Leveraging Social Affect for Identifying Individual Mood

Elaheh Momeni
University of Vienna
Währinger Strasse 29, A-1090
Vienna
momeni@cs.univie.ac.at

Peter Kalchgruber
University of Vienna
Währinger Strasse 29, A-1090
Vienna
kalchgruber@cs.univie.ac.at

Daniela Ramsauer
University of Vienna
Währinger Strasse 29, A-1090
Vienna
ramsauer@cs.univie.ac.at

Reza Rawassizadeh
University of California
Riverside
900 University Ave, 309
Winston Chung Hall Riverside,
CA 92521
rezar@cs.ucr.edu

ABSTRACT

The PREventive Care Infrastructure based On Ubiquitous Sensing (PRECIOUS) project aims to develop a preventive care system to promote healthy lifestyles. One of the goals of the project is the development of a method and application for automatic identification of human mood. To this end, we hypothesize that, in addition to using smart pervasive artifacts, leveraging influential factors from social media signals for inferring individuals' moods may enhance the performance of the mood prediction process and furthermore, may reduce the total sparsity and uncertainty of information regarding this process. Accordingly, this position paper describes how our experiment was conducted and reports on our primary achievements for the development of a mood predictor of social media data. Furthermore, we report on the development of a wearable app which enables us to collect explicit feedback from users for conducting a study to evaluate our approach.

1. INTRODUCTION

The PREventive Care Infrastructure based On Ubiquitous Sensing (PRECIOUS) project [1] aims to develop a preventive care system to promote healthy lifestyles. It consists of three components: (1) "transparent sensors for monitoring user context and health indicators (food intake, sleep and activity) that deliver ambient data about current user behavior"; (2) "users are represented by individual virtual models, which infer health risks and suggest behavioral changes"; and (3) "state-of-the-art motivational techniques originating from gamification and motivational interviews to trigger a set of feedback to change the user habits toward more healthy behavior". Risk factors for type II diabetes as a central use case have been chosen for this project; however, other illnesses provoked by lifestyle and their risk factors can also be investigated using the developed prototype. The system will not only detect and communicate detailed early warning signs, but also provide forecasts of future developments and associated problems.

[1] http://www.thepreciousproject.eu/

One of the goals of the project is the development of a method and application for automatic identification of human mood, as the quality of emotions contributes significantly to the eating habits [5]. Pursuing this goal, sensors and applications of new smart pervasive artifacts (such as smart-phones and smart-watches) can capture diverse spatio-temporal data about an individual from various sensors and applications. The resulting personal life stream data can support powerful inference with regard to the individual's moods and behavior [8].

However, the main challenges of running personal life stream data collections and context-sensing applications are high energy consumption, uncertainty and sparsity of information. Many of these applications that require context information may occasionally need continuous or frequent context monitoring. On the other hand many users use social media platforms for social interactions and express their moods and activities via textual communication and social interactions. These also provide useful signals about the individual. To this end, we hypothesize that leveraging influential factors from social media communications for inferring moods may reduce the total sparsity of information and uncertainty of mood identification process.

For this purpose, this position paper describes the manner in which our experiment was set up and our primary achievements for the development of a mood classifier of social media data. We developed classifiers for moods, "Happy" and "Sad". In addition, as one of the main requirements of the project, we also developed a classifier for "Stress". We predict affective states from explicit mood-oriented sentences, collected via a mechanical turk study. These affect-labeled posts are used in a classification experiment to predict the affective state from posts. Our experimental results indicate a wide variation in classifier performance across different affects and classification algorithms – this may be the result from how patterns and styles when using language vary depending on affective states.

2. RELATED WORK

Recently, some works have investigated development of mood prediction models and methods from individuals' social media communications. For instance in [3] a web-based tool 'MoonPhrases' was created to enable Twitter users to reflect about their mood and well-being. Therefore the Tweets of a user are analyzed with LIWC (Linguistic Inquiry and Word Count[2]) to get information about positive-, negative affect and linguistic styles. These results were visualized with moons and plain text to enable the user to reflect about his/her mood in the past. A similar approach was taken in [4], it was investigated to improve the classification of Tweets in either positive, neutral or negative sentiment. Therefore three classifiers were used, then not the results but the probabilities (or confidences) of the classifiers for each class were compared. Finally the class with the highest average probability is chosen.

Moreover in [6] and [2] messages of Twitter users were interpreted to find out how it is talked about depression in Tweets and how the usage of sentiment words of a depressed person differ from a not depressed person. They have shown that depressed persons show lowered social activity, they post more about themselves and interact less with others. Also depressed persons seem to express more negative emotion with the use of words of certain affect categories such as anger, causation, tentative, communication, and friends. Furthermore Park et al. [6] revealed that depressed person tweet much personal information about their depression. This is also in agreement with the findings of Choudhury et al. [2] where they noticed a higher incidence of Tweets dealing about medical concerns. Moreover depressed users tweet more about relational concerns and religious thoughts than not depressed persons and tend to have smaller, tightly clustered close-knit networks with other users on Twitter.

Furthermore in [11] first a qualitative analysis of past Tweets about health or fitness was done. As a result a taxonomy was created, which classifies the post according to its activity and the sentiments expressed. Second qualitative interviews were done with Twitter users to find out about the motivation of posting. The interviews have shown that all participants didn't actively search for health or fitness related content. They came across the fitness community in Twitter and explored the content of the Tweets but didn't actively post in the beginning.

Nevertheless, our main focus in this project is leveraging social media communications for solving uncertainly and sparsity issues of data for prediction of individuals' mood.

3. EVALUATION FRAMEWORK

In order to investigate our hypothesis, we have developed an evaluation framework. The framework:

- Collects activities of individuals via smart pervasive artifacts using the open source applications UbiqLog [10] implicitly. The framework collects individual users' activities and social interactions in periodical rhythms: (1) social interactions such as text communication via social media platforms, lists of persons contacted, frequency of phone calls, number of communications per hour etc., and (2) physical activities such as mobility state per minute, geographical locations etc. Furthermore, the framework collects explicit feedback (in order to create a ground truth dataset) from users in periodical rhythms by: (1) asking users about their moods and behavioral states (happy, sad, stress, etc.); (2) asking users about current location (work, home, university, shopping, etc.); and (3) asking users about current activity (working, sport, driving, etc.)

- Predicts mood of individuals automatically with regard to three different settings: (1) using social media communications (2) using collected activities and locations, and (3) using both text communications and activities.

- Compares and evaluates performance of mood predictions in three different settings using explicit feedback collected from individuals.

3.1 Mood Classification via Social Media Communications

3.1.1 Data Acquisition

To create a ground-truth dataset from real world social media text communications, we performed a mechanical turk study[3]. We asked each turker to provide us five of their Facebook posts related to "Sad", "Happy", and "Stress" moods. In order to ensure the quality of the work by coders, we requested them to provide the Facebook profile address of popular persons (singer, sportsperson, etc.) and provide us five posts of the selected person with "Happy" tone. The goal was to detect possible inconsistencies and ensure that answers were specific and not given randomly. 100 turkers participated in this study and only those who had already received a total of Human Intelligence Tasks (HITs) higher than 5000 and HIT Approval Rate higher than 98% were accepted. In total, we collected 1500 posts related to three moods (500 happy, 500 sad, and 500 stress). It is important to note that to be allowed to participate turkers had to be active users of social media platforms, such as Facebook, and provide us their user ID.

3.1.2 Experimental Set Up

For developing mood classifiers, we first set up two sets of features for classification:

1. **Text-based Features (TB)**: We use a standard classification feature setup that is common in text and sentiment classification. Posts are represented as vectors of unigram and bigram features. Before feature extraction, posts are lowercased, URLs are removed, and numbers are normalized (canonical form). Next,

[2]http://www.liwc.net/

[3]http://aws.amazon.com/mturk/

feature reduction takes place. First, features that occur fewer than five times are removed. Second, features are subsequently reduced to the top 120 features in terms of log likelihood ratio.

2. **Psycholinguistic Features (PLB)**: We utilized an established source of text analysis dictionary, Linguistic Inquiry and Word Count (LIWC), to develop a set of features. LIWC was demonstrated by previous work [7] as a useful resource for identifying emotions of user-generated content. For LIWC, we used affective-indicative categories like positive/negative emotions, anxiety, sadness, and anger.

Subsequently, we chose three classifier algorithms to evaluate their performance for mood identification: Logistic Regression (LR), Support Vector Machine (SVM), and Bayesian network (BN) classifier. Also, for developing the mood classifier, we used two modeling approaches: (1) Balanced binary class for each mood, meaning three separated classifiers were developed for each mood. (2) Multi-class for all moods, meaning a classifier for predicting all moods with multi classes was developed. For analyzing the influence of the different sets of features on their performance, each classifier was set with all combinations of the feature sets and they were evaluated against each other. Finally, to evaluate the performance of the classifiers, we used four measures: precision (P), recall (R), F1-measure (the harmonic mean between precision and recall) and AUC (Area Under Curve) the Receiver Operator Curve (ROC).

3.1.3 Preliminary Experimental Results

Classification results for different modeling approaches and various combinations of features are given in Table 1-3. The results demonstrate the effectiveness of using text-related features for inferring individual moods. Nevertheless, for all moods, training a classification model using both sets of features shows improved performance compared to the same models trained using one set of features. Surprisingly, we observe that the Bayesian network classifier performs best for almost all moods with different combinations of features. The best performances are observed for the mood "Stress", while the worst are for "Sad". More precisely, in the case of the "Stress" mood, we are able to achieve an F1 score of 0.88, coupled with high precision and recall, when using the Bayesian network classifier in combination with all the features. Similarly, for the same setting, we achieve an F1 score of 0.86, coupled with also high precision and recall for "Happy" mood. However, we find a lower level of F1 score (0.79) when using the same classifier for "Sad" mood, but it is still the best performance setting for this mood. With regard to binary or multi class modeling, we observe that the binary class classifier using both sets of features outperforms other models and, in particular, outperforms models with binary classes using only one set of features.

As the text-related features play an important role for mood classification, we computed a ranked list of terms from a set of 1500 posts for each mood (500 posts for each mood) as an illustrative example. For ranking the terms, we used the Mutual Information (MI) measure from the information theory which can be interpreted as a measure of how much the joint distribution of features X_i (terms in our case) deviate from a hypothetical distribution in which features and categories are independent of each other. Table 4 shows the top 20 terms extracted for each category. Obviously, many of the "Happy" posts contain terms expressing sympathy or commendation. "Sad" posts, on the other hand, often contain negative adjectives.

Features	Classifier	Binary-Class				Multi-Class			
		P	R	F1	ROC	P	R	F1	ROC
TB	LR	0.81	0.82	0.80	0.84	0.50	0.71	0.59	0.84
	SVM	0.84	0.83	0.81	0.68	0.48	0.82	0.61	0.78
	BN	0.86	0.86	0.85	0.91	0.60	0.98	0.70	0.91
PLB	LR	0.81	0.81	0.81	0.85	0.58	0.72	0.64	0.86
	SVM	0.81	0.82	0.79	0.67	0.58	0.70	0.63	0.82
	BN	0.80	0.79	0.79	0.82	0.53	0.75	0.63	0.84
Both	LR	0.85	0.85	0.85	0.91	0.65	0.73	0.71	0.92
	SVM	0.85	0.85	0.84	0.75	0.62	0.77	0.69	0.86
	BN	**0.88**	**0.86**	**0.86**	**0.94**	0.70	0.91	0.80	0.96

Table 1: Results from the evaluation of classification algorithms for "Happy" mood

Features	Classifier	Binary-class				Multi-class			
		P	R	F1	ROC	P	R	F1	ROC
TB	LR	0.79	0.80	0.79	0.76	0.53	0.49	0.50	0.76
	SVM	0.79	0.80	0.79	0.76	0.63	0.51	0.56	0.73
	BN	0.87	0.87	0.88	0.87	0.87	0.73	0.80	0.96
PLB	LR	0.67	0.73	0.68	0.67	0.41	0.42	0.43	0.68
	SVM	0.62	0.74	0.64	0.50	0.42	0.48	0.45	0.66
	BN	0.74	0.71	0.71	0.77	0.42	0.44	0.42	0.69
Both	LR	0.77	0.78	0.77	0.78	0.56	0.54	0.55	0.81
	SVM	0.80	0.81	0.79	0.67	0.58	0.56	0.57	0.75
	BN	**0.89**	**0.89**	**0.88**	**0.92**	0.87	0.76	0.81	0.96

Table 2: Results from the evaluation of classification algorithms for "Stress" mood

3.2 Explicit Feedback Collector via Wearable Device

In order to evaluate and compare usage of different signals from various channels we require explicit feedback from the user. Therefore, we implemented a simple smartwatch data collection application (Figure 1). Both smartphones and wearable devices (e.g. smartwatches) can be used to collect information on human behavior. However, wearables have a higher potential for gathering more personal information due to their close proximity to users. In particular, devices such as fitness trackers and smartwatches have two major advantages over smartphones, being constantly connected to the skin and located on the body. As a result, these features make them more capable than smartphones of collecting physiological and explicit data [9].

The explicit feedback collector app collects users' moods and locations as explicit inputs within the users' physical activities as implicit inputs. In a more technical sense, if the user shakes the watch, then a pop-up appears and allows them to enter their mood manually. This app enables users to

Features	Classifier	Binary-class				Multi-class			
		P	R	F1	ROC	P	R	F1	ROC
TB	LR	0.77	0.78	0.76	0.75	0.50	0.40	0.44	0.75
	SVM	0.78	0.79	0.75	0.61	0.53	0.43	0.48	0.72
	BN	0.83	0.83	0.83	0.87	0.86	0.61	0.71	0.93
PLB	LR	0.71	0.75	0.70	0.68	0.45	0.34	0.39	0.68
	SVM	0.65	0.74	0.64	0.50	0.43	0.30	0.35	0.63
	BN	0.73	0.76	0.71	0.66	0.47	0.27	0.34	0.68
Both	LR	0.76	0.77	0.76	0.79	0.56	0.56	0.56	0.83
	SVM	0.78	0.79	0.77	0.65	0.56	0.57	0.57	0.77
	BN	**0.80**	**0.74**	**0.79**	**0.91**	0.75	0.76	0.75	0.95

Table 3: Results from the evaluation of classification algorithms for "Sad" mood

Stress	Sad	Happy
hours	missing	halloween
late	sad	happy
nervous	miss	awesome
times	rip	favorite
anxiety	sick	fun
test	friend	excited
phone	cold	today
tomorrow	pretty	finally
won't	lost	birthday
minutes	damn	weekend
stressed	ugh	enjoying
break	sucks	dinner
work	put't	cool
start	sad	Love
hour	put	mom
missing	watching	birthday
makes	hot	great
coffee	snow	amp
omg	rain	party
suck	give	blast

Table 4: the top-20 terms extracted for each mood category

explicitly enter their current mood, location (home, work, leisure), and activity four times per day. Physical activity terms are inspired by Google Play services and mood annotation terms are derived from the Circumplex affect model, which contains two orthogonal dimensions: pleasure (from sad to happy), and activeness (from sleepy to aroused).

4. SUMMARY AND FUTURE WORK

In order to investigate the reduction of the total sparsity of information and uncertainty of the mood identification process exploiting social media text communications, this paper describes our experimental set up and our primary achievements for the development of a mood predictor. Furthermore, it reports the development of a wearable app for collecting explicit feedback from users in order to evaluate our approach. We will work on the following steps in future work: (1) Extension of collecting posts with regard to other moods using the circumflex model of affect and collecting posts from other social media platforms such as WhatsApp and Twitter. (2) Development of mood predictor using only signals from smart pervasive artifacts, leveraging available results in related work [1, 8]. (3) Setting up a within subjective study using collected explicit data and developed classifiers in order to investigate impact of predicting moods using social media text communications.

Acknowledgement

The PRECIOUS project has received funding from the European Community's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 611366.

Figure 1: User interface of the prototype implementation of explicit feedback collectors on smartwatch

5. REFERENCES

[1] A. Bogomolov, B. Lepri, M. Ferron, F. Pianesi, and A. S. Pentland. Daily stress recognition from mobile phone data, weather conditions and individual traits. In *Proceedings of the ACM International Conference on Multimedia*, MM '14, pages 477–486, New York, NY, USA, 2014. ACM.

[2] M. d. Choudhury, M. Gamon, S. Counts, and E. Horvitz. Predicting depression via social media. In *Seventh International AAAI Conference on Weblogs and Social Media*, 2013.

[3] M. d. Choudhury, M. Gamon, A. Hoff, and A. Roseway. "moon phrases": A social media faciliated tool for emotional reflection and wellness. In V. Osmani, A. T. Campbell, and P. Lukowicz, editors, *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th Int. Conference on*, 2013.

[4] M. Hagen, M. Potthast, M. Büchner, and B. Stein. Twitter sentiment detection via ensemble classification using averaged confidence scores. In A. Hanbury, G. Kazai, A. Rauber, and N. Fuhr, editors, *Advances in Information Retrieval*, volume 9022 of *Lecture Notes in Computer Science*, pages 741–754. Springer International Publishing, 2015.

[5] M. Macht. How emotions affect eating: A five-way model. *Appetite*, 50(1):1 – 11, 2008.

[6] Minsu Park, Chiyoung Cha, Meeyoung Cha, and Yoorim Kweon. Depressive moods of users portrayed in twitter. *Telecommunications Review*, Jun. 2013.

[7] E. Momeni, C. Cardie, and M. Ott. Properties, prediction, and prevalence of useful user-generated comments for descriptive annotation of social media objects. In *Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM2013)*, Boston, USA, June 2013. AAAI.

[8] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas. Emotionsense: A mobile phones based adaptive platform for experimental social psychology research. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, UbiComp '10, pages 281–290, New York, NY, USA, 2010. ACM.

[9] R. Rawassizadeh, B. Price, and M. Petre. Wearables: Has the Age of Smartwatches Finally Arrived? *Communications of the ACM*, 58(1):45–47, 2015.

[10] R. Rawassizadeh, M. Tomitsch, K. Wac, and A. M. Tjoa. UbiqLog: A Generic Mobile Phone-based Life-log Framework. *Personal and Ubiquitous Computing*, 17(4):621–637, 2013.

[11] R. Teodoro and M. Naaman. Fitter with twitter: Understanding personal health and fitness activity in social media. In *Seventh International AAAI Conference on Weblogs and Social Media*, 2013.