

# Happiness Recognition from Mobile Phone Data

Andrey Bogomolov  
SKIL Telecom Italia Lab,  
University of Trento,  
Via Sommarive, 5  
I-38123 Povo - Trento, Italy  
Email: andrey.bogomolov@unitn.it

Bruno Lepri  
MIT Media Lab (USA),  
Fondazione Bruno Kessler  
Via Sommarive, 18  
I-38050 Povo - Trento, Italy  
Email: lepri@fbk.eu

Fabio Pianesi  
Fondazione Bruno Kessler  
Via Sommarive, 18  
I-38050 Povo - Trento, Italy  
Email: pianesi@fbk.eu

**Abstract**—In this paper we provide the first evidence that daily happiness of individuals can be automatically recognized using an extensive set of indicators obtained from the mobile phone usage data (call log, sms and Bluetooth proximity data) and “background noise” indicators coming from the weather factor and personality traits. Our final machine learning model, based on the Random Forest classifier, obtains an accuracy score of 80.81% for a 3-class daily happiness recognition problem. Moreover, we identify and discuss the indicators, which have strong predictive power in the source and the feature spaces, discuss different approaches, machine learning models and provide an insight for future research.

## I. INTRODUCTION

The pursuit of happiness is a right that every American possesses according to the Declaration of Independence. To be happy is one of the major goals, if not the ultimate goal, of human beings. During the last decades, the consequences and benefits of happiness have also come into the focus of research. In a recent review, Lyubomirsky et al. [18] showed that happy people are successful in many life domains and that this success is at least partly due to their happiness. Happy people are more social, altruistic, active, like themselves and others. They have strong bodies and immune systems. For this reason, there was an explosion of research on happiness and subjective well-being. Several studies dealt with the problem of how to measure happiness. Generally, researchers relied on self-reports which are sometimes coupled with informant data, interviews by trained clinicians, unobtrusive observations of nonverbal expressions and physiological assessments (see [10]). Respondents typically are asked to rate their levels of positive and negative affect over a particular period of time or to make a judgment of their overall life quality.

The development of the mobile phones technology created a new pervasive way. Digital applications started to replace paper-based questionnaires to collect happiness data. For example, “Track your Happiness” [1] is a web based platform for tracking and visualizing user’s happiness. The main weakness of this tool is that the final measuring method is still based on self-reported surveys.

Instead, our paper proposes to use smartphone usage patterns, including social interactions, to recognize users’ daily happiness automatically. Nowadays, smartphones allow for unobtrusive and cost-effective access to previously inaccessible sources of data related to daily social behavior [17], [29].

Recently, the social psychologist Geoffrey Miller wrote “The Smartphone Psychology Manifesto” in which he argues that the smartphone should be taken seriously as new research tools for psychology. In his opinion, these tools could revolutionize all fields of psychology and other behavioral sciences making these disciplines more powerful, sophisticated, international, applicable and grounded in real-world behavior [22].

In this work we formulated the automatic recognition of daily happiness as a 3-class classification problem based on information concerning: a) peoples activity which is detected through their smartphones; b) weather conditions; c) personality traits. The first information type is represented by features extracted from call and sms logs and from Bluetooth hits, able to capture (i) amount of call, sms and proximity; (ii) diversity of call, sms, and proximity; and (iii) regularity in user behaviors. Concerning weather conditions, both commonsense and the literature suggest that they can have an important impact on daily happiness states. Finally, individual dispositions (as captured by personality traits) are expected to interact with situational and contextual aspects (including weather conditions) to play a role in determining daily happiness states.

Classification experiments are performed through a variety of approaches, including support vector machines, neural networks and random forests, yielding the accuracy up to 80.81%.

## II. RELATED WORK

The automatic recognition of happiness using smartphones has not been the main target of the research community so far. Some previous works tried to estimate only a few possible determinants of happiness. BeWell [16] monitors three daily behaviors (activity, social interaction, and sleep patterns), describes the effect of these behaviors on well-being, and provides some feedback to the user.

More recently, Muaremi et al. [25] proposed to use smartphones for automatic measurement of individual and social happiness level.

Other previous works tried to focus on mood assessment using mobile phones. Moturu, et al. [24] explored the associations between sleep, mood and sociability by analyzing mobile-phone-generated social communication data and self-reported mood and sleep data. Rachuri et al. [26] proposed EmotionSense, a novel system for social psychology study of user emotion based on mobile phones.

A significant work is the one conducted this year by LiKamWa et al. [27], in which the authors reported a smartphone software system, called “MoodScope”, able to infer the mood of its user, based on how the smartphone is used.

### III. DATASET: LIVING LABORATORY

We exploited a dataset, called “Friends and Family”, capturing more than eight complete weeks of 117 subjects living in a married graduate student residency of a major US university, collected between 21 February, 2010 and 16 July, 2011. During this period, each participant was equipped with an Android-based cellular phone incorporating a sensing software explicitly designed for collecting mobile data. This software runs in a passive manner and does not interfere with the normal usage of a phone [2]. The collected *smartphone activity data* consists of: i) call logs, ii) sms logs, iii) proximity data, obtained by scanning near-by phones and other Bluetooth devices every five minutes, and iii) self-reported surveys about personality traits (“Big Five”) and daily happiness. Social interactions were derived from Bluetooth proximity detection data in a manner similar to those in previous reality mining studies [11], [20]. In total, the source dataset includes 33497 phone calls, 22587 sms and 1460939 Bluetooth hits.

*Ground-truth data*, self-perceived daily happiness, was reported by the participants filling a seven item categorical scale with mapping: 1 = “being very unhappy”, 4 = “being neither unhappy, nor happy”, and 7 = “being very happy”. Most of the people for most of the time described themselves as having been from moderately to quite happy during the day ( $mean = 4.84, median = 5.00, skew = -0.39, kurtosis = -0.07$ ). The distribution of this scores has a negative skew – the density is moved to the higher region of happiness score. The distribution also has negative excess kurtosis which is called platykurtic distribution, that means that the focus group reported a specific daily happiness score more often, than neutral. Within person and between-subject variance for happiness scores (Fig. 1) show, that within-person daily happiness variance range is much wider than between-person. This property shows that within-person emotional state for each ground-truth class is seen. Following our hypothesis,

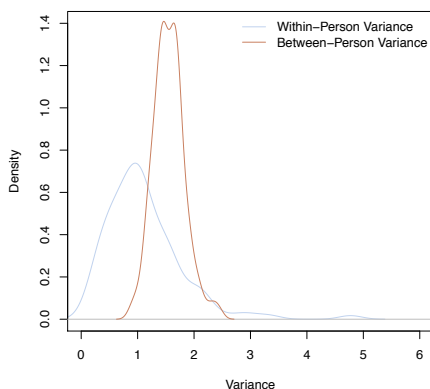


Fig. 1. Within person and between-subject variance

the “background noise data” we used is the personality data (“Big Five” personality traits score) and generally available daily weather data. Our hypothesis was supported by several

studies in social psychology [31], [19], [28], [5], [8]. Big Five personality traits were measured by asking the subjects to use 1-5 point scale to answer the online version of the 44 questions Big Five questionnaire developed by John et al. [13]. *Daily weather* and happiness relation is supported by [14] and [9] research results. In our experiments we used the following weather parameters: mean temperature, pressure, total precipitation, humidity, visibility and wind speed metrics.

### IV. FEATURE EXTRACTION

Taking inspirations from previous works on personality prediction using mobile phone data [6], [23], we derived a basic set of 25 call and sms features and 9 proximity features. The extracted features fall into four broad categories: (i) general phone usage, (ii) diversity, (iii) active behaviors and (iv) regularity.

*General phone usage* features capture the total number of outgoing, incoming and missed calls, the total number of sent and received sms.

We also extracted the outgoing to incoming calls ratio, missed to (outgoing + incoming) calls ratio and sms sent/received ratio. Concerning *regularity features*, we measured the time elapsed between calls, the time elapsed between sms exchanges and the time elapsed between call and sms. More precisely, we consider both the average and variance of the inter-event time of ones’ call, sms, call+sms. 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 different. *Diversity* measures how evenly an individual’s activity 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 and (iii) number of unique contacts. We compute the diversity features both for calls and sms. In particular, the entropy of an individual is the ratio between his/her total number of contacts and the relative frequency at which he/she interacts with them. The more one interacts equally often with a large number of contacts the higher the entropy will be. From the Bluetooth hits collected we filtered the events where  $RSSI = 0$ , assuming this case being the best proxy for close social proximity in space. RSSI is an 8-bit integer that denotes whether the received power level is within the Golden Receiver Power Range (GRPR). The lower and upper thresholds of GRPR are loosely bound, leaving them to be device specific. That means that RSSI is merely a relative parameter. The RSSI parameter is intended to be used for power control purpose. The receiver sends increase or decrease transmission power request to the transmitting side depending on whether the perceived RSSI level. A positive or negative RSSI, measured in dB, means the received power level is above or below the GRPR, respectively. A zero implies that it is ideal. From the filtered Bluetooth proximity data we extracted *Bluetooth proximity features* following three broad categories: (i) general proximity information, (ii) diversity, and (iii) regularity.

For each feature from our basic subset we calculated mean, median, min, max, quantiles with a linear step of 0.05, 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. Moreover, to capture the possible

influence of the previous happiness states on the current one we computed each basic subset of features using backward moving windows of different size (2, 3, 4 and 5 days). In order to get bias corrected empirical entropy estimates we applied *Miller-Madow correction* for entropy calculation [21], which is explained in Equation 1.

$$\hat{H}_{MM}(\theta) \equiv - \sum_{i=1}^p \theta_{ML,i} \log \theta_{ML,i} + \frac{\hat{m} - 1}{2N}, \quad (1)$$

where  $\hat{m}$  is a number of bins with nonzero  $\theta$ -probability. The likelihood function is given as the product of probability density functions  $P(\theta) = f(x_1; \theta) f(x_2; \theta) \cdots f(x_n; \theta)$  for a random sample  $X_1, \dots, X_n$ .  $\theta_{ML}$  is the maximum likelihood estimate of  $\theta$ , which maximizes  $P(\theta)$ .

## V. METHODOLOGY AND EXPERIMENTAL RESULTS

We formulated the automatic recognition of daily happiness as a classification problem with 3 classes (not happy, neutral and happy). The ground truth labels for classification problem were set to  $-1$  for "not happy" class, when the (*score*  $< 4$ ),  $0$  for "neutral" (*score*  $= 4$ ) and  $1$  for "happy" daily mood (*score*  $> 4$ ). We separated all the data at random following the uniform distribution in a training and a control test set in proportion of 80:20. To let optimization algorithms converge more efficiently the feature matrix was normalized by each column to  $[-1, 1]$  interval and centered to 0. Then, we trained a three families of classifiers: (i) Support Vector Machines, (ii) Neural Networks and (iii) Random Forests.

For *support vector machines model* [30] we used the following *decision function* for each class:

$$D(x) = \text{sign}\left(\sum_{i=1}^l y_i \alpha_i K(x_i, x) + b\right) \quad (2)$$

The classification problem was solved using *C-classification algorithm* as the optimization problem:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \alpha^\top \mathbf{Q} \alpha - \mathbf{e}^\top \alpha \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l, \\ & \mathbf{y}^\top \alpha = 0, \end{aligned} \quad (3)$$

where  $\mathbf{e}$  is the unity vector,  $C$  is the upper bound,  $\mathbf{Q}$  is a positive semidefinite matrix with  $l$  by  $l$  dimensions,  $Q_{ij} \equiv y_i y_j K(x_i, x_j)$  and the kernel:  $K(x_i, x_j) \equiv \phi(x_i)^\top \phi(x_j)$ .

Our multi-class classification problem is solved as a one-versus-all case by a winner-takes-all strategy, in which the classifier with the highest output function is assigned to the class. Linear kernel did not provide accurate enough results thus only Gaussian radial basis [4] kernel results are reported in the paper.

The second family of statistical models used to solve our classification problem is based on *random forest algorithm*. Random forests is 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 [3]. The decision boundary is formed according to the margin function. Given an ensemble of tree classifiers  $h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_K(\mathbf{x})$  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:

$$mg(\mathbf{X}, Y) = \text{avg}_k I(h_k(\mathbf{X}) = Y) - \max_{j \neq Y} \text{avg}_k I(h_k(\mathbf{X}) = j), \quad (4)$$

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:

$$PE^* = P_{\mathbf{X}, Y}(mg(\mathbf{X}, Y) < 0), \quad (5)$$

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:

$$I_A(x) = \begin{cases} 1 & \iff (x \in A) \\ 0 & \text{otherwise} \end{cases} \begin{cases} 1 & \iff \exists x \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Random Forests classifiers were trained with a stepwise 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 100.

The third approach for a solution is based on multi-layer perceptron *neural network*. In a parametric functional form it is expressed as:

$$f_\theta(\mathbf{x}) = S(\langle \mathbf{w}, \tilde{\mathbf{x}} \rangle + b), \quad (7)$$

where

$$\tilde{\mathbf{x}} = S(\mathbf{W}^{\text{hidden}} \mathbf{x} + \mathbf{b}^{\text{hidden}}), \quad (8)$$

with parameters

$$\theta = \{\mathbf{W}^{\text{hidden}}, \mathbf{b}^{\text{hidden}}, \mathbf{w}, b\}, \quad (9)$$

and the decision function is a sigmoid function

$$S = \frac{1}{1 + e^{-x}} \quad (10)$$

The happiness recognition problem was solved as the following optimization problem:

$$\hat{\theta} = \arg \min_{\theta} \hat{R}_\theta(f_\theta, D_n), \quad (11)$$

where the sum of empirical risk and regularization term is expressed

$$\hat{R}_\theta(f_\theta, D_n) = \sum_{i=1}^n L(f_\theta(\mathbf{x}^{(i)}), t^{(i)}) + \lambda \Omega(\theta). \quad (12)$$

Neural network training was performed using classical back-propagation algorithm. Neural network topology was iteratively searched for the best balance between empirical error minimization and maximum  $\kappa$  metric using 10-fold cross validation.

In order to select the final approach and the model, we compared these three families by means of a set of accuracy and  $\kappa$  metrics. The  $\kappa$  measures pairwise agreement among a set of functions which are making classification decisions with correction for an expected chance agreement [7]:

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)} \quad (13)$$

$\kappa = 0$  if there is no agreement more than expected by chance following the empirical distribution.  $\kappa = 1$  when there is a *max* agreement.  $\kappa$  is a state-of-the-art statistics about how significantly the classification model is different from chance. More importantly, it makes the interpretation of what what the model has learned – being an intuitive task. Also there is some critique in the literature of this metric, but we found it highly relevant for our model search and comparison task.  $\kappa$  statistic has the property of more robust measure than the simple percent classes agreement measure or the area under the ROC curve taking into account the agreement occurring by chance and not equally balanced classes in the training set.

During the learning and model selection process we used a random sampling with replacement to generate a new set of data for each fold from the basic training set following leave-one-out 10-fold cross validation scheme. We adopted this strategy in order to deal with potential data loss in cases when calls, sms and Bluetooth proximities existed in the real world but were not registered in the database.

Given the unbalanced classes of the ground truth daily happiness labels (that people were more happy than unhappy during the experiment) we found a solution to prevent data overfitting by the way we trained the models and searched in the possible classifier’s space. Our *structural risk minimization* as opposed to empirical risk minimization solution, to prevent the data overfitting, was incorporated by working with a regularization penalty into the learning process, by balancing the model’s complexity against its fitting the training data and by sampling of the model training sets in the way they mimic the empirical distributions without most probable erroneous outliers.

$\kappa$  metrics comparison for each model type using 10-fold cross validation strategy is provided in Table II. Random Forest classifier showed an average accuracy for this classification problem better than the best SVM-based model with radial basis kernel for 10.09% and better than neural network classifier for 5.56%. The conservative measure of agreement  $\kappa$  comparison explains the problem of unbalanced classes for an SVM classifier. SVM did good only for majority classes. As  $\kappa$  has a tendency to take the observed categories’ frequencies as given, which causes the effect of underestimating agreement for a category that is most commonly used. Neural network classifier captured relations between the variables in a pretty stable manner among the unbalanced classes, comparing to SVM. This best tradeoff between generalization capabilities of a neural network and high overfitting for a neural network is still less efficient in comparison to the Random Forest classifier results.

We found that distribution of the estimated performance metrics does not vary substantially among each fold, that means a good generalization despite of possible existence of heterogeneous data in each fold and the “noise” coming from the resampling procedure. Given the reason discussed above, as we trained the models on artificially generated data from each fold of the training sample, we have lower accuracy and  $\kappa$  results for the cross validation results than we get on the native data.

The final model is based on a Random Forests algorithm and uses a 111-dimensional feature vector. We identified the

TABLE I. 10-FOLD CV MODEL COMPARISON: ACCURACY

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
SVM	0.6404	0.6417	0.6434	0.6432	0.6442	0.6462
NeuralNetwork	0.6529	0.6606	0.6718	0.6708	0.6810	0.6865
RandomForest	0.6891	0.7000	0.7093	0.7081	0.7162	0.7247

TABLE II. 10-FOLD CV MODEL COMPARISON: KAPPA

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
SVM	0.0012	0.0056	0.0133	0.0115	0.0162	0.0226
NeuralNetwork	0.1381	0.1708	0.2055	0.1963	0.2173	0.2570
RandomForest	0.3094	0.3358	0.3565	0.3533	0.3733	0.3950

variables that have more predictive power for the Random Forest model by calculating mean decrease in Gini index.

This model shows 80.81% accuracy on the training set and 80.36% accuracy on the test set (Table V). In this table we provide the final happiness recognition model performance metrics and statistical estimates comparison. To show generalization power of the model, we calculated detailed metrics for each fold during 10-fold cross validation process.

TABLE III. FINAL CLASSIFIER HAPPINESS RECOGNITION MODEL CONFUSION MATRIX FOR TRAINING SET

	-1	0	1
-1	782	119	75
0	153	1170	145
1	600	903	6448

TABLE IV. FINAL CLASSIFIER HAPPINESS RECOGNITION MODEL CONFUSION MATRIX FOR TEST SET

	-1	0	1
-1	197	30	14
0	34	274	37
1	152	243	1616

The confusion matrices for this model for each training and test sets are provided in Table III and Table IV. These matrices show in details that there is a major agreement between classes. But Table IV shows that there are still a lot of misclassified samples as “unhappy” (152 out of 2597 samples) which are, in fact, the happy state for the actual class. 243 samples are classified as “neutral” being actually “happy”. The best result we have for the majority “happy” class (1616 correctly classified). These tables prove us good generalization capabilities of the model – for each the training and the test sets we have similar results.

To show our result visually in a simple form, what actually the final model delivers better than “by chance”, we report the multi-class area under the ROC curve [32], [12]. The final model ROC curve plotted for the test subset is provided in Fig. 2. The multi-class area under the ROC curve (AUC) = 0.844 means a good reduction of missclassification error for each class combination [15]. The area between the ROC curve and the diagonal is, in fact, what our model does better than “by change”. The area between the ROC curve and the perpendicular lines going to the upper left corner [1.0, 1.0] is the measure of what is not explained by our model, what causes the misclassification error.

## VI. DISCUSSION

The present study provides the first evidence that daily happiness of individuals can be automatically predicted using

TABLE V. PERFORMANCE METRICS

	Training set	Test set
Accuracy	0.8081	0.8036
Kappa	0.5879	0.5743
AccuracyLower	0.8004	0.7878
AccuracyUpper	0.8156	0.8187
AccuracyNull	0.6415	0.6419
AccuracyPValue	2.139e-303	8.826e-73
McNemarPValue	5.647e-208	1.738e-57

a set of indicators obtained from mobile phone data (call logs, sms logs and Bluetooth interactions) and additional indicators related to weather factors and individuals' personality traits. Our final model based on a Random Forest classifier using 111-dimensional feature vector obtains an accuracy measure of 80.81% for a 3-class classification problem.

We found that weather factor and personality traits, employed as a "filter" on a decision tree in the feature space, provide much better accuracy,  $\kappa$  and the area under the ROC curve metrics for the recognition model. But weather or personality traits data used as a single group of independent variables does not provide a feasible result.

These results are interesting because the previous studies in social psychology focus mainly on the associations between happiness or subjective well-being and Extraversion or Neuroticism. Instead, our work shows the important contribution played also by Conscientiousness, Openness and Agreeableness. Moreover, these results open us the possibility of creating a multi-step stochastic model in which first we estimate the personality and then we use the estimated variables as independent variables for a happiness recognition problem.

Regarding the role played by mobile phone data in predicting daily happiness, it's interesting to note the significant contribution of the proximity features. Among the top 30 features used for happiness recognition, 10 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. Social proximity, measured by Bluetooth interactions, has strong predictive power for daily happiness recognition based on median, 0.75, 0.9 and 0.95 quantiles from the distribution of the time intervals, measured in seconds, for which a Bluetooth identifier is seen.

Moreover, features capturing the diversity in co-location interactions are in the top 30 list (e.g. entropy of proximity contacts). This result seem to confirm previous studies in social psychology that found associations between people's happiness and the richness in terms of the amount and the diversity of people's social interactions. 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.

This result is consistent with previous studies on happiness showing that happy people are usually more social and have more interaction and exchange. In addition, a role of the regularity in call patterns (average and standard deviation values of the time elapsed between two calls) and active behavior of an individual (the ratio of calls initiated by the individual) also play predictive role. The latter indicator could be easily explained making reference to the standard definition of happy people as more active people given by social psychologists.

The role of sms interactions is less evident from our

investigation. The only feature related to sms interactions among the top 30 predictive features is a feature related to an active behavior of the individual and more specifically is the latency in replying to a text message that we define as the median value that people take to answer a text. The predictive power of the sms data needs further investigation.

Among the *limitations of the present study* we list the following: our sample comes from a population living in the same environment (our subjects were all married graduate students living in a campus facility of a major US university); the non-availability of proximity data concerning the interaction with people not participating in the data collection, a fact that is common to many other studies of this type. The first two problems are at least partially attenuated by the large variability of the sample in terms of provenance and cultural background, which can be expected to correspond to a wide palette of interaction behaviors that efficaciously counterbalance the effects of living-place homogeneity.

Moreover, the FUNF framework used for collecting this data did not account for the data loss in cases when calls, sms and Bluetooth proximities existed in the real world but were not registered in the database as "NA". The source data loss, when the battery is out, was also not registered and was unpredictable during the data collection. In order to solve these potential problems, we proposed and used a random sampling with replacement for each fold in our learning task thus generating a better set from the training dataset.

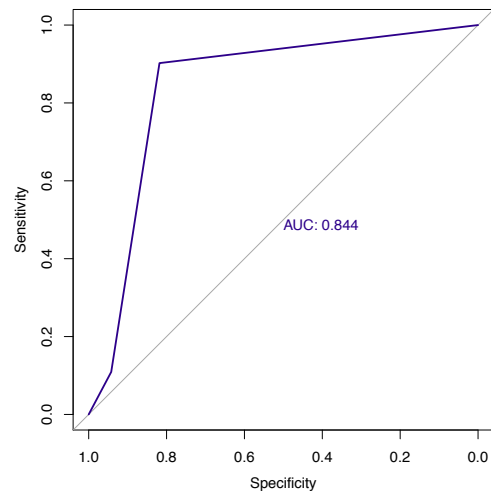


Fig. 2. Final Model ROC curve

## VII. CONCLUSION

The goal of our paper is to investigate the feasibility of automatic recognition of people's daily happiness from mobile phone data. To achieve this goal, we formulated the automatic recognition of daily happiness as a 3-class classification problem based on the information concerning: a) peoples activity, as detected through their smartphones; b) the weather conditions; c) personality traits. The first information type is represented by features extracted from call and sms logs and from Bluetooth hits, able to capture (i) amount of call, sms

and proximity; (ii) diversity of call, sms, and proximity; (iii) user active behaviors; and (iv) regularity in user behaviors.

Despite the limitations of this study discussed above, we believe that our results (80.81% for a 3-class classification problem) have provided compelling evidence that individual daily happiness can be reliably predicted from smartphone usage data and from additional indicators related to the weather factors and individual dispositions (personality traits). Hence, on a practical side our results are a first important step towards automatic systems able to predict people's daily happiness and towards engineering a construct that is not only a goal of people's life but also a means for reaching other goals and for facilitating the desirable behaviors and outcomes.

#### ACKNOWLEDGMENT

The work of Andrey Bogomolov was supported by Telecom Italia Semantics and Knowledge Innovation Laboratory (SKIL) with research grant T. The work of Bruno Lepri was supported by the Marie Curie COFUND – 7th Framework PERSI project.

We would like to thank MIT Media Lab for helping us with the data collection.

#### REFERENCES

- [1] Track your happiness web application. <http://www.trackyourhappiness.org>, Access date: 12-JUN-2013.
- [2] Nadav Aharoni, 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.
- [3] Leo Breiman. Random forests. *Mach. Learn.*, 45(1):5–32, October 2001.
- [4] Martin D. Buhmann and M. D. Buhmann. *Radial Basis Functions*. Cambridge University Press, New York, NY, USA, 2003.
- [5] Charles S. Carver and Michael F. Scheier. Origins and Functions of Positive and Negative Affect: A Control-Process View. *Psychological Review*, 97(1):19–35, January 1990.
- [6] 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.
- [7] Jacob Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37–46, April 1960.
- [8] Paul T. Costa and Robert R. McCrae. Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, 38:668–678, 1980.
- [9] 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.
- [10] Ed Diener. Assessing subjective well-being: Progress and opportunities. *Social Indicators Research*, 31(2):103–157, 1994.
- [11] Nathan Eagle and Alex Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.
- [12] David J. Hand and Robert J. Till. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning*, 45:171–186, 2001.
- [13] 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.
- [14] Matthew C. Keller, Barbara L. Fredrickson, Oscar Ybarra, Stephane Cote, Kareem Johnson, Joe Mikels, Anne Conway, and Tor Wager. A Warm Heart and a Clear Head: The Contingent Effects of Weather on Mood and Cognition. *Psychological Science*, 16(9):724–731, September 2005.
- [15] Wojtek J. Krzanowski. Misclassification rates. In *Encyclopedia of Statistics in Behavioral Science*. 2005.
- [16] Nicholas Lane, Mashfiqui Mohammad, Mu Lin, Xiaochao Yang, Hong Lu, Shahid Ali, Afsaneh Doryab, Ethan Berke, Tanzeem Choudhury, and Andrew Campbell. Bewell: A smartphone application to monitor, model and promote wellbeing. *IEEE*, 4 2012.
- [17] Nicholas D. Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T. Campbell. A survey of mobile phone sensing. *Comm. Mag.*, 48(9):140–150, September 2010.
- [18] Sonja Lyubomirsky, Laura King, and Ed Diener. The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin*, 131(6):803–855, November 2005.
- [19] Sonja Lyubomirsky, Chris Tkach, and Robin M. Dimatteo. What are the differences between happiness and self-esteem? *Social Indicators Research*, 78:363–404, 2006.
- [20] Anmol Madan, Manuel Cebrián, David Lazer, and Alex Pentland. Social sensing for epidemiological behavior change. In *UbiComp*, pages 291–300, 2010.
- [21] G. Miller. Note on the bias of information estimates. *Information theory in psychology: Problems and methods*, 2(95):100, 1955.
- [22] Geoffrey Miller. The smartphone psychology manifesto. *Perspectives on Psychological Science*, 7(3):221–237, 2012.
- [23] Yves-Alexandre Montjoye, Jordi Quoidbach, Florent Robic, and Alex (Sandy) Pentland. Predicting personality using novel mobile phone-based metrics. In Ariel M. Greenberg, William G. Kennedy, and Nathan D. Bos, editors, *Social Computing, Behavioral-Cultural Modeling and Prediction*, volume 7812 of *Lecture Notes in Computer Science*, pages 48–55. Springer Berlin Heidelberg, 2013.
- [24] Sai T. Moturu, Inas Khayal, Nadav Aharoni, Wei Pan, and Alex Pentland. Using social sensing to understand the links between sleep, mood, and sociability. In *SocialCom/PASSAT*, pages 208–214, 2011.
- [25] 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.
- [26] Kiran K. Rachuri, Mirco Musolesi, Cecilia Mascolo, Peter J. Rentfrow, Chris Longworth, and Andrius Aucinas. Emotionsense: a mobile phones based adaptive platform for experimental social psychology research. In *UbiComp*, pages 281–290, 2010.
- [27] Nicholas D. Lane Lin Zhong Robert LiKamWa, Yunxin Liu. Moodscope: Building a mood sensor from smartphone usage patterns. [http://niclane.org/pubs/mobisys\\_moodscope.pdf](http://niclane.org/pubs/mobisys_moodscope.pdf), 2013.
- [28] Cantor N. Sanderson, C.A. A life task perspective on personality coherence: Stability versus change in tasks, goals, strategies, and outcomes. In D. Cervone & Y. Shoda (Eds.), *The Coherence of Personality: Social-cognitive Bases of Consistency, Variability, and Organization*, pages 372–392, New York, NY, USA, 1999. Guilford Press.
- [29] Jacopo Staiano, Bruno Lepri, Nadav Aharoni, 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, UbiComp '12*, pages 321–330, New York, NY, USA, 2012. ACM.
- [30] Vladimir N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.
- [31] Joar Vitters. Personality traits and subjective well-being: emotional stability, not extraversion, is probably the important predictor. *Personality and Individual Differences*, 31(6):903 – 914, 2001.
- [32] Matthew S. Wandishin and Steven J. Mullen. *Multiclass ROC Analysis*. American Meteorological Society, 2009.