Pervasive Stress Recognition for Sustainable Living

Pervasive Stress Recognition for Sustainable Living

Andrey Bogomolov*, Bruno Lepri†, Michela Ferron†, Fabio Pianesi†, and Alex (Sandy) Pentland‡
*SKIL Telecom Italia Lab, University of Trento
Via Sommarive, 5, I-38123 Povo - Trento, Italy
E-mail: andrey.bogomolov@unitn.it
†Fondazione Bruno Kessler
Via Sommarive, 18 I-38123 Povo - Trento, Italy
E-mail: {lepri, ferron, pianesi}@fbk.eu
‡The Media Lab, Massachusetts Institute of Technology
Cambridge, MA, USA. E-mail: pentland@mit.edu

***Abstract*—In this paper we provide the evidence that daily stress can be reliably recognized based on human behavior metrics derived from the mobile phone activity (call log, sms log, bluetooth interactions). We introduce an original approach for feature extraction, selection, recognition model training and discuss the experimental results based on Random Forest and Gradient Boosted Machine algorithms. Random Forest based model showed low variance comparing to the GBM-based one, thus winning the bias-variance tradeoff and preventing overfitting, given the noisy source data. Potential impact of the technology is reducing stress and enhancing subjective well-being for sustainable living.**

I. Introduction

Psychological stress is a dynamic process eliciting a negative emotional response and it occurs when an individual feels that the environmental demands exceed his/her adaptive capacity [37]. It is a well-known condition in modern life and research has shown that the amount of cumulative stress can play a role in a diverse range of physical, psychological and behavioural conditions, such as anxiety, low self-esteem, depression, social isolation, irritability, cognitive impairments, sleep and immunological disorders, neurodegenerative diseases and other medical conditions [12], [17], [18] Hence, measuring stress in daily life situations has become an important challenge [36].

Nowadays, smartphones provide a novel opportunity for unobtrusive and cost-efficient access to previously inaccessible sources of data related to daily social behavior [14]. Smartphones are able to capture and store huge amounts of behavioral data. The social psychologist Geoffrey Miller in "The Smartphone Psychology Manifesto" argued that smartphones should be seriously considered as a new research tool for social psychology. In his opinion, such tools could revolutionize all fields of psychology and behavioral sciences, making these scientific disciplines more powerful, sophisticated, international, applicable, and grounded in real-world behavior [32]. In the pervasive and ubiquitous computing community, several works have started to use smartphone activity data to detect and predict personality traits [10], [41], mood states [29], and daily happiness [35], [6].

In our paper, we formulated the automatic recognition of daily stress as a 2-class classification problem (stressed and not stressed) based on information concerning: a) people activities, as detected through their smartphones; b) weather conditions; c) personality traits. People activities are represented by features extracted from calls, sms logs and bluetooth hits. We created a large number of models for different weather conditions and individual personality traits, given the fact that they influence emotional and mental states substantially. All these models were combined in one ensemble model.

Classification experiments were performed using a variety of approaches, including decision trees, neural networks and support vector machines. The best solution for our classification problem were found using an ensemble of decision tree classifiers based on Random Forest and Generalized Boosted Model algorithms.

The article is structured as follows. Section II reviews previous work on stress detection. Section III describes the data collection while Section IV describes our approach to stress recognition and, in particular, the features we extracted, the best feature space selection, the learning algorithms we used, and the best model selection. The experimental results together with the data and model limitations are discussed in Section V. Section VI sets out social implications of automatic stress recognition for sustainable living. Summary and future work expectations are defined in Section VII.

II. Related Work

A substantial amount of research on stress recognition focused on physiological measurements to detect stress (see for example [25], [28]). In particular, several methods have been based on physiological signals, such as heart-rate variability, skin conductance, temperature, respiration, blood pressure or muscle activity. Despite their high reliability in stress detection, technologies based on this approach present considerable weaknesses because they need to be carried at all times in order to allow continuous monitoring.

A set of less obtrusive techniques used in stress detection are based on voice analysis and in particular on variations in speech production, and take into consideration different acoustic features, such as phonetic variations, pitch, glottal pulse or spectral slope [30]. However, this methodology relies on high acoustic quality, which is not always achievable in natural settings, and its reliability can be undermined by large individual differences in the correlation between speech and emotion [40]. Moreover, despite being less obtrusive than other physiological sensors, stress detection based on speech may still pose privacy problems related to the recording and analysis of human voice.

A promising and reliable alternative that can overcome the major shortcomings posed by these methodologies is activity recognition from mobile phones usage patterns. In particular, research in this field has mainly investigated general behavioural change and peoples relational dynamics [4], [16]. researchers have focused on the use of mobile phones to better understand individuals affective state [29] and stress levels [5], [39].

LiKamWa and colleagues [29] focused on mood recognition and developed MoodScope, a software system that detects users mood from mobile phone usage data, such as phone calls, sms logs, email messages, application use, web browsing histories and location changes. The system achieved an initial 66% accuracy of participants' daily mood, gradually improving to 93% after two months of training, with phone calls and categorized applications being the most relevant features.

Bauer and Lukowicz [5] addressed the problem of stress recognition, and monitored 7 students during a two weeks exam session followed by two weeks of non-stressful period. They recorded data such as users location and social interactions through Bluetooth proximity, phone calls and sms logs, and detected an average behaviour modification of 53% for each participant during the exam session.

Sano and Picard [39] collected 5 days of data for 18 participants integrating a wrist sensor with morning/evening surveys (Big Five, mood, sleep quality, tiredness and stress levels, or use of technological devices) and mobile phone usage features (such as calls, sms logs, mobility patterns, screen on/off mode). The authors applied correlational analysis to find the most relevant features and reached a 75% accuracy using machine learning to classify whether the participants were stressed or not. This multifactorial approach achieves a good accuracy level, which is comparable to the results presented in this paper. However, we tested our approach on more reliable dataset capturing the daily of larger amount of subjects (117 vs 18) for eight weeks.

In general, these previous studies suggest smartphones as a valuable source of rich data from real life, which can be exploited to gain beneficial insights about peoples affective state and stress levels, in order to develop new context-aware mobile services that could better support psychological wellbeing.

III. Our Dataset

Inside the Friends and Family longitudinal study (see for more details [3]), we collected a dataset capturing daily stress data of 117 subjects for more than eight weeks. During this period, each participant was equipped with an android smartphone and with a sensing software explicitly designed for this data collection. The source data consists of call logs, sms logs, social proximity data, obtained by scanning near-by phones and other Bluetooth devices every five minutes and ground truth survey bases self-reported data about personality ("Big Five" personality traits) and daily stress for each day of the experiment. Social interactions were derived from Bluetooth proximity data in a similar way to previous reality mining studies [15]. The FUNF phone sensing framework [1] was used to detect Bluetooth devices in the user's proximity. The Bluetooth scan was performed every 5 minutes in order to keep from draining the battery while achieving a high temporal resolution of social interactions tracking.

For the data analysis described in this paper we used 33497 phone calls, 22587 SMS, and 1460939 Bluetooth hits.

Additionally, the participants were asked to fill daily surveys about their self-perceived stress level. The stress information was reported by the participants filling a seven items scale with 1 = "not stressed", 4 = "neutral" and 7 = "extremely stressed". For the stress recognition task presented in this paper

TABLE I. Recorded Daily Stress

Min.	0.000
1st Qu.	3.000
Median	4.000
Mean	3.793
3rd Qu.	5.000
Max.	7.000

the survey based labels data was transformed from 1 to 4 to "not stressed" class, and from 5 to 7 – to "stressed". The ground truth labels distribution for our recognition task tend to be approximately balanced by the classes.

In addition, personality was measured by asking subjects to answer on a 1-5 point scale to the online version of the Big Five questionnaire developed by John et al. [27], which owes its name to the five traits taken as a constitutive of people's personality: Extraversion vs. Introversion; Emotional stability vs. Neuroticism; Agreeableness vs. Disagreeableness; Conscientiousness vs. Unconscientiousness; Openness to experience. We obtained personality scores by computing a sum for each item in the personality trait questionnaire.

A partial version of the data, consistent with privacy and legal limitations, is publicly shared for the research community [2]. We hope that the Friends and Family study and similar initiatives may be a first step for dealing with the lack of empirical naturalistic pervasive computing data in the real world [34].

IV. Our Approach

Based on previous findings in social psychology [44], we tested if extraverted people, as opposed to introverted people, would be stressed or not given the same social interactions and we found that we needed to create separate models for extraverted and introverted people.

The idea to filter the *weather implications* on the independent variables was the second novelty of our approach. It was found to be consistent with environmental psychology studies showing a significant effect of temperature, hours of sunshine and humidity on mood [26], [38], [13]. In our experiments, we used the following weather parameters: mean temperature, pressure, total precipitation, humidity, visibility and wind speed metrics, extracted from public sources.

Based on previous works that characterize social interactions by means of mobile phone data and usage of social interactions to predict personality traits and daily happiness states [33], [6], we derived 25 call and sms features and 9 proximity features. These features, grouped in four broad categories, characterize *general phone usage, diversity, active behavior and regularity*.

More precisely, features for general phone usage capture the total number of outgoing, incoming and missed calls, and the total number of sent and received sms. Moreover, we also extracted the outgoing to incoming calls ratio, missed to (outgoing + incoming) calls ratio, and sms sent/received ratio.

Diversity measures how evenly an individual's time is distributed among others. In our case, the diversity of user behavior is addressed by means of three kinds of features: (i) entropy of contacts, (ii) unique contacts to interactions ratio, (iii) number of unique contacts. We computed the diversity features both for calls and sms. For entropy calculation, we applied *Miller-Madow correction* [31]. Miller-Madow correction, dealing with the data quality problems, to get bias-corrected empirical entropy estimate.

Concerning regularity features, we measured the time elapsed between calls, the time elapsed between sms exchanges and the time elapsed between call and sms. We consider both the average and variance of the inter-event time of one's call, sms, call+sms. It is worthwhile to note that even though two users have the same inter-event time for both call and sms, their mean inter-event times for call+sms can be very different.

Our contribution to the feature space creation is based on using sliding window functions to capture intertemporal influence of the stress recognition independent variables. For example, for each feature from the basic feature subset we calculated second order features, such as mean, median, min, max, 99%, 95% quantiles, quantiles for the cases of 0.5, 1, 1.5 and 2 standard deviations from the mean (applying Chebyshev's inequality), variance and standard deviation functions. Analyzing the time domain, for each basic feature subset we calculated the same functions for 2 and 3 days backward moving window to check if some events from the near past influenced the current stress state.

For feature selection we used mean decrease in Gini Coefficent [23]. The Gini coefficient measures the inequality between the values of a frequency distribution, – in our case the dimensions of the feature vector. A Gini coefficient of 0 expresses perfect equality, where all values have the same predictive power. A Gini coefficient of 1 expresses maximal inequality among the variables. However, for most features, values close to 1 are very unlikely in practice.

A. Classification Algorithms and Model Selection

We formulated the automatic recognition of daily stress as a classification problem with two classes ("not stressed" or "stressed"). The ground truth labels for classification problem were set to 0 for "not stressed", where label score $<= 4$ and 1 for "stressed", where label score > 4.

We separated all the data at random, following an uniform distribution, in a training set and in a control test set fixing the proportion of 80:20. To let optimization algorithms converge more efficiently the feature matrix was centered and normalized by each column [7].

Applying a grid search approach we trained a number of sets of classifiers: support vector machines, neural networks, ensemble of tree classifiers based on a Breiman's Random Forest (RF) and Friedmans Generalized Boosted Model (GBM) [19] algorithms with different parameters.

Multiple regression models, *support vector machines model* [43] with linear and Gaussian radial basis [9] kernels and multi-layer perceptron *neural network* did not provide good classification results or required building separate models for each personality type and weather conditions.

Random forest algorithm produces a combination of tree predictors, such that each tree is dependent on the values of a random vector sampled independently with the same distribution for all the classification trees in the forest [8]. The decision boundary is formed according to the margin function. Given an ensemble of tree classifiers $h_1(\mathbf{x}), h_2(\mathbf{x}), ..., h_K(\mathbf{x})$ and if the training set is drawn at random from the empirical distribution of the random vector $Y, \mathbf{X}$ the margin function is defined as:

$$\begin{aligned} mg(\mathbf{X}, Y) = avg_k I(h_k(\mathbf{X}) = Y) - \\ \max_{j!=Y} avg_k I(h_k(\mathbf{X}) = j), \end{aligned} \tag{1}$$

where $I(\cdot)$ is the characteristic function. The margin function measures the distance between the average votes at $(\mathbf{X}, Y)$ for the right class and the average vote for any other class. For this model the generalization error function is:

$$PE^* = P_{\mathbf{X},Y}(mg(\mathbf{X}, Y) < 0), \tag{2}$$

where $P_{\mathbf{X},Y}$ is the probability over $\langle \mathbf{X}, Y \rangle$ space. For any event $A \subset \Omega$ of the feature space the characteristic function $I(\cdot)$ of A is:

$$I_A(x) = \begin{Bmatrix} 1 & \iff (x \subset A) \\ 0 & otherwise \end{Bmatrix} \begin{Bmatrix} 1 & \iff \exists x \\ 0 & otherwise \end{Bmatrix} \tag{3}$$

For the second ensemble model (GBM-based) we adopted a greedy function approximation and the stochastic gradient boosting strategy, which are described in [20] and [21]. The optimization problem was formulated as finding a function, $\hat{f}(\mathbf{x})$, that minimizes the loss function $\Psi(y, f)$:

$$\hat{f}(\mathbf{x}) = \arg\min_{f(\mathbf{x})} {}_{y|\mathbf{x}} [\Psi(y, f(\mathbf{x}))|\mathbf{x}] \tag{4}$$

The implementation implementer for this paper solution is described in Algorithm 1.

In order to find the best model, we trained a number of models and selected the best one based on κ metrics for the 10-fold validation strategy. The Cohen's κ measures pairwise agreement among a set of functions which are making classification decisions with correction for an expected chance agreement [11]. $\kappa = 0$ if there is no agreement more than expected by chance following the empirical distribution. $\kappa = 1$ when there is a max agreement. κ is a state-of-the-art statistics about how significantly the classification model is different from chance. More importantly, it makes the interpretation of the scale of what the model has learned to be an intuitive task. κ statistic has the properties of more robust and conservative measure to show what we have learned from the data than F1 and area under the ROC curve metrics.

Algorithm 1: Gradient Boosting Implementation for Stress Recognition Problem

Input: $\vec{X} = \vec{x}_{i}{}_{n=1}^{N}$ such that $\vec{x} \in H$;
$\vec{Y} = \vec{y}_{1}{}_{n=1}^{N}$ such that $\vec{y} \in F_2$
Output: $\hat{f}(\vec{x})$ such that
$\hat{f}(\vec{x}) = \arg\min_{\rho} \sum_{i=1}^{N} \Psi(y_i, \rho)$.

begin
$\Psi \longleftarrow$ Bernoulli
$T \longleftarrow$ number of trees
$K \longleftarrow$ terminal nodes limit
$p \longleftarrow$ subsampling rate
$\hat{f}(\vec{x}) \longleftarrow \emptyset$
for $i \in T$ **do**
Compute gradient:

$$z_i = -\frac{\partial}{\partial f(\vec{x}_i)} \Psi(y_i, f(\vec{x}_i)) \Bigg|_{f(\vec{x}_i)=\hat{f}(\vec{x}_i)}$$

Select $p \times N$ observations from the feature space
Fit the tree, limited to K terminal nodes
Compute the optimal terminal node:
$\rho_k = \arg\min_{\rho} \sum_{\vec{x}_i \in S_k} \Psi(y_i, \hat{f}(\vec{x}_i) + \rho)$
for $j \in length(\vec{x})$ **do**
Update $\hat{f}(\vec{x})$, such that

$$\hat{f}(\vec{x}) \leftarrow \hat{f}(\vec{x}) + \lambda \rho_{j(\vec{x})}$$

end
end
return $\hat{f}(\vec{x})$
end

During the learning and model selection process we followed a leave-one-out 10-fold cross validation strategy. We adopted this strategy in order to prevent data overfitting and to deal with potential data loss in cases when calls, sms and Bluetooth proximities existed in the real world but were not registered by the mobile application.

The performance metrics used to evaluate the models are: accuracy, sensitivity, specificity and Cohen's κ. For detailed analysis we provide model confusion matrix [42]. The significance of the results is supported by the proper statistical tests [22] and is shown in the tables.

V. Experimental Results and Discussion

The final feature vector selected for the recognition task has 111 dimensions. Interestingly, for the computational power of the current devices it may be considered being a low-dimensional feature vector, that can be computed not only on powerful servers but even on a smartphone.

Random Forests were trained through a step-wise increase of the number of trees equal to the upper limit of 2^{11}. Optimal number of trees for model generalization as measured by mean misclassification rate for 10-fold cross validation strategy is estimated to be 112 trees.

Recognition model based on Random Forest algorithm shows 90.68% accuracy on the training set and 72.51% accuracy on the test set (Table II).

Metric	Value
Accuracy	0.7239
95% CI	(0.7063, 0.741)
P-Value [Acc > NIR]	< 2.2e-16
Kappa	0.3707
Sensitivity	0.5037
Specificity	0.8486

TABLE II. RF-based Model Performance Metrics

Instead, GBM-based model showed the results provided in Table III. Despite better Accuracy and Cohen's κ metrics for Random Forest based model, the choice of the more adequate learning approach depends on the concrete application task. Specificity of 92.52% using GBM-based model in comparison with 84.86% using random forest algorithm can be favorable if we are focused more on "not stressed" class, rather than stress detection.

Metric	Value
Accuracy	0.7135
95% CI	(0.6957, 0.7308)
P-Value [Acc > NIR]	3.156e-16
Kappa	0.2992
Sensitivity	0.3397
Specificity	0.9252

TABLE III. GBM-based Model Performance Metrics

The confusion matrices for each model test sets are provided in Table IV and Table V. These matrices show that there is a major agreement between classes. But Table IV shows that random forest based solution is better for detecting stress.

	0	1
0	1404	460
1	254	479

TABLE IV. Confusion Matrix for RF-based Model

	0	1
0	1534	620
1	124	319

TABLE V. Confusion Matrix for GBM-based Model

The investigation of the most important predictors of daily stress reveals interesting associations.

All the personality traits contributed significantly in predicting the daily stress variable. These results are interesting because previous studies in social psychology focused on Neuroticism, Extraversion and Conscientiousness. Instead, our work shows the important contribution played also by Agreeableness and Openness for the automatic classification of daily stress.

With regard to weather, we found a confirmatory association between temperature, humidity, wind speed, pressure, total precipitation and visibility and stress.

Regarding mobile phone data, it is interesting to note the significant contribution of the proximity features. Among the top 30 features used for stress recognition, 12 features are proximity ones calculated from the Bluetooth data. In particular, an interesting predictive role is played by the time intervals for which an id is seen. In addition, features capturing the diversity in co-location interactions are in the top 30 list (e.g. entropy of proximity contacts). This result seems to confirm previous studies in social psychology that found associations between people's stress and the richness in terms of the amount and the diversity of people's social interactions [24]. Instead, among the 30 less predictive features we can find the number of times in which the least common ID is seen. Interestingly, the results obtained using proximity features seem to confirm previous findings in social psychology: in particular, the relevant role played by interactions with strong ties in determining the stress level of a subject. For sure, this result requires further investigation.

As for call interactions, we can infer the role played by general phone usage features such as the number of incoming calls and the number of outgoing calls. On the contrary, the role of sms interactions for predicting daily stress is less evident from our investigation. The only feature related to sms interactions among the top 30 predictive features is subject active behavior and more specifically is the latency in replying to a text message. Therefore, the predictive power of the sms data needs further investigation.

The limitations of our study include the following: (1) our sample comes from a population living in the same environment – a campus facility of a major US university; (2) non-availability of proximity data concerning the interaction with people not participating in the data collection; (3) missing data coming out of battery issues; (4) absence of location data, which in fact could have been tracked.

VI. Social Implications for Sustainable Living

Stress has become a major problem in our society. New information and communication technologies, ubiquitous connectivity, information overload, increased mental workload and time pressure are all elements contributing to increase general stress levels. While in some cases people may realize that they are undergoing stressful situations, for example under intense pressure such as an approaching deadline, severe and chronic stress may be more difficult to detect. Moreover, stress is not necessarily always perceived as negative, and may be considered the norm in a modern and demanding society. Nonetheless, while slightly increased stress levels may be functional for productivity, prolonged and severe stress can be at the source of several physical dysfunctions like headache, sleep or immunological disorders, unhealthy behaviours such as smoking and bad eating habits, as well as of psychological and relational problems that can deeply affect peoples social life.

Our technology provides a cost-effective, unobtrusive, widely available and reliable tool for stress recognition. It detects daily stress levels with a 72.39% accuracy combining real life data from different sources, such as personality traits, social relationships (in terms of calls, text messages and Bluetooth proximity), and weather data. The development of a reliable stress recognition system is a first but essential step toward wellbeing and sustainable living, and its scope can be extended to different areas of applicability. Providing people with a tool capable of gathering rich data about real life, and transforming them into meaningful insights about stress levels, paves the way to a new generation of context-aware technologies that can target clinics as well as the corporate sector and common citizens. This technology can inform the design of ICT based medical decision support systems for the assessment and treatment of psychological stress. With such a tool, therapists could monitor and record patients daily stress levels, access longitudinal data, identify recurrent or significant stressors and modulate treatment accordingly.

In work environments, where stress has become a serious problem not only affecting productivity and leading to occupational issues, but also causing health diseases, our system could be extended and employed for early detection of stress related conflicts and stress contagion in professional social networks, and for supporting balanced workloads. Awareness is a first but crucial step to motivate people to change their behaviour and take informed and concrete steps toward a healthy lifestyle and an appropriate stress balance. Mobile applications developed on the basis of our technology could be targeted to the general public and provide contextual feedback to increase peoples awareness of their stress levels, alerts when they reach a warning threshold, and suggest stress management and relaxation techniques when appropriate.

VII. Conclusion and Future Work

In this paper we provide a new evidence that daily stress can be reliably recognized based on human behavior metrics derived from mobile phone data.

The scientific novelty is focused on the original approach for feature extraction, selection, and the ensemble recognition model which combines a number of models for each different weather conditions and personality dispositions.

Despite the limitations discussed above, we believe that our solution, resulting the 72.39% accuracy and 0.37 Cohen's κ metrics for 2-class classification problem, have provided substantial proof that individual daily stress can be predicted from smartphone data. The experimental results based on Random Forest and Gradient Boosted Machine algorithms are discussed, showing that Random Forest based algorithm is better to detecting stress, and Gradient Boosted Machine based model provides better results for detecting "not stressed" daily states. Random Forest based model shows low variance comparing to the GBM-based one, thus winning the bias-variance tradeoff and preventing overfitting, given the noisy source data.

Individual, social and business implications of the proposed technology for sustainable living are significant and practical, given the fact of unobtrusive and cost-effective way of data collection, low-dimensional feature space discovered and the algorithmic efficiency of the proposed stress recognition model.

Future work will be focused on different feature subsets interaction and on a multi-step stress recognition model development, on the first step predicting personality from user's mobile phone activity and on the next step – detecting stress in a compeletely automated way.

Acknowledgments

The work of Andrey Bogomolov was partially supported by Telecom Italia Semantics and Knowledge Innovation Laboratory (SKIL) with research grant T.

References

[1] FUNF Open Sensing Framework. http://www.funf.org, Access date: 12-JUN-2013.

[2] Friends and Family Dataset. http://realitycommons.media.mit.edu/friendsdataset.html, Access date: 15-JAN-2014.

[3] Nadav Aharony, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing*, 7(6):643–659, 2011.

[4] Nadav Aharony, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. The social fmri: measuring, understanding, and designing social mechanisms in the real world. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 445–454. ACM, 2011.

[5] Gerald Bauer and Paul Lukowicz. Can smartphones detect stress-related changes in the behaviour of individuals? In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 423–426. IEEE, 2012.

[6] Andrey Bogomolov, Bruno Lepri, and Fabio Pianesi. Happiness recognition from mobile phone data. In *Social Computing (SocialCom), 2013 International Conference on*, pages 790–795, 2013.

[7] G. E. P. Box and D. R. Cox. An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252, 1964.

[8] Leo Breiman. Random forests. *Mach. Learn.*, 45(1):5–32, October 2001.

[9] Martin D. Buhmann and M. D. Buhmann. *Radial Basis Functions*. Cambridge University Press, New York, NY, USA, 2003.

[10] Gokul Chittaranjan, Jan Blom, and Daniel Gatica-Perez. Mining large-scale smartphone data for personality studies. *Personal Ubiquitous Comput.*, 17(3):433–450, March 2013.

[11] Jacob Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37–46, April 1960.

[12] S. Cohen, Kessler R. C., and L. U. Gordon. *Measuring stress: A guide for health and social scientists*. Oxford University Press, USA, 1997.

[13] Jaap J. A. Denissen, Ligaya Butalid, Lars Penke, and Marcel A. G. Van Aken. The effects of weather on daily mood: A multilevel approach. *Emotion Researcher*, 8(5):662–667, 2008.

[14] Wen Dong, Bruno Lepri, and Alex Pentland. Modeling the co-evolution of behaviors and social relationships using mobile phone data. In *MUM*, pages 134–143, 2011.

[15] Nathan Eagle and Alex Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.

[16] Nathan Eagle, Alex Sandy Pentland, and David Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106(36):15274–15278, 2009.

[17] Tobias Esch, Gregory L Fricchione, George B Stefano, et al. The therapeutic use of the relaxation response in stress-related diseases. *Med Sci Monit*, 9(2):23–34, 2003.

[18] G. Fink. *Stress consequences: Mental, neuropsychological and socioeconomic*. Elsevier, San Diego, CA, 2010.

[19] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997.

[20] Jerome H Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, pages 1189–1232, 2001.

[21] Jerome H Friedman. Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4):367–378, 2002.

[22] Salvador Garcia and Francisco Herrera. An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons. *Journal of Machine Learning Research*, 9:2677–2694, December 2008.

[23] C Gini. Concentration and dependency ratios. *Rivista di Politica Economica*, 87:769–792, 1997.

[24] William J. Goode. A theory of role strain. *American Sociological Review*, 25(4):pp. 483–496, 1960.

[25] J.A. Healey and R.W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *Intelligent Transportation Systems, IEEE Transactions on*, 6(2):156–166, 2005.

[26] Edgar Howarth and Michael S Hoffman. A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology*, 75(1):15–23, 1984.

[27] Oliver P. John and Sanjay Srivastava. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In Lawrence A. Pervin and Oliver P. John, editors, *Handbook of Personality: Theory and Research*, pages 102–138. Guilford Press, New York, second edition, 1999.

[28] E. Jovanov, A. O'Donnell Lords, D. Raskovic, P.G. Cox, R. Adhami, and F. Andrasik. Stress monitoring using a distributed wireless intelligent sensor system. *Engineering in Medicine and Biology Magazine, IEEE*, 22(3):49–55, 2003.

[29] Robert LiKamWa, Yunxin Liu, Nicholas D. Lane, and Lin Zhong. Moodscope: building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, MobiSys '13, pages 389–402, New York, NY, USA, 2013. ACM.

[30] Hong Lu, Denise Frauendorfer, Mashfiqui Rabbi, Marianne Schmid Mast, Gokul T. Chittaranjan, Andrew T. Campbell, Daniel Gatica-Perez, and Tanzeem Choudhury. Stresssense: detecting stress in unconstrained acoustic environments using smartphones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 351–360, New York, NY, USA, 2012. ACM.

[31] G. Miller. Note on the bias of information estimates. *Information theory in psychology: Problems and methods*, 2(95):100, 1955.

[32] Geoffrey Miller. The smartphone psychology manifesto. *Perspectives on Psychological Science*, 7(3):221–237, 2012.

[33] Yves-Alexandre Montjoye, Jordi Quoidbach, Florent Robic, and Alex(Sandy) Pentland. Predicting personality using novel mobile phone-based metrics. In ArielM. Greenberg, WilliamG. Kennedy, and NathanD. Bos, editors, *Social Computing, Behavioral-Cultural Modeling and Prediction*, volume 7812 of *Lecture Notes in Computer Science*, pages 48–55. Springer Berlin Heidelberg, 2013.

[34] S. Moran, I.L. de Vallejo, K. Nakata, R. Conroy-Dalton, R. Luck, P. McLennan, and S. Hailes. Studying the impact of ubiquitous monitoring technology on office worker behaviours: The value of sharing research data. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 902–907, 2012.

[35] Amir Muaremi, Bert Arnrich, and Gerhard Tröster. A survey on measuring happiness with smart phones. In *6th International Workshop on Ubiquitous Health and Wellness (UbiHealth 2012)*, 2012.

[36] K. Plarre, A. Raij, S.M. Hossain, A.A. Ali, M. Nakajima, M. Al'absi, E. Ertin, T. Kamarck, S. Kumar, M. Scott, Daniel Siewiorek, A. Smailagic, and L.E. Wittmers. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, pages 97–108, 2011.

[37] Cohen S, Janicki-Deverts D, and Miller GE. Psychological stress and disease. *JAMA*, 298(14):1685–1687, 2007.

[38] Jeffrey L Sanders and Mary S Brizzolara. Relationships between weather and mood. *The Journal of General Psychology*, 107(1):155–156, 1982.

[39] Akane Sano and Rosalind W Picard. Stress recognition using wearable sensors and mobile phones. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 671–676. IEEE, 2013.

[40] Klaus R Scherer, Didier Grandjean, Tom Johnstone, Gudrun Klasmeyer, and Thomas Bänziger. Acoustic correlates of task load and stress. In *INTERSPEECH*, 2002.

[41] Jacopo Staiano, Bruno Lepri, Nadav Aharony, Fabio Pianesi, Nicu Sebe, and Alex Pentland. Friends don't lie: inferring personality traits from social network structure. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 321–330. ACM, 2012.

[42] Stephen V. Stehman. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1):77–89, 1997.

[43] Vladimir N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.

[44] M. Vollrath and S. Torgersen. Personality types and coping. *Personal. Individ. Differ.*, 29:367–378, 2000.