger, disgust, sadness, fear, and neutral	It is not associated with a predefined set of emotion categories	Pleased and active	Positive, negative, and neutral	Happy, content, anxious, and depressed	Happiness, sadness, anger, surprise, fear, and disgust	Neutral, angry, happy, sad
Data collection apparatus	A PC video game	Smartphones	Smartphones	Smartphones	Smartphones	Social networks	Wearable devices and smartphones	Smartphones	Smartphones
Data categories for individual's profile	Bodily expressions	Date, time, call logs, text message logs, location, user's occupation, gender, and weather information	Typing behavior, touchscreen usage patterns, device movements, ambient light conditions, location, time, weather and temperature	Usage patterns of applications and smartphone's acceleration and position	Phone calls, text messages, emails, application usage patterns, visited locations, and browser history	Friends' emotions, correlation of the current emotion with previous ones, and individual's environment and activities	Biometrics, physical activities, sleep quality, diet, and user's environment	Environment, contact, application usage, and activity	Smartphone's accelerometer data and typing behavior of the user
Number of individuals involved in the field study	Not applicable	Not applicable	1 individual	Not applicable	32 individuals	469,707 individuals	6 individuals (5 volunteers and 1 from a dataset)	30 individuals	3 individuals
Duration of the field study	Not applicable	Not applicable	14 days	Not applicable	60 days	90 days	25 days	1 month	1 month
Overall number of records	Not applicable	Not applicable	314 records	Not applicable	1,920 records	9,869 records	300 records	Not specified (more that 2000 aprox)	307 records
Model type(s)	Dynamic decision network	Bayesian network	Bayesian network	Simple moving average	Least-squares multiple linear regression	Dynamic factor graph	Support vector machines and C4.5 decision trees	A machine learning algorithm proposed by the authors and based on factor graph model	Naive Bayes, J48 decision tree, lazy instance-based learning, multi-response linear regression, and support vector machines

Table 1 (continued)

	Authors' pre-research	Conati (2002)	Oh et al. (2010)	Lee et al. (2012)	Kim and Choi (2011)	LiKamWa et al. (2013)	Zhang et al. (2010)	Soleimani-jadian et al. (2018)	Zhang et al. (2018)	Zualkernan et al. (2017)
Model evaluation technique	Empirical evaluation using tenfold cross-validation	Discussion of a sample scenario	No evaluation was carried out	Empirical evaluation using tenfold cross-validation	No evaluation was carried out	Empirical evaluation using Leave-One-Out Cross-Validation	Metropolis Hastings algorithm	Empirical evaluation using tenfold cross-validation	Comparison of performance metrics with a decision tree, a support vector machine, and a logistic regression	Empirical evaluation using tenfold cross-validation
Requires feedback from participants to re-train the model	No	No	No	Yes	No	No	No	No	No	No
Creates an individual model using data of one participant	Yes	Not applicable	Not applicable	Yes	Not applicable	Yes	No	No	Yes	No
Creates a general model using data of all the participants	Yes	Not applicable	Not applicable	No	Not applicable	Yes	Yes	Yes	No	Yes
Creates gender-specific models	Yes	Not applicable	Not applicable	No	No	No	No	No	No	No
Assumes knowledge about the emotions of surrounding people	No	No	No	No	No	No	Yes	No	No	No
Assumes knowledge about previous individual's emotions	No	No	No	No	No	No	Yes	Yes	No	No

Table 1 (continued)

Authors' pre-research	Conati (2002)	Oh et al. (2010)	Lee et al. (2012)	Kim and Choi (2011)	LiKamWa et al. (2013)	Zhang et al. (2010)	Soleimani-jadian et al. (2018)	Zhang et al. (2018)	Zualkernan et al. (2017)
The trained model requires collecting data for a given time window to recognize emotions	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No
Nature of contextual data	Digital-world situations	Both real-world and digital-world situations	Both real-world and digital-world situations	Digital-world situations	Both real-world and digital-world situations	Digital-world situations	Both real-world and digital-world situations	Digital-world situations	Digital-world situations

- the participants' perception about the thermal sensations in the city where this study took place.
- *Physical affliction* The physical afflictions recorded by the participants are *sore, tired, hungry, menstruating, sleepy, sick, and none*. The term *none* indicates the absence of physical afflictions, whereas the term *sick* indicates the presence of any other physical affliction. We did not include an exhaustive list of physical afflictions to facilitate data input. The included physical afflictions were mentioned by the participants of the guided brainstorming session. We collected data on physical afflictions due to evidence presented in Hutchings et al. (2001) indicating a relationship between negative emotions and physical afflictions.
- *Location* As in the case of the activity types, the participants were provided with a predefined set of common locations: *home, school, fitness center, restaurant, job, doctor, transport, cafeteria, bar, and party*. Nevertheless, due to the numerous locations that exist, the participants were also allowed to manually record any other location. We collected data on participants' locations because, as indicated by Sandstrom et al. (2017), there is a relationship between individuals' emotions and their location.
- *Company* Similar to the data input mechanism of locations and activities, the participants were provided with a predefined set of common companion types: *no one, friends, mother, father, girlfriend, boyfriend, strangers, sister, and brother*. However, the participants were also allowed to manually record any other type of companion. We collected data on participants' companions because interpersonal interaction may affect individuals' emotions (Dimotakis et al. 2011). It should be noted that sometimes an individual may be interacting with different types of companions, e.g., family and friends, for such cases, we instructed participants to make a choice and indicate a single predominant companion type. In doing so, we avoided unnecessary companion categories.
- *Part of the day* The date and time (of each record) were used to establish a temporal order of records. As indicated by Stone et al. (2006), people may have emotional behavioral patterns associated with the part of the day.

3.3 Data validation process

With the aim of guaranteeing consistent and valid data, we conducted a manual data validation process where the following validation rules were applied:

- We removed all the records of users that entered less than one record per day.
- We removed all the records from users that entered obvious, fictitious information (in the case of the data

categories where users were allowed to manually input information).

- We also removed all the records from users that entered a relatively large number of records, one after the other almost at the same time.
- We removed all the records from users that entered multiple duplicated records.

When any of the validation rules applied to the records of a particular user, we removed all his/her records because we assumed that all his/her records may have been compromised. It should be noted that none of these validation rules applied to the records of the 32 participants (described in Sect. 3.1).

3.4 Materials

A mobile application called *Applings* was developed to support the data collection process (see Fig. 1). This mobile application was designed to allow participants to collect the two types of data used in the present study: emotions (happy, sad, angry, upset, scared, surprised, bored, nervous, and embarrassed) and contextual information (activity, thermal sensation, physical affliction, location, company, and part of the day). In addition, using *Applings*, participants can also register their age and gender. The application was installed in the smartphone of each participant. The technical requirements necessary to install the mobile application are Android operating system version 4.4 (or higher) and

Fig. 1 Data collection mobile application



12MB of disk space. No other special characteristic was required on smartphones as all collected data was stored on a remote server in real time.

The mechanisms implemented for data collection in the mobile application are (1) manual data recording by participants and (2) automatic data recording by the application itself:

- *Emotion* Participants selected the experienced emotion from a predefined set of options (see Fig. 1a).
- *Activity* The data collection mobile application provided the participants with a predefined set of common activity types and with an option to manually record any other activity (see Fig. 1b).
- *Thermal sensation* This information was manually selected by participants from a predefined set of options (see Fig. 1c).
- *Physical affliction* This information was manually selected by participants from a predefined set of options (see Fig. 1d).
- *Location* The mobile application provided participants with a set of predefined options to choose from as well as the option to manually record any other location (see Fig. 1e).
- *Company* The data collection mobile application provided the participants with a predefined set of companion types and with an option to manually record any other type of companion (see Fig. 1f).
- *Part of the day* This data was automatically collected by the mobile application.

By manually self-collecting participants' context, we aimed to obtain an accurate description of the context at least according to what participants perceived. It should be noted that automated data collection mechanisms are possible, for instance, the type of company might be inferred using surrounding audio sensors. Nevertheless, this work is focused on identifying emotions from user's context and fully automated data collection mechanisms are out of the scope.

3.5 Procedure

An informative session was held with all the participants to explain the data collection mechanism implemented by the mobile application. In addition, participants were provided with a *user manual*.

The data collection period lasted 20 days. As part of the data collection strategy, participants were instructed to record experienced emotions and contextual data as follows: *for the next twenty days we would like you to use the developed mobile application to keep a record of your emotions and your context. Whenever possible,*

create a record for each time you switch from one context to another or from one emotion to another.

4 Construction of supervised machine learning models for emotion classification

To explore the feasibility of automatically determining whether an individual expresses a positive or a negative emotion according to a given context, we built commonly used supervised machine learning and statistical classification models, namely (1) a back-propagation multilayer perceptron artificial neural network (for a detailed description see (Witten et al. 2016), p. 256), (2) a random forest [for a detailed description see (Breiman 2001)], (3) a logistic regression with a ridge estimator [for a detailed description see (Le Cessie and Van Houwelingen 1992)], and (4) a naive Bayes classifier [for a detailed description see (John and Langley 1995)].

All the models were created using the Waikato Environment for Knowledge Analysis (Weka) version 3.8.2 (Hall et al. 2009). It should be noted that all the models were validated by using tenfold cross-validation. The performance metrics used to evaluate the models are true positive (TP) rate, false positive (FP) rate, precision, recall, F-measure, and area under the ROC curve. These performance metrics are commonly used to evaluate machine learning classification models (Fawcett 2006). Detailed descriptions of these performance metrics can be found in Fawcett (2006).

We built three types of models: *general* models, *individual* models and *gender-specific* models. The *general* models were created using data of 32 participants (for a total of 1,262 records) with the objective of building a model capable of determining emotion types for all the participants. The *individual* models were created using data of only one (male) participant, the participant that provided the largest number of records (with 182 records). The *individual* models were created with the objective of exploring the feasibility of building specific models for a single individual. It should be noted that we did not create *individual* models for the other (31) participants because they did not provide sufficient records to build robust models. The *gender-specific* models of males were created using data of 26 participants (for a total of 1,024 records) and the *gender-specific* models of females were created using data of 6 participants (for a total of 238 records). The *gender-specific* models were created with the objective of exploring whether grouping participants' records by gender improves the performance of the machine learning models.

4.1 Data preprocessing

It should be mentioned that due to (1) the relatively small size of the sample (consisting of 1262 records collected from 32 participants) and (2) the fact that there were very few records of certain types of emotions, it was unfeasible to build robust supervised machine learning models for emotion classification taking into account the nine types of emotions. So, to increase the number of records per emotion type, we categorized emotions into *positive* and *negative* accordingly, which is an emotion categorization commonly found in emotion-related literature, see (Becerra et al. 2019; Granat et al. 2017; Morrison et al. 2016).

In addition to be provided with predefined sets of options, the participants were also allowed to manually record additional activities, locations, and companions. This caused a myriad of activity types, location types, and companion types. So we grouped them into relevant categories accordingly. The categories for *activity* are *university-related* activities, *leisure* activities, *work-related* activities, *food intake* activities, and *other* activities. The categories for *location* are *home*, *university campus*, and *other* location. The categories for *companion* are *family members*, *friends*, *classmates*, *strangers*, *nobody*, and *other*.

To define the *part of the day*, we grouped timestamps into *morning*, *afternoon*, *evening*, and *late nights* (from midnight to before dawn). However, it is acknowledged that the *date* records were disregarded because these were insufficient to detect a long-term pattern.

An example of a record of a user’s context associated with an emotion is as follows: {Emotion type : *positive*, Context : [activity = *leisure*, thermal sensation = *warm*, physical affliction = *none* , location = *home* , company = *friends*, part of the day = *evening*]}. From the 32 participants, we obtained a total of 1262 contexts associated with their corresponding emotion categories either *positive* or *negative*. Using those records, we built 16 supervised machine learning classification models (namely 4 *general* models, 4 *individual* models, 4 models of females, and 4 models of males), which received as input the user’s context whereas their output is the emotion category.

4.2 Control parameters

The control parameters of the models were determined based on a tuning to maximize their effectiveness as measured by the area under the ROC curve. We selected the area under the ROC curve (as the primary performance measure) because as stated by Swets (1988), it is the preferred measure of accuracy of classifier systems. In addition, the area under the ROC curve is not affected by the class distribution, which is important in this case, because for the *individual* models only 18.68% of the instances are labeled as *negative* and for the *general* models only 40.25% of the instances are labeled as *negative* (see Fig. 2).

The control parameters of the *general* models are as follows. The neural network’s input layer has 29 neurons (because all the attributes were nominal and were converted

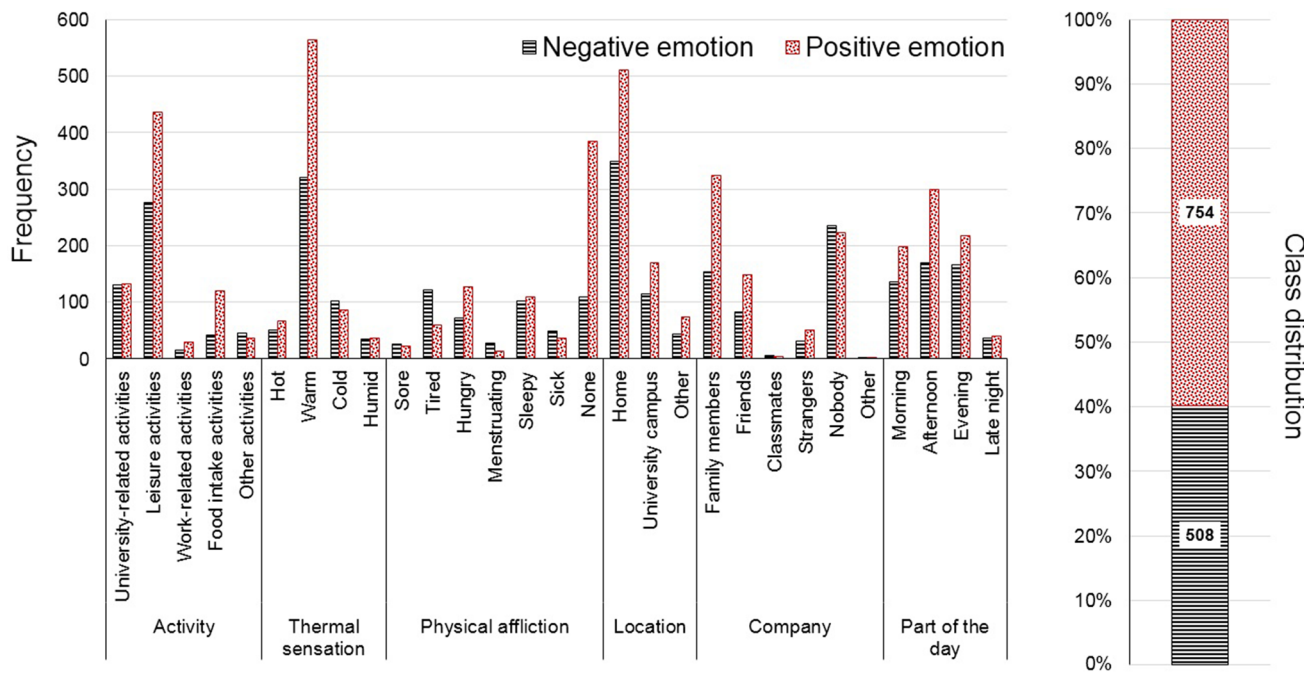


Fig. 2 Histograms of each data category by emotion class and class distribution

into binary numeric attributes increasing the number of input neurons), the hidden layer has 7 sigmoid neurons, and the output layer has 2 neurons (one for each class). The number of epochs was set to 500 with a decaying learning rate of 0.05 and a momentum of 0.25. For the random forest model, the number of trees was set to 100, and from the best of 2 (out of the 6) features, one of the features was selected at random to split the sample for each decision tree node. In addition, the maximum depth of the trees was set to unlimited. For the logistic regression model, we used the default parameters, i.e., the number of iterations was set to unlimited to allow for the convergence of the model and the ridge parameter was set to 1.0×10^{-8} . For the naives Bayes classifier using categorical features, the Weka suite does not have parameters to be set.

The control parameters of the *individual* models are as follows. The neural network’s input layer and output layer also has 29 neurons and 2 neurons, respectively, however, the hidden layer has 1 sigmoid neuron. The number of epochs was also set to 500 but with a decaying learning rate of 0.5 and a momentum of 0.25. For the random forest model, the number of trees was also set to 100, and from the best of 4 (out of the 6) features, one of the features was selected at random to split the sample for each decision tree node. In addition, the maximum depth of the trees was also

set to unlimited. For the logistic regression model and the naives Bayes classifier, we used the default parameters.

The control parameters of the *gender-specific* models were the same as those of the *general* models.

It should be noted that all the models were created using a meta-classifier to allow for cost-sensitive classification by reweighting training instances according to user-defined costs for each class. This cost-sensitive classifier enabled us to introduce a matrix cost where misclassifying an instance as a *negative* emotion costs twice as much to misclassifying an instance as a *positive* emotion. This was made under the assumption that the models can be used to introduce an intervention mechanism whenever an individual is experiencing negative emotions, and as a consequence, misclassifying a *negative* emotion as *positive* prevents the intervention mechanism from acting.

5 Results and analysis

From the evaluation of the supervised machine learning models for emotion classification (Figs. 3, 4, 5, 6, 7, and 8), an information gain analysis (Fig. 9), and the histograms of each feature (Fig. 2), we drawn five main observations.

Fig. 3 Percentage of correctly classified instances for all the models

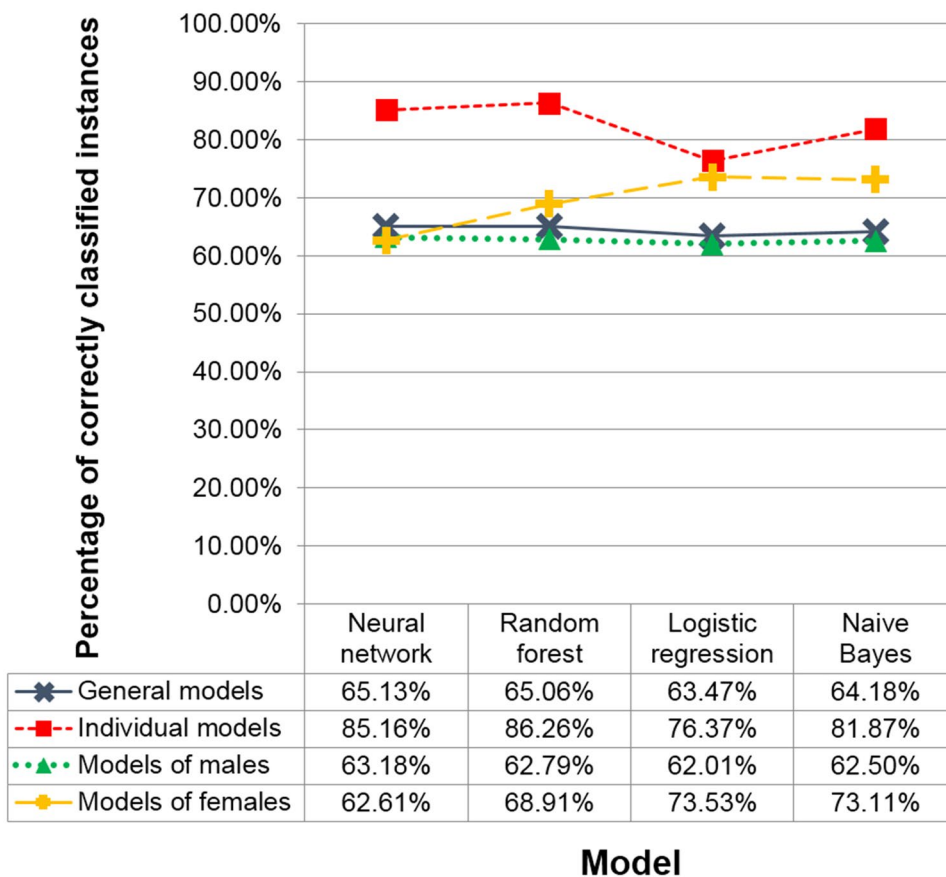


Fig. 4 Confusion matrices for the emotion classifiers

General models				Individual models			
	negative	positive			negative	positive	
Neural network	372	136	negative	Neural network	23	11	negative
	304	450	positive		16	132	positive
Random forest	344	164	negative	Random forest	20	14	negative
	277	477	positive		11	137	positive
Logistic regression	380	128	negative	Logistic regression	17	17	negative
	333	421	positive		26	122	positive
Naive Bayes	375	133	negative	Naive Bayes	14	20	negative
	319	435	positive		13	135	positive
Models of males				Models of females			
	negative	positive			negative	positive	
Neural network	219	140	negative	Neural network	149	0	negative
	237	428	positive		89	0	positive
Random forest	210	149	negative	Random forest	136	13	negative
	232	433	positive		61	28	positive
Logistic regression	230	129	negative	Logistic regression	139	10	negative
	260	405	positive		53	36	positive
Naive Bayes	231	128	negative	Naive Bayes	139	10	negative
	256	409	positive		54	35	positive

Observation 1 The areas under the ROC curve for both *individual* and *general* models ranged from 0.703 to 0.838 as shown in Figs. 5 and 6.

Analysis Four types of supervised machine learning models for emotion classification built from contextual data of either 32 participants (with 1262 instances) or a single participant (with 182 instances) were to some extent capable of determining whether the emotions of participants were either *negative* or *positive*. It should be noted that all the models were evaluated using tenfold cross-validation [a commonly adopted mechanism for validating classifiers (Forman and Scholz 2010)]. In addition, according to Swets, a model

with an area under the ROC curve greater than 0.7 is useful. By having eight correctly validated, useful *individual* and *general* models with an area under the ROC curve greater than 0.7, we obtained evidence supporting the feasibility of automatically determining whether an individual expresses a *positive* emotion or a *negative* emotion from a given context.

Observation 2 According to the information gain criterion, the most informative feature is *physical affliction* and the least informative feature is *location* (see Fig. 9).

Analysis To evaluate how informative a feature is with respect to the emotion class, all the attributes were ranked

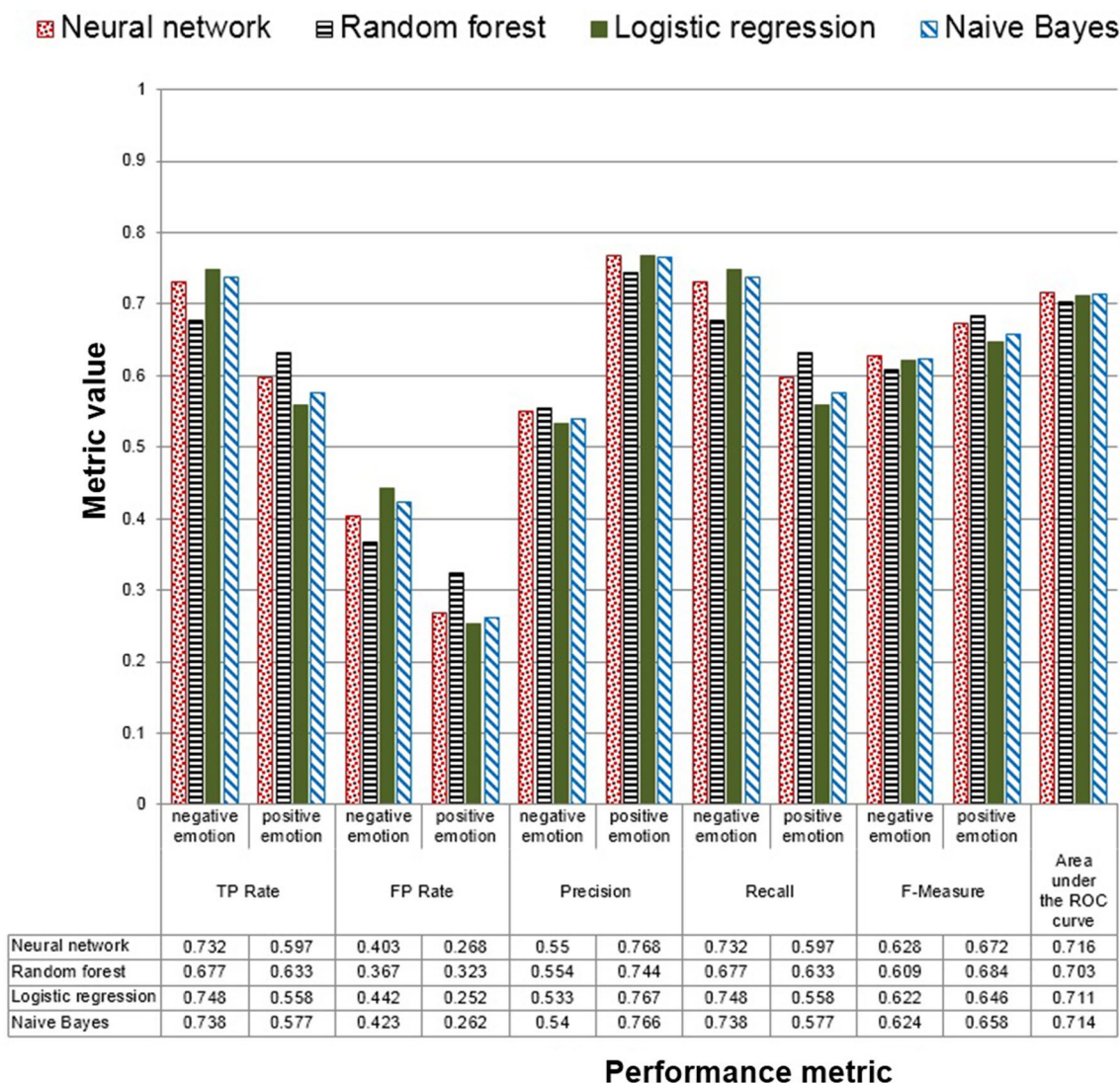


Fig. 5 Performance metrics for all the general models separated by emotion class: negative and positive

based on information gain (also using Weka). In the context of this work, the information gain indicates the feature’s relevance for emotion classification. Fig. 9 shows the range and distribution of information gain for each feature. It should be noted that although in general the most informative feature with respect to the emotion class was *physical affliction*, there was even one participant whose most informative feature was *location* (the least informative feature when taking into account the overall data from all the participants). This suggests that models should be created for each individual. This claim is also supported by the fact that the average percentage of correctly classified instances of the individual models (82.42%) was higher than the average percentage of correctly classified instances of the general models (64.46%), see Fig. 3. Additional evidence that supports this claim is the fact that LiKamWa et al. (2013) also found that

the average percentage of correctly classified instances of their individual models (up to 93%) was higher than the performance of their general models (66%). However, it should be noted that the models of LiKamWa et al. determine daily average moods, whereas our models determine emotion categories for every time an individual switches from a context to another. It is also important to notice that LiKamWa et al. validated their models using Leave-One-Out Cross-Validation (LOOCV), nonetheless, models validated using LOOCV suffer from overfitting (Liu et al. 2007). In contrast, we validated our models using tenfold cross-validation.

Observation 3 For the *general* and *individual* models, the best performance with respect to TP rate of *negative* emotions was achieved by the multilayer perceptron artificial neural network with a rate of 73.2% and 67.6%, respectively.

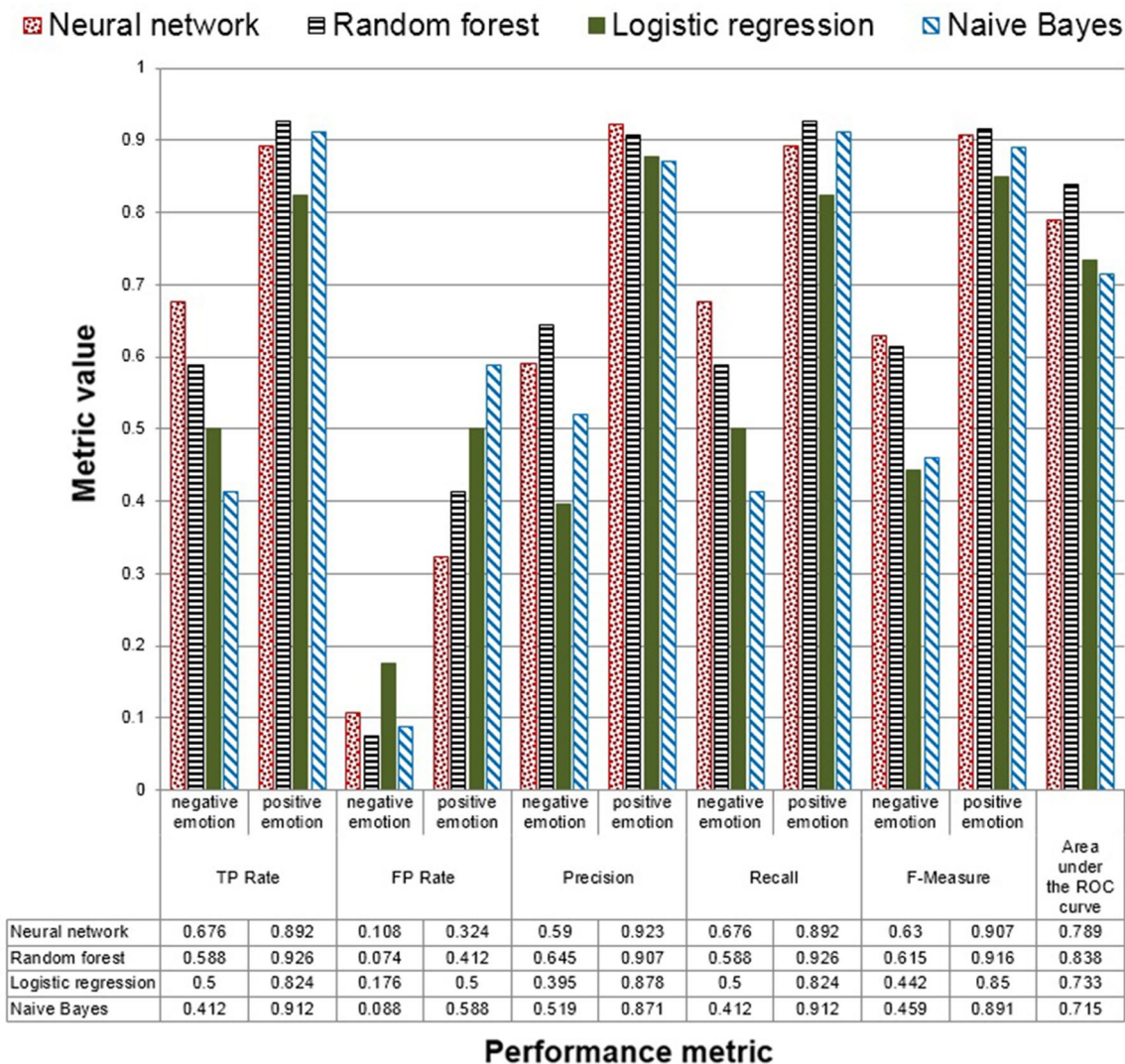


Fig. 6 Performance metrics for all the individual models separated by emotion class: negative and positive

Analysis By building a cost-sensitive multilayer perceptron artificial neural network, we aimed to reduce the error in misclassifying *negative* emotions as *positive* emotions due to the potential future applications of this work to improve the emotional state of individuals via an intervention mechanism. As shown in Figs. 5 and 6, on average, 7 out of 10 *negative* emotions were detected using the *general* and *individual* neural network models. However, when using the *individual* neural network model, 10.8% of the *positive* emotions were classified as *negative* and when using the *general* neural network model 40.3% of the *positive* emotions were also classified as *negative*. These FP rates of *negative* emotions would cause unnecessary interventions. So a future intervention mechanism should be designed in such a way that potential unnecessary interventions would not be intrusive.

Observation 4 In general, the machine learning models for emotion classification of females are better than the models for emotion classification of males when considering percentage of correctly classified instances and area under the ROC curve.

Analysis On average, the percentage of correctly classified instances of the models for emotion classification of females was 69.54%, which was higher than that of the models for emotion classification of males (62.62%), see Fig. 3. In addition, on average, the area under the ROC curve of the models for emotion classification of females was 0.743, which was also higher than that of the models for emotion classification of males (0.662), see Figs. 7 and 8. This result may seem counter-intuitive because the models for emotion classification of males were built using

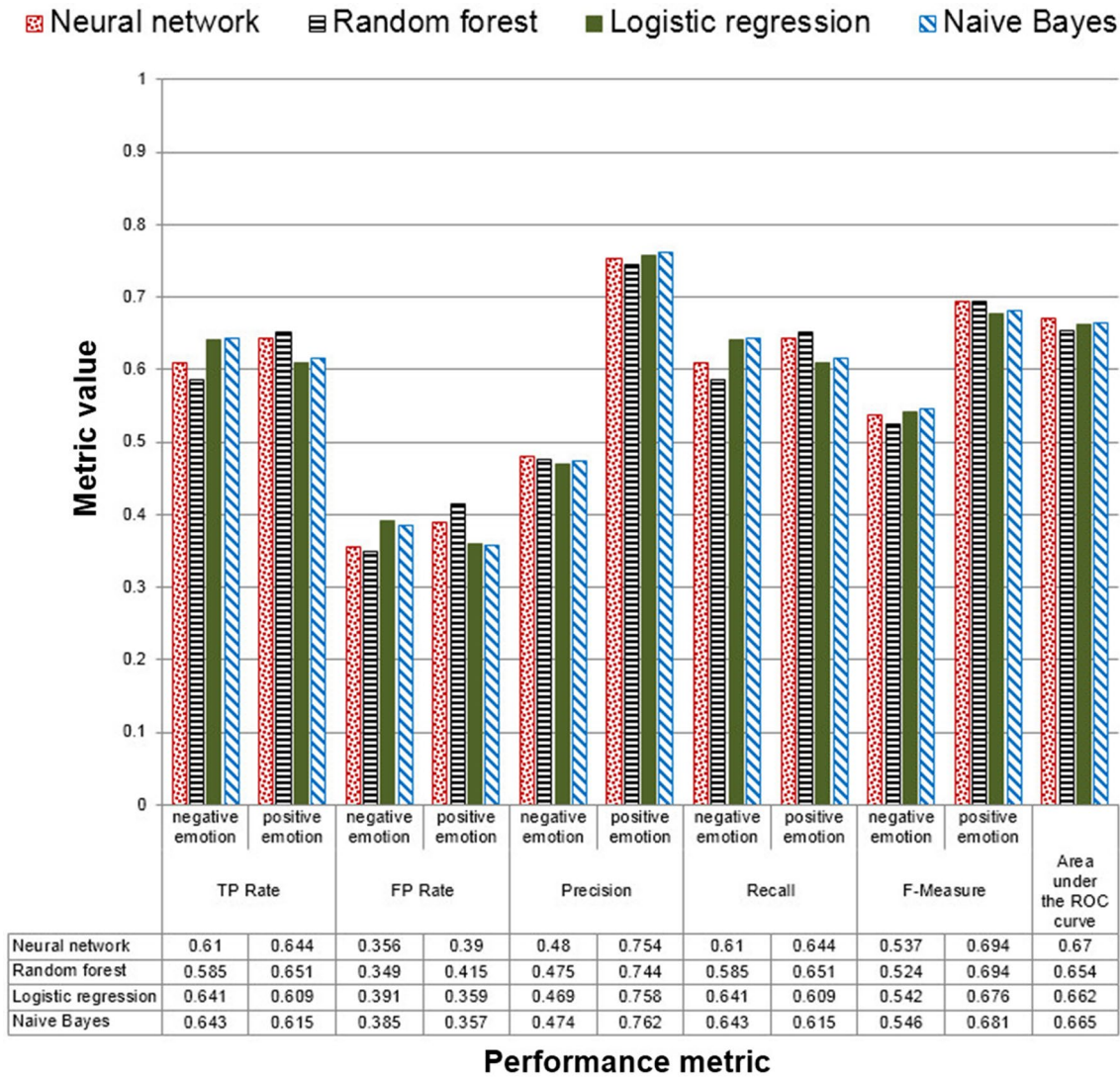


Fig. 7 Performance metrics for all the models of males separated by emotion class: negative and positive

data of 26 participants, whereas the models for emotion classification of females were built using data of only 6 participants (for a total of 238 records). In fact, this relative small number of records were insufficient to create a robust neural network model for emotion classification of females, which classified all the instances (either *negative* or *positive*) as *negative* (see Fig. 4). Nonetheless, the percentage of correctly classified instances of the random forest model, the logistic regression model, and the naive Bayes model for emotion classification of females was higher than that of the *general* models (see Fig. 3). This fact as well as the result analyzed in *Observation 3* regarding the *individual* models outperforming the *general* models, may suggest that machine learning models for emotion classification should be built from contextual data of either (1) specific groups of individuals or (2) single individuals.

Furthermore, this result may also suggest that there is a stronger relationship between women’s self-reported emotions and contextual aspects than the relationship between men’s self-reported emotions and context. Nevertheless, we acknowledge that more data from more male and female participants as well as more experiments may be required to confirm this result.

Observation 5 According to the histogram shown in Fig. 2, the three most common contextual aspects associated with *positive* emotions are (1) engaging in leisure activities, (2) feeling a warm thermal sensation, and (3) being at home (Fig. 2). However, the histogram (Fig. 2) also shows that those same contextual aspects are also the most associated with *negative* emotions.

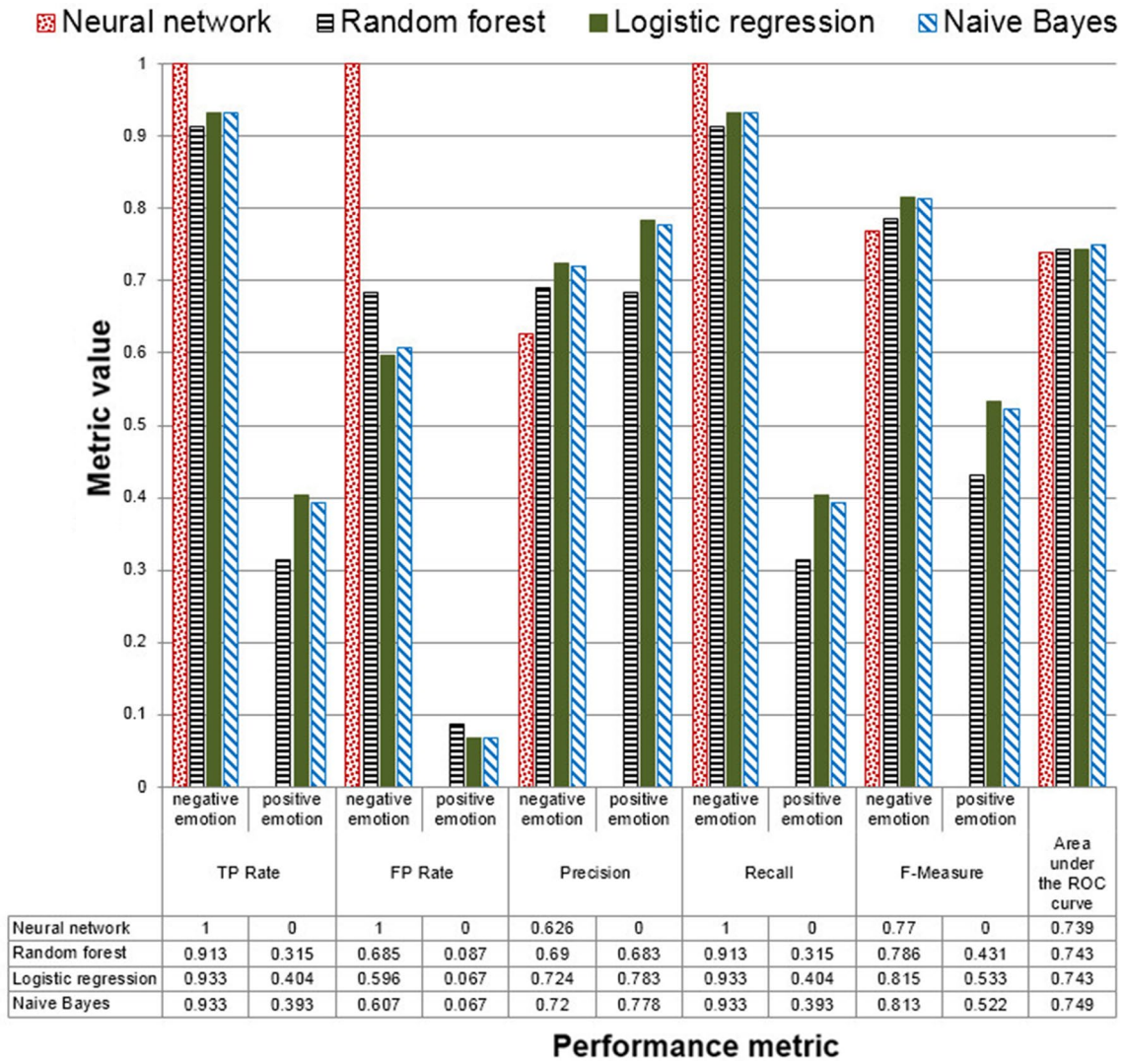


Fig. 8 Performance metrics for all the models of females separated by emotion class: negative and positive

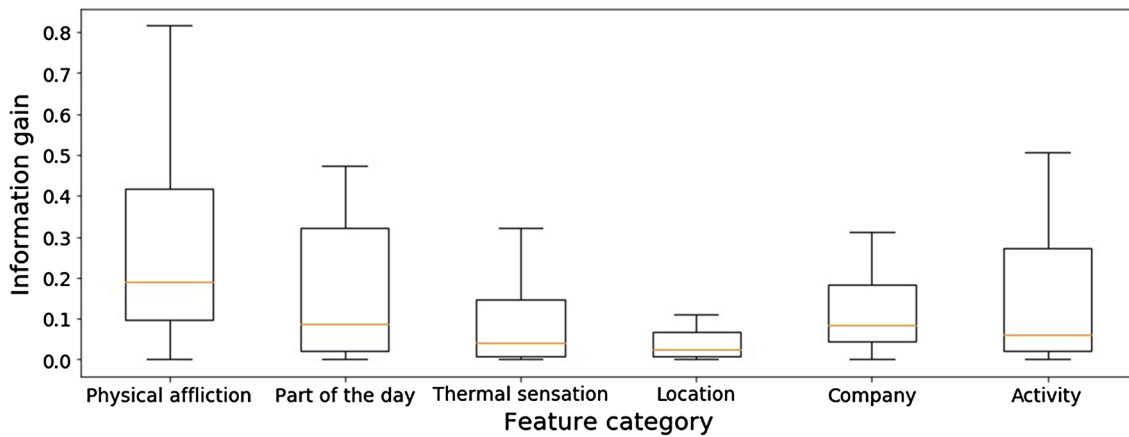


Fig. 9 Range and distribution of information gain for each feature

Analysis The statistical analysis indicates that, while at home, individuals experienced both *negative* and *positive* emotions numerous times. Whereas this seems counter-intuitive, the fact that the individuals were at home suggests that the individuals interacted with other family members, and in this regard, family relationships are complex (Sprenkle and Piercy 2005), which may have caused both *negative* and *positive* emotions. With respect to the warm thermal sensation, the fact that some individuals expressed either *negative* or *positive* emotions while feeling a warm sensation may have been circumstantial because during the study the weather remained warm. In regard to the emotional bipolarity in leisure activities, a potential reason is that even though the participants were engaged in leisure activities, the participants did not enjoy them. From this observation, we can conclude that emotions are multifactorial, i.e., no single contextual aspect can serve as a predictor of whether an individual will express *negative* or *positive* emotions, and thus, in order to determine the emotion category experienced by individuals, a combination of features should be used (as performed by the 16 machine learning models proposed in this work).

6 Discussion

In this paper, we have examined the feasibility of automatically recognizing individuals' emotions from contextual information using machine learning techniques. Our results confirm that, to some extent, emotions of young adults (aged between 18 and 22 years) who pursue an engineering degree can be recognized from contextual aspects. However, it remains an open research question as to whether it is feasible to recognize emotions from contextual information for individuals with different profiles other than undergraduate engineering students (e.g., artists or art students) and/or of different age groups (e.g., seniors). With respect to individuals' profiles and emotions, Sheldon (1994) found differences in emotional behaviors between scientists and artists. With respect to age and emotions, Martin-Krumm et al. (2018) provide empirical evidence indicating that emotional behaviors evolve as individuals mature. This suggests that further research is required to investigate emotion recognition from contextual information for individuals with different profiles and/or of different age groups.

In addition, the results obtained from our *individual*, *general*, and *gender-specific* models, suggest that machine learning models for emotion recognition from contextual information should be built for either specific groups or individuals. In this regard, the research study by Mesquita et al. (2017) states that individuals' emotional experience is culturally shaped and constituted. Moreover, another study that found group differences in emotions was conducted by

Martin-Krumm et al. (2018) who found that there are gender differences with respect to emotions. In fact, Martin-Krumm et al.'s results may help to explain the differences in performance found between our models for emotion classification of females and males. As a consequence of these group differences in emotions, in order to improve the performance of our machine learning models, the context should be enriched with other culture-related features and/or gender-related features.

It is acknowledged that this research relies on self-reported data about emotions and context. In this regard, self-report methods are commonly used to collect data on emotions (Kanjo et al. 2015), however, self-reported data may be biased as participants may not be able to correctly indicate their emotions. In the context of this research effort, a possible way to avoid this limitation is to conduct controlled experiments with predefined, fabricated situations where emotional reactions of participants can be monitored and analyzed by psychologists. This solution may also have other limitations caused by the fabricated contexts, in which participants may not be themselves and experienced emotions may not represent those that may be experienced in real situations. Another possible way to avoid this limitation is to build a fully automated system to collect data about participants' emotions, activities, companions and/or physical afflictions, among other contextual aspects, which may be a very challenging endeavor that is out of the scope of this work. Nonetheless, in order to have a pervasive and ubiquitous emotion recognition system, the construction of an Internet-of-Things platform for automated data collection is in progress, see (Salido Ortega et al. 2018). Finally, to attenuate this limitation, we explained to the participants all the information to be collected and made emphasis on the fact that their participation in the study was entirely voluntary. In addition, no strict periodic data input was required from them. In doing so, we aimed to have a relative large number of individuals willing to participate in the study with no intention to input noisy data.

Another limitation of our work is the relative small size of the sample. In this regard, our field study involved 32 participants as in LiKamWa et al. (2013). However, it is worth mentioning that this type of studies usually have fewer participants as in (Oh et al. 2010; Lee et al. 2012; Soleimanejadrian et al. 2018; Zhang et al. 2018; Zualkernan et al. 2017) or do not present an empirical evaluation as in (Conati 2002; Kim and Choi 2011). It should be noted that the study by Zhang et al. (2010) involved 469,707 participants, however, they extracted data from social networks (a digital-world context), which prevented them to take advantage of real-world contextual information. It is worth mentioning that by collecting data from as many participants as possible, we were able to properly validate our emotion recognition system using tenfold cross-validation while preventing

overfitting [unlike the use of LOOCV as in LiKamWa et al. (2013)]. Nevertheless, we acknowledge that more data from potentially many more participants may be required to build machine learning models for emotion recognition with a negligible error rate.

7 Conclusions

The novelty of this work is that, to the best of our knowledge, it is the first work in automatically recognizing emotions based solely on a *real-world* context characterized by day-to-day individual's physical situations in which emotions are actually experienced. The main contribution of this work is to provide empirical evidence of the feasibility of emotion recognition from contextual information using four types of machine learning models, namely multilayer perceptron neural networks, random forest, logistic regression, and a naive Bayes classifier. It is worth mentioning that we built *individual* models, *general* models, and *gender-specific* models with performance comparable to those obtained by emotion recognition system based on either *digital-world* contextual information or a combination of *digital-world* and *real-world* contextual information. In addition, another contribution is our information gain analysis of context features with respect to a binary emotion categorization.

This research effort lays the foundations for further work to study emotions and their association with contextual features. Among the potential future research directions are (1) to develop an automatic intervention mechanism so as to improve affective states of individuals supported by contextual information and (2) to explore potential causes of either negative or positive emotions, which may be inferred from individuals' context.

Acknowledgements J. O. Gutierrez-Garcia gratefully acknowledges the financial support from the Asociación Mexicana de Cultura, A.C. This work was supported by PFCE 2019.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

- Ahn H, Picard RW (2005) Affective-cognitive learning and decision making: a motivational reward framework for affective agents. In: Tao J, Tan T, Picard RW (eds) *Affective computing and intelligent Interaction*. ACII 2005. Lecture notes in computer science, vol 3784. Springer, Berlin, Heidelberg, pp 866–873
- Alegre U, Augusto JC, Clark T (2016) Engineering context-aware systems and applications: a survey. *J Syst Softw* 117:55–83
- Ashkanasy NM, Daus CS (2002) Emotion in the workplace: the new challenge for managers. *Acad Manage Exec* 16(1):76–86
- Becerra R, Preece D, Campitelli G, Scott-Pillow G (2019) The assessment of emotional reactivity across negative and positive emotions: development and validation of the perth emotional reactivity scale (pers). *Assessment* 26(5):867–879
- Bechara A (2004) The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. *Brain Cogn* 55(1):30–40
- Bechara A, Damasio H, Damasio AR (2000) Emotion, decision making and the orbitofrontal cortex. *Cereb Cortex* 10(3):295–307
- Bellavista P, Corradi A, Fanelli M, Foschini L (2012) A survey of context data distribution for mobile ubiquitous systems. *ACM Comput Surv (CSUR)* 44(4):1–45
- Brave S, Nass C (2007) Emotion in human-computer interaction. In: Sears A, Jacko JA (eds) *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*. CRC Press, Boca Raton, pp 103–118
- Breazeal C (2003) Emotion and sociable humanoid robots. *Int J Hum Comput Stud* 59(1):119–155
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
- Cabanac M (1981) Physiological signals for thermal comfort. In: Cena K, Clark JA (eds) *Bioengineering, thermal physiology and comfort*. Elsevier, Amsterdam, pp 181–192
- Calvo RA, D'Mello S (2010) Affect detection: an interdisciplinary review of models, methods, and their applications. *IEEE Trans Affect Comput* 1(1):18–37
- Conati C (2002) Probabilistic assessment of user's emotions in educational games. *Appl Artif Intell* 16(7–8):555–575
- Costa H, Macedo L (2012) Affective computing. Tech Report, University of Coimbra, Coimbra, Portugal. <https://eden.dei.uc.pt/~hpcosta/docs/papers/techReports/TAMC-stateOfTheArt.pdf>
- Croy I, Olgun S, Joraschky P (2011) Basic emotions elicited by odors and pictures. *Emotion* 11(6):1331–1335
- Deng J, Xu X, Zhang Z, Frühholz S, Schuller B (2017) Universum autoencoder-based domain adaptation for speech emotion recognition. *IEEE Signal Process Lett* 24(4):500–504
- Dimotakis N, Scott BA, Koopman J (2011) An experience sampling investigation of workplace interactions, affective states, and employee well-being. *J Organ Behav* 32(4):572–588
- Ekman P (2000) Basic emotions. In: Dalgleish T, Power M (eds) *Handbook of cognition and emotion*. Wiley, pp 45–60
- Ekman PE, Davidson RJ (1994) *The nature of emotion: fundamental questions*. Oxford University Press, Oxford
- El Ayadi M, Kamel MS, Karray F (2011) Survey on speech emotion recognition: features, classification schemes, and databases. *Pattern Recogn* 44(3):572–587
- Farmer R, Sundberg ND (1986) Boredom proneness-the development and correlates of a new scale. *J Pers Assess* 50(1):4–17
- Fasel B, Luettin J (2003) Automatic facial expression analysis: a survey. *Pattern Recogn* 36(1):259–275
- Fawcett T (2006) An introduction to roc analysis. *Pattern Recogn Lett* 27(8):861–874
- Forman G, Scholz M (2010) Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *ACM SIGKDD Explor Newslett* 12(1):49–57
- Fredrickson BL (2001) The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *Am Psychol* 56(3):218–226
- Giatsoglou M, Vozalis MG, Diamantaras K, Vakali A, Sarigiannidis G, Chatzisavvas KC (2017) Sentiment analysis leveraging emotions and word embeddings. *Expert Syst Appl* 69:214–224
- Granat A, Gadassi R, Gilboa-Schechtman E, Feldman R (2017) Maternal depression and anxiety, social synchrony, and infant regulation of negative and positive emotions. *Emotion* 17(1):11–27

- Grünerbl A, Muaremi A, Osmani V, Bahle G, Oehler S, Tröster G, Mayora O, Haring C, Lukowicz P (2015) Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE J Biomed Health Inform* 19(1):140–148
- Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The weka data mining software: an update. *ACM SIGKDD Explor Newslett* 11(1):10–18
- Hutchings CV, Shum KW, Gawkrödger DJ (2001) Occupational contact dermatitis has an appreciable impact on quality of life. *Contact Dermat* 45(1):17–20
- John GH, Langley P (1995) Estimating continuous distributions in Bayesian classifiers. In: *Proceedings of the eleventh conference on uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., San Francisco, pp 338–345
- Junot A, Paquet Y, Martin-Krumm C (2017) Passion for outdoor activities and environmental behaviors: a look at emotions related to passionate activities. *J Environ Psychol* 53:177–184
- Kanjo E, Al-Husain L, Chamberlain A (2015) Emotions in context: examining pervasive affective sensing systems, applications, and analyses. *Pers Ubiquit Comput* 19(7):1197–1212
- Keller MC, Fredrickson BL, Ybarra O, Côté S, Johnson K, Mikels J, Conway A, Wager T (2005) A warm heart and a clear head: the contingent effects of weather on mood and cognition. *Psychol Sci* 16(9):724–731
- Kim HJ, Choi YS (2011) Emosens: affective entity scoring, a novel service recommendation framework for mobile platform. In: *Proceedings of the 5th ACM conference on recommender system*, pp 1–4
- Kim M, Chong SC, Chun C, Choi Y (2017) Effect of thermal sensation on emotional responses as measured through brain waves. *Build Environ* 118:32–39
- Kořakowska A (2018) Usefulness of keystroke dynamics features in user authentication and emotion recognition. In: Hippe Z, Kulikowski J, Mroczek T (eds) *Human-computer systems interaction. Advances in intelligent systems and computing*, vol 551. Springer, Cham, pp 42–52
- Kreibig SD (2010) Autonomic nervous system activity in emotion: a review. *Biol Psychol* 84(3):394–421
- Le Cessie S, Van Houwelingen JC (1992) Ridge estimators in logistic regression. *Appl Stat* 41(1):191–201
- Lee H, Choi YS, Lee S, Park I (2012) Towards unobtrusive emotion recognition for affective social communication. In: *IEEE consumer communications and networking conference*, pp 260–264
- Lewis M (2008) The emergence of human emotions. In: Lewis M, Haviland-Jones JM, Feldman Barrett L (eds) *Handbook of emotions*, 3rd edn. Guilford Press, New York, pp 304–319
- LiKamWa R, Liu Y, Lane ND, Zhong L (2013) Moodscope: building a mood sensor from smartphone usage patterns. In: *Proceedings of the 11th annual international conference on Mobile systems, applications, and services*, Taipei, Taiwan, pp 389–402
- Liu KH, Huang DS, Li B (2007) Improving the performance of ICA based microarray data prediction models with genetic algorithm. In: *2007 IEEE congress on evolutionary computation*, Singapore, pp 606–611
- Loewenstein G, Lerner JS (2003) The role of affect in decision making. In: Davidson RJ, Sherer KR, Goldsmith HH (eds) *Handbook of affective science*, Oxford University Press, pp 619–642
- Martin-Krumm C, Fenouillet F, Csillik A, Kern L, Besançon M, Heutte J, Paquet Y, Delas Y, Trousselard M, Lecorre B et al (2018) Changes in emotions from childhood to young adulthood. *Child Indic Res* 11(2):541–561
- Mesquita B, Boiger M, De Leersnyder J (2017) Doing emotions: the role of culture in everyday emotions. *Eur Rev Soc Psychol* 28(1):95–133
- Morrison AS, Mateen MA, Brozovich FA, Zaki J, Goldin PR, Heimberg RG, Gross JJ (2016) Empathy for positive and negative emotions in social anxiety disorder. *Behav Res Ther* 87:232–242
- Nalepa GJ, Kutt K, Bobek S (2019) Mobile platform for affective context-aware systems. *Future Gener Comput Syst* 92:490–503
- Nass C, Takayama L, Brave S (2006) Socializing consistency: from technical homogeneity to human epitome. In: Zhang P, Galletta DF (eds) *Human-computer interaction and management information systems: foundations*. M. E. Sharpe, Armonk, NY, pp 373–391
- Oh K, Park HS, Cho SB (2010) A mobile context sharing system using activity and emotion recognition with bayesian networks. In: *7th international conference on ubiquitous intelligence & computing and 7th international conference on autonomic & trusted computing*, Xian, Shaanxi, pp 244–249
- Ortony A, Clore GL, Collins A (1990) *The cognitive structure of emotions*. Cambridge University Press, Cambridge
- Panda R, Malheiro RM, Paiva RP (2019) Novel audio features for music emotion recognition. *IEEE Trans Affect Comput* 1:1–1
- Pekrun R, Vogl E, Muis KR, Sinatra GM (2017) Measuring emotions during epistemic activities: the epistemically-related emotion scales. *Cogn Emot* 31(6):1268–1276
- Phelps EA (2006) Emotion and cognition: insights from studies of the human amygdala. *Annu Rev Psychol* 57:27–53
- Politou E, Alepis E, Patsakis C (2017) A survey on mobile affective computing. *Comput Sci Rev* 25:79–100
- Ptaszynski M, Rzepka R, Araki K (2010) On the need for context processing in affective computing. *Proc Fuzzy Syst Sympos Jpn Soc Fuzzy Theory Intell Inform* 26:920–924
- Rodríguez PM, Del Pino DA, Alvaredo RB (2011) De lo psicológico a lo fisiológico en la relación entre emociones y salud. *Revista Psicología Científica* 13(19):34–39
- Royet JP, Zald D, Versace R, Costes N, Lavenne F, Koenig O, Gervais R (2000) Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: a positron emission tomography study. *J Neurosci* 20(20):7752–7759
- Salido Ortega MG, Rodríguez LF, Gutiérrez-García JO (2018) Energy-aware data collection from the internet of things for building emotional profiles. In: *Third international conference on fog and mobile edge computing (FMEC)*, Barcelona, pp 234–239
- Sandstrom GM, Lathia N, Mascolo C, Rentfrow PJ (2017) Putting mood in context: using smartphones to examine how people feel in different locations. *J Res Pers* 69:96–101
- Seo J, Laine TH, Sohn KA (2019) Machine learning approaches for boredom classification using eeg. *J Ambient Intell Human Comput*. <https://doi.org/10.1007/s12652-019-01196-3>
- Sheldon KM (1994) Emotionality differences between artists and scientists. *J Res Pers* 28(4):481–491
- Soleimanejadani P, Zhang M, Liu Y, Ma S (2018) Mood detection and prediction based on user daily activities. In: *First Asian conference on affective computing and intelligent interaction (ACII Asia)*, Beijing, pp 1–6
- Sprenkle DH, Piercy FP (2005) Pluralism, diversity, and sophistication in family therapy research. *Res Methods Fam Therapy* 2:3–18
- Stone AA, Schwartz JE, Schkade D, Schwarz N, Krueger A, Kahneman D (2006) A population approach to the study of emotion: diurnal rhythms of a working day examined with the day reconstruction method. *Emotion* 6(1):139
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240(4857):1285–1293
- Tagar MR, Federico CM, Halperin E (2011) The positive effect of negative emotions in protracted conflict: the case of anger. *J Exp Soc Psychol* 47(1):157–164
- Tangney JP, Wagner P, Fletcher C, Gramzow R (1992) Shamed into anger? the relation of shame and guilt to anger and self-reported aggression. *J Pers Soc Psychol* 62(4):669

- Wegrzyn M, Vogt M, Kireclioglu B, Schneider J, Kissler J (2017) Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS One* 12(5):1–15
- Wharton AS, Erickson RI (1993) Managing emotions on the job and at home: understanding the consequences of multiple emotional roles. *Acad Manag Rev* 18(3):457–486
- Wingenbach TS, Ashwin C, Brosnan M (2018) Sex differences in facial emotion recognition across varying expression intensity levels from videos. *PLoS One* 13(1):1–18
- Witten IH, Frank E, Hall MA, Pal CJ (2016) *Data mining: practical machine learning tools and techniques*. Morgan Kaufmann, Burlington
- Yin Z, Zhao M, Wang Y, Yang J, Zhang J (2017) Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Comput Methods Programs Biomed* 140:93–110
- Zhang Y, Tang J, Sun J, Chen Y, Rao J (2010) Moodcast: emotion prediction via dynamic continuous factor graph model. In: *IEEE international conference on data mining*, Sydney, NSW, pp 1193–1198
- Zhang X, Li W, Chen X, Lu S (2018) Moodexplorer: towards compound emotion detection via smartphone sensing. *Proc ACM Interact Mobile Wear Ubiquitous Technol* 1(4):176:1–176:30
- Zualkernan I, Aloul F, Shapsough S, Hesham A, El-Khorzaty Y (2017) Emotion recognition using mobile phones. *Comput Electr Eng* 60:1–13

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.