Using Emotions to Predict User Interest Areas in Online Social Networks

Yoad Lewenberg
School of Computer Science and Engineering
The Hebrew University of Jerusalem, Israel
Email: yoadlew@cs.huji.ac.il

Yoram Bachrach
Microsoft Research
Cambridge, United Kingdom
Email: yobach@microsoft.com

Svitlana Volkova
Center for Language and Speech Processing
Johns Hopkins University,
Baltimore MD 21218, USA
Email: svitlana@jhu.edu

***Abstract*—We examine the relation between the emotions users express on social networks and their perceived areas of interests, based on a sample of Twitter users.**

Our methodology relies on training machine learning models to classify the emotions expressed in tweets, according to Ekman's six high-level emotions. We then used raters, sourced from Amazon's Mechanical Turk, to examine several Twitter profiles and to determine whether the profile owner is interested in various areas, including sports, movies, technology and computing, politics, news, economics, science, arts, health and religion.

We find that the propensity of a user to express various emotions correlates with their perceived degree of interest in various areas. We present several models that use the emotional distribution of a Twitter user, as reflected by their tweets, to predict whether they are interested or disinterested in a topic or to determine their degree of interest in a topic.

I. Introduction

Social networks, such as Facebook, Google+ and Twitter are becoming increasingly important methods of communication and social interaction between people. Social media is also a prominent economic domain [1], with social networks becoming one of the most important advertising platforms [2], [3], [4].

The huge numbers of people using such services and the enormous volume of data readily available from such networks make them a very compelling research target for social scientists and data-mining experts. Such researchers have examined many issues using social network data, including characterizing user attributes [5], [6], [7], the emotions and sentiments users express [8], [9], [10], [11], and the language they use to express themselves [12], [13], [14], [15].

Much of the revenue of social networking websites comes from online advertising [16], [17], [18]. A key advantage of online advertising over traditional forms of advertising is the ability to target very specific audiences, or even personalizing the advertising content based on the characteristics of the target user [19], [20]. Similarly, many recommender systems attempt to provide users with *personalized* recommendations on items such as movies, books or music, based on a profile of the user [21]. One approach is collaborative filtering, where the system generates a recommendation based on items consumed by users similar to the target user in terms of the items they have consumed in the past [22]. Earlier work already discusses method of building a "fingerprint" of a user in collaborative filtering systems based on the items they have previously consumed, typically using a dimensionality reduction or matrix factorization approach [23], [24], [25], [26]. [1]

Such dimensionality reduction techniques and fingerprinting methods offer a succinct way to represent a person so as to provide them good recommendations. However, they have a key disadvantage: the way they represent the preferences of an individual has no clear and easily interpretable meaning. Many marketing or advertising managers are interested in a characterization of a user that is easily understandable to humans, including features such as their demographic traits, social characterization or hobbies.

Clearly, an important piece of information regarding a consumer for which we are trying to personalize advertising content is their interest areas [33], [34], [35], [36]. For example, a person interested in technology is likely to be a better advertising target for a company selling computer gadgets than a person who has no interest in technology. Unfortunately, information regarding the interests of a user is not always readily available, so it needs to be inferred from other information regarding the consumer. One possible source of information which can form the basis for such inference are the *emotions* projected by a user of an online social network.

Our contribution: We investigate the relation between the emotions expressed by users in online social networks and their areas of interest.

We use a sample of target users from Twitter, and apply machine learning modules to classify their tweets according to the emotions expressed in them. We use the opinions of users sourced from Amazon's Mechanical Turk to tag these target users with their areas of interest. We then correlate the distribution over the emotions expressed by a user and their perceived areas of interest.

Our results indicate that there are indeed correlations between the emotions expressed by a user and their areas of interest. We show how to use these correlations to build a machine learning model that can predict the interests of a user given the emotions they express in their social network profile.

[1] Some such earlier work also describes various ways to reduce the space and time complexity of such fingerprints, at the expense of their recommendation quality [27], [28], [29], [30], [31], [32].

Finally, we discuss some of the implications of our work to online advertising.

II. Related Work

We briefly discuss some related work dealing with emotion analysis in texts and social media, and personal analytics based on online social network information.

Emotion detection in social media

Our analysis is based on the expression of *emotions* in social media. Similarly to *sentiments*, emotions are affective states, and are states of consciousness where a person experiences internal sensations.

Analysis of emotions has been applied to various forms of short texts, including instant messaging, emails and blogs [37], [38], [39], [40]. Recent work has also examined emotions in social media [41], [42], [43].

We use the emotion classification system proposed by Ekman [44], which captures six high-level emotions: joy, sadness, surprise, anger, fear and disgust. Researchers have used supervised machine learning methods to classify the emotions expressed in several forms of texts using this framework [45], [46], [47], [43].

Online Social Network Analytics

Earlier work investigated using information from social media to infer user properties, such as gender [48], [49], [50], [51], [52], [10], age [53], [54], ethnicity [55], personality [6], [9] [2] and even political and sexual preferences [56], [7].

Some such methods use not only the information from the target's profile and posted texts, but also use information about the network structure (such as their neighbors in the social graph) [61], [62], [63], [64].

In contrast to such earlier work, we examine the interest areas of the target user, rather than their demographic traits, and we focus on the relation between these interests and the emotions expressed by these users.

III. Data

We begin by describing the data used in our study, including our sample of Twitter users, and the annotations regarding the users' expressed emotions and areas of interest.

A. Twitter User Sample

Our study is based on a set of Twitter users, sampled at random from the Twitter data feed.

As Twitter contains many accounts that belong to celebrities or commercial entities, rather than real users, we have first made sure that our targets were standard users. To achieve this, we have hired workers from the Amazon's Mechanical Turk platform to examine each potential target, and let us know whether the account belongs to a normal user or whether it is a fake account, an account of a commercial entity, an account of a celebrity or some other non-standard user.

Each potential target account was evaluated by at least three crowdsourced reviewers, and was only selected to participate in our set of target users if all reviewers tagged the account as an account belonging to a normal user. Following this screening, we were left with a target account set of $|U^t| = 891$ users. We have extracted the Tweets of all the users U^t, amounting to $|T^t| = 2,644,940$ tweets.

B. Emotion Annotations

In order to study the relation between expressed emotions and areas of interest we need to tag each tweet of each of our sampled users with a tag capturing the emotion expressed in the tweet. As we are using Ekman's emotion framework, each tweet can be tagged by one of Ekman's six high level emotions: joy, sadness, surprise, anger, fear and disgust.

One possible way of obtaining emotion annotation is asking crowdsourced workers to examine each tweet and tag it with the emotion contained in it. However, this is extremely costly, due to the large number of tweets in our data. We thus use an alternative approach, and train a machine learning model using a distant supervision technique based on hashtags. Hashtags allow a Twitter user to enter a "keyword" associated with their tweet, expressing and idea or issue the tweet relates to. Many users use hashtags to express their emotion (for example *"I just got fired from my job! #sad"*), and such emotion hashtagged tweets can be leveraged to construct a dataset for training machine learning models. Similarly, many users use emoticons, such as ":-)" or ":-(", to express their feelings, so we can use emoticons in a similar way to hashtags.

We construct an emotion annotation dataset, which is similar to existing emotion tag datasets that rely on hashtags [45], [43]. We first build a list of emotion hashtags. We start with the hashtags of the Ekman's six emotions (i.e. #joy, #surprise and so on). We then expand the list with synonyms and derivations of these emotions. To do this, we compile a synonym list using WordNet-Affect, Google Synonyms and Roget's Thesaurus. The expanded list contains, for example, #miserable (relating to sadness) or #afraid (relating to fear). In total the expanded emotion hashtags list contains 360 emotion hashtag and emoticons.

After compiling the emotion hashtags list, we sample random tweets from Twitter's 1% feed, and select those that contain one of the emotional hashtags. Each such tweet forms a labeled datapoint, with the tweet, consisting of the list of words comprising the tweet, and the emotional label (the emotion the hashtag refers to). This results in an tweet-emotion dataset, consisting of over 50,000 datapoints (tweet-emotion pairs).

The machine learning model we use to tag tweets with emotions is a log-linear model. The model is trained on our tweet-emotion dataset, using the scikit-learn toolkit [65].[3] We have chosen a logistic-regression model and not alternatives such as support vector machines or various forms of neural networks as it can be quickly trained, and as earlier work on predictive analytics in social media [66] shows it has a good performance.

[2] Personality also correlates with other "digital footprints", such as browsing and website preferences or product purchase behavior [57], [58], [59], [60] but the prediction accuracy is generally higher when using data from online social networks.

[3] Available on http://scikit-learn.org/

Given a tweet $t \in T^t$, we denote by $E(t)$ the dominant emotion expressed in the tweet. A logistic regression classifier takes a tweet t as input and outputs a probability distribution over the possible emotions (a result of a softmax operator). We denote by $\mathrm{P}(E(t) = e \mid t)$ the probability that the expressed emotion is e given the tweet t. The predicted dominant emotion of the tweet is the emotion maximizing this probability:

$$\Phi_E(t) = \mathrm{argmax}_e \mathrm{P}(E(t) = e \mid t) \quad (1)$$

The features used as inputs for the emotion classifier are:

- LEXICAL FEATURES: Binary unigram bag-of-word features (the classification quality did not improve when using n-grams of higher orders, or count based features such as normalized frequency counts rather than binary features).
- HASHTAG EMOTION LEXICON FEATURES: the scores for unigram features of the Hashtag Emotion Lexicon (see details regarding these features in [43]).
- STYLE BASED FEATURES: The existence of elongated words (such as *Yeeeeeeees, awaaaaaaaaay*; fully capitalized words (such as *PLEASE, AMAZING*); emotions (of various forms, including both positive and negative ones), Mixed or repeating punctuation marks (such as *??!, !!!*); the number of hashtags.
- NEGATION FEATURES: transporting a unigram feature to a modified one to words appearing between a negation and a clause-level punctuation mark [67].
- PART OF SPEECH FEATURES: part-of-speech tags obtained using from the Twitter Part-Of-Speech tagger.[4]

The quality of emotion identification varies slightly between the different emotions. Using a 10 fold cross-validation, we obtained the following F_1 scores: disgust 0.92, anger 0.8, joy 0.79, fear 0.77, surprise 0.64, sadness 0.62 (an improved performance for all emotions over recently developed emotion classifiers [45], [43].

The trained emotion classifier $\Phi_E(t)$ can take a previously unobserved tweet, and tag it with the emotion expressed in the tweet. As this is a logistic regression, we actually obtain the probability of the tweet belonging to each of the six emotion classes. When required to generate a single tag, we can simply return the most probable emotion given the tweet (Equation 1).

Interest Areas Annotations

Our analysis of the relation between the emotions a Twitter user expresses and their areas of interest requires annotating each target user in our Twitter sample U^t with their areas of interest. We have selected the following possible areas of interest: sports, movies, technology computing, politics, news, economics, science, arts, health and religion. We denote the set of potential interests as I.

The degree to which a target user is interested in each of these areas was determined using raters crowdsourced from Amazon's Mechanical Turk. Each target was examined by at least three different raters. The average number of raters per target was 3.49, with a standard deviation of 0.783. Each rater had to examine the profile, and rank the degree of interest of the target in each of our interest areas, on a 5-level scale: completely uninterested, somewhat uninterested, neutral (no strong indication regarding interest or disinterest in profile), somewhat interested, very interested. Annotators received a constant payment for each profile they examined. To avoid annotators providing random answers, we have included gold-set multiple choice questions (such as "what is 2+2?"), and annotators were excluded if they failed these questions. [5]

We denote the set of annotators by A. We refer to all the interest ratings given by a single annotator to a single target user as an interest rating row. In total, our dataset contains $r = 3,442$ interest annotation rows, given by 691 distinct annotators.

IV. METHODOLOGY

We now describe our methodology. Our goal is to determine whether it is possible to predict a social network user's areas of interest from the emotions they express in the social network. Clearly, any single tweet made by a user may express a different emotion (or even multiple emotions), resulting in a very highly dimensional data. One possibility for simplifying the space is to *aggregate* the emotional tags across *all* the tweets generated by a target user.

A. User Emotion Scores

One method for aggregating emotions into a single score per emotion is simply characterizing a user by the *proportion* of their tweets containing a specific emotion.

Our emotion classifier is based on logistic regression, so it actually outputs a *probability distribution* over the possible emotions. Given a tweet i, we denote the probabilities assigned to each emotion expressed in the tweet as $p_i^{joy}, p_i^{surprise}, \ldots, p_i^{disgust}$.

Denote the set of emotions as $E = \{\text{joy, sadness, surprise, anger, fear, disgust}\}$. We define the *dominant* emotion expressed in a tweet i as the mode of the emotion distribution returned by our classifier, i.e. the emotion:

$$\arg\max_{e \in E} p_i^e$$

Denote the set of tweets of a given user $u \in U^t$ as T^u. The *proportion score* of an emotion $e \in E$ for a user u, denoted by q_u^e, is simply the proportion of u's tweets where the dominant emotion is e.

Another alternative to aggregate the emotions expressed in a user tweet is by using the probabilities assigned by our classifier (rather than only the mode of the distribution). The *distribution score* of an emotion $e \in E$ for a user u, denoted by s_u^e, is the average probability assigned by our emotional classifier to that emotion across all of u's tweets:

$$s_u^e = \frac{1}{|T^u|} \cdot \sum_{i \in T^u} p_i^e$$

[4]http://www.ark.cs.cmu.edu/TweetNLP/

[5]Alternative designs may reward participants based on their performance, or using a contest between annotators [68], [69], [70], but have a more elaborate structure. As interest areas are a somewhat subjective judgment, we chose not to employ such more intricate designs.

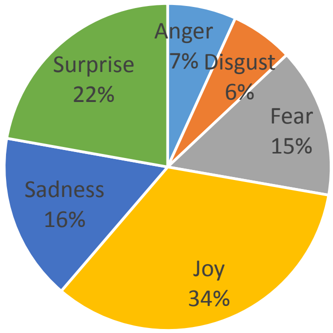


Fig. 1: Mean emotion proportion scores for Ekman's six basic emotions

Thus, the proportion score assigns a single point to the dominant emotion in a tweet and zero points to the remaining emotions, while the distribution score distributes a single point between the different emotions, according to the probability assigned by our emotional classifier. Both scores are normalized by the number of user tweets, so they fall in the range $[0, 1]$.

Figure 1 shows the emotion proportion scores for Ekman's six basic emotions, averaged across all users in our data. It shows that on average, users express some emotions much more than others. The most prevalent emotion was joy, followed by surprise. Negative emotions were more rare, with disgust being the least expressed emotions - only 6% of the tweets express disgust.

Figure 1 only presents the mean proportion scores for the various emotions. Our goal is predicting user interest using their relative propensity of expressing various emotions. This goal can only be achieved if users differ in their emotion scores. We have thus also examined the distribution of emotion scores across users.

Figure 2 presents a histogram of the emotion scores across users (the x-axis is the proportion emotion score bin, and the y-axis is the proportion of users in our dataset exhibiting this emotion score).

Figure 2 shows a large variability in emotion scores across users, especially for the more prevalent emotions. Interestingly, the histograms are quite different from a normal distribution. For example, the distribution of joy has a long right tale, indicating that quite a few users express joy very often. In contrast, the sadness distribution has a long left tale, indicating there are quite a few users who express very little sadness, but having a user who expresses a very large proportion of sad tweets is rare (despite this being a relatively common emotion).

B. User Interest Score

Each sample Twitter user $u \in U^t$in our dataset was examined by several annotators sourced from Amazon's Mechanical Turk. Denote the set of annotators who examined a target user u as A^u. For each interest area $i \in I$, each annotators $a \in A^u$ assigns an interest score, $r^i_{a,u}$ reflecting the degree to which a believes u is interested in i.

The possible interest scores for $r^i_{a,u}$ are $\{0, 1, 2, 3, 4\}$ (where 0 stands for "completely uninterested", 1 stands for "somewhat uninterested" and so on in our 5-level scale, ending with 4 which stands for "very interested").

The interest level score for a given target user $u \in U^t$ for the interest area $i \in I$ is simply the average interest score assigned by the raters to that user and interest: [6]

$$k^i_u = \frac{1}{|A^u|} \cdot \sum_{a \in A^u} r^i_{a,u}$$

Figure 3 presents a histogram of the user interest scores across all users in our dataset. It shows that some topics present a high variability across users in their interest in the topics, while others are more uniform. For example, there are many users who are either very interested or very disinterested in sports, whereas most users seem to have a moderate degree of interest in travel an leisure.

Further, some topics attract, on average, more user interest than others. For example, very few users are very interested in business and economics, whereas a large portion of the users are at least moderately interested in movies and television.

C. Interested and Disinterested Users

We may sort the users according to their degree of perceived interest in an area. For example, we can sort the users by their degree of interest in sports, k^{sports}_u, from the user perceived to be the most interested in sport to the user perceived to be the least interested in sports. We partition this sorted list of users into five parts (each containing 20%) of the users; We refer to the first part as "users who are *interested* in the area", and the last part as "users who are *disinterested* in the area"; the three middle parts are users who are neither very interested in the area nor very disinterested in it.

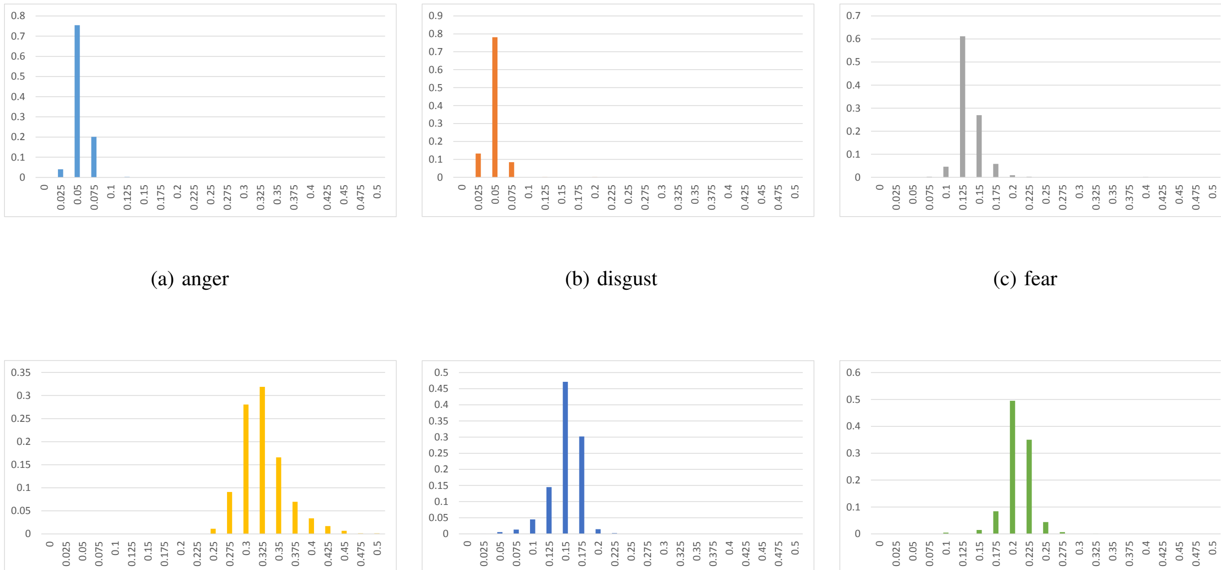
While we may refer to a subset of users as users who are "interested in sports" or "users who are disinterested in sports", it is important to remember that this is not based on an objective measure (such as the amount of time they spend browsing sports related websites), or on self-reports by these users. Rather, these are the results of impressions these users leave on our set of crowdsourced annotators, i.e. they are based on perceptions about this users.

D. Relating Emotions and Interest Areas

We use several tools to study the relation between the emotions expressed by users and their interest areas. First, we apply statistical tests to determine whether there is indeed a statistically significant relation between a user's propensity to express a certain emotion and their degree of interest in various areas. To achieve this, we rely on the Mann-Whiteney U test.

After testing for statistical significance of each emotion in isolation, we attempt to generate a prediction regarding a user's

[6]We note that more sophisticated methods, such as Bayesian methods or probabilistic graphical models, can aggregate the opinions of multiple raters while taking into account different performance levels of the annotators [71], [72], [73], [74]. Nonetheless, methods such as majority vote or average scores have been show to still achieve high performance in many tasks [75], [76]. We chose the method of average scores for its simplicity.



(a) anger (b) disgust (c) fear

(d) joy (e) sadness (f) surprise

Fig. 2: Histogram of the emotion scores across different users

areas of interest, based on the emotions they express. More concretely, we build a logistic-regression model which takes a user's emotion scores for each of Ekman's six emotions (either the proportion score or the distribution score), and determine whether that user is interested or disinterested in some area of interest. We then evaluate the quality of these prediction model, by examining the area under the ROC curve (AuC).

V. Results

We begin by checking whether users who are interested in an area tend to express emotions that are different than those not interested in an area. For each interest area i and each emotion e, we first examined the Pearson correlation between the interest scores of a user in the area i and their propensity to display a certain emotion (as measured by that emotion's distribution score). As we have several raters for each user, we have averaged the scores given by the different rates for that user. The Pearson correlations are given in Table I.

Table I shows that the propensity of expressing various emotions is correlated with many user interests. For example, users who are interested in sports tend to express less anger and more fear and surprise; users interested in arts express more joy and less disgust; those interested in religion express more joy and less surprise.

Noting the many correlations above, we checked whether it is possible to separate people who are interested in a topic from those disinterested in it by examining their propensity to display certain emotions. Given a specific interest i (such as "Sports" or "Movies and Television"), we compare two disjoint user populations: H_i, those interested in the area, consisting of the 20% of the user with highest perceived interest in i, and L_i, those disinterested in the area, consisting of the 20% of the users with lowest perceived interest in i. We note that $H_i \cup L_i$ only contain 40% of the entire user population.

Given a specific emotion, $e \in E$, we can check whether people interested in i express the emotion e more (or less) than those disinterested in e. More precisely, we test whether H_i and L_i have significantly different emotional scores with regard to e (either the proportion score or the distribution score), using a Mann-Whiteney U test. The p-values for all emotions and interests are given in Table II.

Table II shows that for all our studied interest areas, there are at least three emotions (and usually more) where users who are interested in an area display a significantly different propensity (≤ 0.05) to display an emotion than those disinterested in the area.

However, statistically significant differences between groups (those interested in an area and those disinterested in it) could still occur even if the differences in emotional scores between the groups is not very large. To check the effect size, we examine the differences in emotional scores between those groups. For every emotion e and every interest i, we examine H_i and L_i (those interested and disinterested in i), and compute the difference in the mean emotional scores for e between the groups. We denote this emotional difference value as $\delta_{i,e}$.

A high value of $\delta_{i,e}$ indicates that people interested in i tend to express the emotion e much more than those disinterested in it, while negative values of $\delta_{i,e}$ indicate that those disinterested in i express e more than those interested in i (with values close to zero indicating that both groups express e roughly the same). Table III presents the emotional difference values.

Table III shows that different interest areas are correlated

(a) Arts, Painting, Dance (b) Business, Economics, Finance (c) Fashion (d) Health, Medicine

(e) Politics, Society, News (f) Religion, Spiritualism (g) Science, Environment (h) Sports

(i) Movies, Television (j) Technology, Computing, Internet (k) Travel, Leisure

Fig. 3: Histogram of user interest scores, for each of the interest areas, across different users

	Anger	Disgust	Fear	Joy	Sadness	Surprise
Sports	**-0.1491**	-0.0183	**0.2259**	-0.1001	-0.1205	0.2039
Movies and Television	-0.104	-0.0922	**-0.192**	-0.0012	0.1151	0.1723
Technology, Computing and Internet	0.0176	**-0.1808**	0.0384	0.0683	-0.1444	0.0983
Politics, Society and News	0.0082	-0.0859	**0.3194**	0.1124	**-0.3411**	-0.0896
Business, Economics and Finance	-0.0275	-0.1715	0.2249	0.1575	**-0.3263**	-0.0286
Arts, Painting and Dance	-0.0879	-0.1583	-0.1619	**0.2003**	-0.0516	-0.0227
Science and Environment	-0.0272	**-0.1891**	0.1395	0.1808	**-0.2579**	-0.0584
Health and Medicine	-0.0719	-0.1234	0.0462	**0.1938**	-0.1819	-0.0948
Religion and Spiritualism	-0.1101	-0.1	0.0736	**0.2265**	-0.156	**-0.2091**
Travel and Leisure	-0.1258	-0.1742	-0.1375	**0.2038**	-0.0351	-0.0424
Fashion	0.005	0.0607	**-0.3076**	0.0269	0.2216	-0.0448

TABLE I: Pearson correlations between users' interest scores and propensity to express various emotions (emotion distribution score). The greatest values of every row and column are embolden.

with different emotional differences. For some interests, there are large differences between those interested in the topic and those disinterested in it in their propensity to express some emotions. For instance, those interested in politics express much more fear but much less sadness than those disinterested in it; those interested in fashion express less fear and more sadness than those disinterested in it; those interested in travel and leisure express more joy and less disgust than those disinterested in it; users interested in business express more joy and fear and less sadness and disgust than those disinterested in business.

Perhaps unsurprisingly, the results in Table III are quite similar to Table I. However, we note that Table I is based on the Pearson correlations between interests and emotions for the *entire user population*, whereas Table III is only based on the users who are *very interested* or *very disinterested* in various topics.

Our results so far indicates that it should be possible to predict whether a user is likely to be very interested or very disinterested in an area, based solely on their propensity to display various emotions. We now discuss such predictive models and their accuracy.

First, we tried to build a model to separate users very interested in a topic from those very disinterested in it. For each interest area $i \in I$, we train a classification model, which attempts to predict whether a target user u is a user who is very interested in the topic i, so $u \in H_i$, or whether u is very disinterested in i, so $u \in L_i$, based solely on the user's propensity to express each of Eackman's six basic emotions. This is a classification model M_i^c, which takes as input the six emotion distribution scores, $s_u^{joy}, s_u^{sadness}, \ldots, s_u^{disgust}$. It outputs a binary label: 0 if u is predicted to be in L_i and 1 if u is predicted to be in H_i. The training and test data for the model includes only users from $L_i \cup H_i$ (i.e. we removed users with moderate degrees of interest in i). We evaluated the performance of the models using 10-fold cross validation.

Figure 5 presents the prediction quality for the classification models M_i^c, as measured by the Area Under the ROC Curve. It shows that for all the studied interest areas we studied, it is possible to determine whether a person is very interested or very disinterested in the topic based only on their propensity to express various emotions. The best predictions were achieved for sports, fashion and politics (with a high AuC of roughly 80%), and the worst were achieved for religion, movies and health (with AuC slightly above 60%).

The M_i^c only allow deciding whether a user is very interested or very disinterested in a topic. However, to what degree can we predict the *degree* of interest of a user in a topic on a *fine grained scale*, based on the emotions they express?

For each interest area $i \in I$, we built a regression model, which attempts to predict a target user's u *degree* of interest in the topic i, based on their propensity to express each of Eackman's six basic emotions. This is a regression model M_i^r, again taking as input the six emotion distribution scores, and outputting a prediction for k_u^i - user u's degree of interest in i. Each such model was trained as a linear regression model, using the population of all out Twitter users, and evaluated using 10-fold cross validation.

Figure 5 presents the prediction quality for the regression models M_i^r, as measured in their R^2 values.

The best predictions were achieved for sports, politics and business (with a R^2 of roughly 0.35), and the worst were achieved for health, travel and arts (with R^2 of roughly 0.2). These results indicate that while it is possible to predict a user's degree of interest in all studied areas using their expressed emotions, though the quality of the predictions varies across interests.

VI. Conclusions

All users express all the basic emotions on social media, but the relative quantity of each emotion differs across people. We examined the relation between user interests and the emotions their express in social media. Our results show several interesting correlations between a user's interest areas and their propensity to express certain emotions. Further, our results show that using the proportions of emotions a user expresses in social media it is possible to determine whether they are interested or disinterested in various topics.

Our results have several direct applications in online advertising. First, our models can be used to predict user interests so as to personalize advertising content delivered to users. Clearly, there are other methods to predict user interests. For example, one may try to directly mine the words in a user's post to determine their interest. However, we note that emotions are expressed on a very wide variety of environments, making our methodology quite general. Further, even following very few posts by a user we can begin to track their emotional pattern, whereas it may a while for users to directly discuss their interests.

Another application of our approach is social and psychological research. Our models allow making prediction regarding user interests and their expressed emotions. These models can be used to analyze self-disclosure processes, in both offline and online environments. It stands to reason that users would be reluctant to disclose their emotional state to new people, but discussing general interests seem to require less intimacy. Our models can be used to track the issues discussed by two people, and thus to shed light about how people decide to disclose information about themselves.

We note that despite the promising results, our research does have several limitations. First, our interest scores are based on the evaluation of user profiles by strangers recruited via crowdsourcing. As such, these scores may inaccurately reflect the true interest of users. Similarly, our method for determining emotional scores is based on a machine learning model, which may not be completely accurate. Finally, this is an observational study, so we cannot make causal inference. In particular, it is unclear whether people who are interested in a certain topic tend to adopt a specific emotional pattern, or whether people who experience various emotions tend to be attracted to some areas of interest.

Additionally, despite the good performance of our predictive models, we note that for several interests there still remains a significant portion of unexplained variance in the degree of interest. In other words, there are some cases where users who are interested in different topics exhibit similar patterns in the emotions they express. Nonetheless, our results illustrate that emotions form a useful characterization of a user, which captures much information about their preferences and interests. It is possible to represent each person in a low-dimensional *emotional spectrum*, reflecting the relative degree to which they express each emotion, in a way that retains much of the important information regarding the interests of these users. This representation is easily understandable by people, it characterizes users as "angry", "easily surprised", "sad" or "joyful".

Our research leaves several questions open for future work.

First, our results focus on Ekman's six basic emotions. Other models offer a more fine-grained list of emotions, or alternative hierarchy of emotions. Would it be possible to improve the predictive performance of user interests using such alternative representations?

Second, our study uses Twitter data. Can similar results be obtained for other social networks? For example, would it be possible to use emotions expressed non-verbally in Flickr images (as perceived by viewers) to determine the interests of the profile owner.

Finally, what other dimensions of user information can be

	Anger	Disgust	Fear	Joy	Sadness	Surprise
Sports	≤ 0.001		≤ 0.001	≤ 0.01	≤ 0.001	≤ 0.001
Movies and Television	≤ 0.01		≤ 0.001			≤ 0.001
Technology, Computing and Internet		≤ 0.001		≤ 0.05	≤ 0.01	≤ 0.05
Politics, Society and News		≤ 0.001	≤ 0.001	≤ 0.01	≤ 0.001	
Business, Economics and Finance		≤ 0.001	≤ 0.001	≤ 0.01	≤ 0.001	
Arts, Painting and Dance	≤ 0.05	≤ 0.001	≤ 0.001	≤ 0.001		
Science and Environment		≤ 0.001	≤ 0.01	≤ 0.001	≤ 0.001	
Health and Medicine		≤ 0.01		≤ 0.001	≤ 0.05	≤ 0.01
Religion and Spiritualism	≤ 0.001	≤ 0.001		≤ 0.001	≤ 0.05	≤ 0.001
Travel and Leisure	≤ 0.001	≤ 0.001	≤ 0.001	≤ 0.001		
Fashion		≤ 0.05	≤ 0.001		≤ 0.001	≤ 0.05

TABLE II: Statistically significant differences between people interested in a topic and disinterested in the topic, in terms of their propensity to express emotions (captured by the emotional distribution scores). Values shown are p-values from a Mann-Whiteney U test (with missing value indicating no statistically significant effect: $p > 0.05$).

	Anger	Disgust	Fear	Joy	Sadness	Surprise
Sports	**-41.999**	-7.756	58.824	-31.569	-26.514	**61.646**
Movies and Television	-34.414	-23.183	-53.72	-7.757	36.672	**58.755**
Technology, Computing and Internet	-0.255	**-47.978**	1.559	18.273	-29.358	26.637
Politics, Society and News	-1.535	-33.465	79.082	36.429	**-90.748**	-22.129
Business, Economics and Finance	-7.532	-40.692	55.453	38.68	**-81.124**	-6.516
Arts, Painting and Dance	-20.585	-35.798	-50.59	**54.899**	-13.104	-7.754
Science and Environment	-13.49	-42.784	33.008	45.043	**-62.25**	-14.054
Health and Medicine	-16.945	-28.569	11.122	**47.254**	-38.786	-30.752
Religion and Spiritualism	-32.103	-27.354	8.001	**58.285**	-33.086	-48.125
Travel and Leisure	-27.008	-39.866	-33.703	**49.421**	-8.808	-12.969
Fashion	-1.233	14.58	**-88.412**	10.962	60.505	-12.21

TABLE III: Emotional difference values $\delta_{i,e}$ (multiplied by 100 for easy comparison). The greatest values of every row and column are embolden.

used to predict their interests? For example, can information about a user's demographic traits or personality be used to predict their interests?

REFERENCES

[1] E. Qualman, *Socialnomics: How social media transforms the way we live and do business*. John Wiley & Sons, 2010.

[2] T. L. Tuten, *Advertising 2.0: social media marketing in a web 2.0 world*. Greenwood Publishing Group, 2008.

[3] W. G. Mangold and D. J. Faulds, "Social media: The new hybrid element of the promotion mix," *Business horizons*, vol. 52, no. 4, pp. 357–365, 2009.

[4] D. L. Hoffman and M. Fodor, "Can you measure the roi of your social media marketing," *MIT Sloan Management Review*, vol. 52, no. 1, pp. 41–49, 2010.

[5] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting personality from Twitter," in *Proceedings of SocialCom/PASSAT*, 2011.

[6] Y. Bachrach, M. Kosinski, T. Graepel, P. Kohli, and D. Stillwell, "Personality and patterns of Facebook usage," in *Proceedings of WebSci*, 2012, pp. 24–32.

[7] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," *National Academy of Sciences*, 2013.

[8] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[9] S. Mohammad and S. Kiritchenko, "Using nuances of emotion to identify personality," in *Proceedings of ICWSM*, 2013.

[10] S. Volkova, T. Wilson, and D. Yarowsky, "Exploring demographic language variations to improve sentiment analysis in social media," in *Proceedings of EMNLP*, 2013.

[11] Y. Yang, J. Jia, S. Zhang, B. Wu, J. Li, and J. Tang, "How do your friends on social media disclose your emotions?" in *Proc. AAAI*, vol. 14, 2014, pp. 1–7.

[12] A. M. Kaplan and M. Haenlein, "Users of the world, unite! the challenges and opportunities of social media," *Business horizons*, vol. 53, no. 1, pp. 59–68, 2010.

[13] M. Zappavigna, *Discourse of Twitter and social media: How we use language to create affiliation on the web*. A&C Black, 2012.

[14] H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, L. Dziurzynski, S. M. Ramones, M. Agrawal, A. Shah, M. Kosinski, D. Stillwell, M. E. Seligman *et al.*, "Personality, gender, and age in the language of social media: The open-vocabulary approach," *PloS one*, vol. 8, no. 9, p. e73791, 2013.

[15] D. Bamman, J. Eisenstein, and T. Schnoebelen, "Gender identity and lexical variation in social media," *Journal of Sociolinguistics*, vol. 18, no. 2, pp. 135–160, 2014.

[16] E. K. Clemons, S. Barnett, and A. Appadurai, "The future of advertising and the value of social network websites: some preliminary examinations," in *Proceedings of the ninth international conference on Electronic commerce*. ACM, 2007, pp. 267–276.

[17] M. Trusov, R. E. Bucklin, and K. Pauwels, "Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site," *Journal of marketing*, vol. 73, no. 5, pp. 90–102, 2009.

[18] D. G. Taylor, J. E. Lewin, and D. Strutton, "Friends, fans, and followers: do ads work on social networks?" *Business Faculty Publications*, 2011.

[19] W.-S. Yang, J.-B. Dia, H.-C. Cheng, and H.-T. Lin, "Mining social networks for targeted advertising," in *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*, vol. 6. IEEE, 2006, pp. 137a–137a.

[20] C. E. Tucker, "Social networks, personalized advertising, and privacy controls," *Journal of Marketing Research*, vol. 51, no. 5, pp. 546–562, 2014.

[21] P. Resnick and H. R. Varian, "Recommender systems," *Communications of the ACM*, vol. 40, no. 3, pp. 56–58, 1997.

[22] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*. ACM, 2001, pp. 285–295.

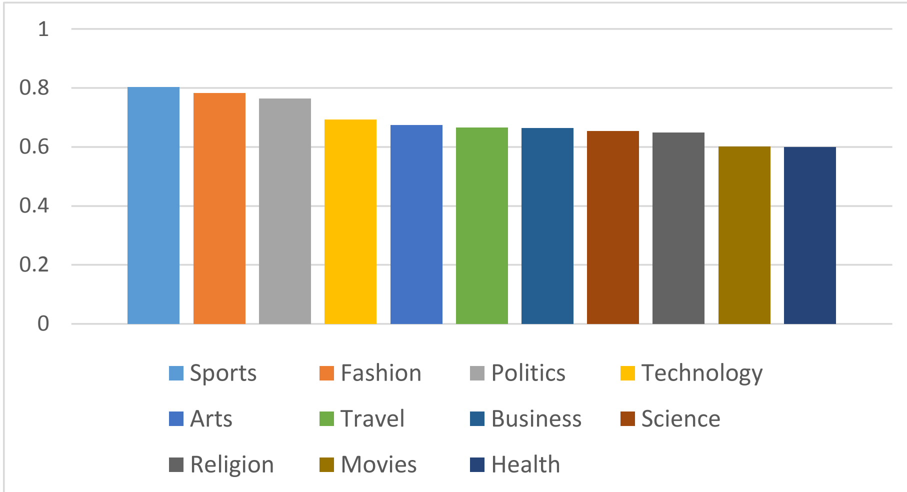


Fig. 4: Prediction quality, in Area Under the ROC Curve (AuC), for the interest degree classification models M_i^c

[23] ——, "Application of dimensionality reduction in recommender system-a case study," DTIC Document, Tech. Rep., 2000.

[24] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, no. 8, pp. 30–37, 2009.

[25] Y. Bachrach, R. Herbrich, and E. Porat, "Sketching algorithms for approximating rank correlations in collaborative filtering systems," in *SPIRE*, 2009.

[26] Y. Bachrach, E. Porat, and J. S. Rosenschein, "Sketching techniques for collaborative filtering," in *IJCAI*, Pasadena, California, July 2009.

[27] Y. Bachrach and R. Herbrich, "Fingerprinting ratings for collaborative filteringtheoretical and empirical analysis," in *SPIRE*. Springer, 2010, pp. 25–36.

[28] P. Li and C. König, "b-bit minwise hashing," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 671–680.

[29] A. Dasgupta, R. Kumar, and T. Sarlós, "Fast locality-sensitive hashing," in *KDD*. ACM, 2011, pp. 1073–1081.

[30] Y. Bachrach and E. Porat, "Sketching for big data recommender systems using fast pseudo-random fingerprints," in *ICALP*, 2013, pp. 459–471.

[31] Y. Bachrach, Y. Finkelstein, R. Gilad-Bachrach, L. Katzir, N. Koenigstein, N. Nice, and U. Paquet, "Speeding up the xbox recommender system using a euclidean transformation for inner-product spaces," in *RecSys*, 2014.

[32] Y. Bachrach and E. Porat, "Fingerprints for highly similar streams," *Information and Computation*, 2015.

[33] D. B. Montgomery and A. J. Silk, "Clusters of consumer interests and opinion leaders' spheres of influence," *Journal of Marketing Research*, pp. 317–321, 1971.

[34] S.-S. Weng and M.-J. Liu, "Feature-based recommendations for one-to-one marketing," *Expert Systems with Applications*, vol. 26, no. 4, pp. 493–508, 2004.

[35] Y. Bachrach, S. Ceppi, I. A. Kash, P. Key, F. Radlinski, E. Porat, M. Armstrong, and V. Sharma, "Building a personalized tourist attraction recommender system using crowdsourcing," in *AAMAS*, 2014, pp. 1631–1632.

[36] P.-T. Chen, J. Z. Cheng, Y.-W. Yu, and P.-H. Ju, "Mobile advertising setting analysis and its strategic implications," *Technology in Society*, vol. 39, pp. 129–141, 2014.

[37] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: Liwc and computerized text analysis methods," *Journal of language and social psychology*, vol. 29, no. 1, pp. 24–54, 2010.

[38] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and trends in information retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.

[39] C. Strapparava and R. Mihalcea, "Learning to identify emotions in text," in *Proceedings of the 2008 ACM symposium on Applied computing*. ACM, 2008, pp. 1556–1560.

[40] J. T. Hancock, C. Landrigan, and C. Silver, "Expressing emotion in text-based communication," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007, pp. 929–932.

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

[42] S. Stieglitz and L. Dang-Xuan, "Emotions and information diffusion in social mediasentiment of microblogs and sharing behavior," *Journal of Management Information Systems*, vol. 29, no. 4, pp. 217–248, 2013.

[43] S. M. Mohammad and S. Kiritchenko, "Using hashtags to capture fine emotion categories from tweets," *Computational Intelligence*, 2014.

[44] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, vol. 6, no. 3-4, pp. 169–200, May 1992. [Online]. Available: http://dx.doi.org/10.1080/02699939208411068

[45] W. Wang, L. Chen, K. Thirunarayan, and A. P. Sheth, "Harnessing Twitter" big data" for automatic emotion identification," in *Proceedings of SocialCom*, 2012, pp. 587–592.

[46] K. Roberts, M. A. Roach, J. Johnson, J. Guthrie, and S. M. Harabagiu, "Empatweet: Annotating and detecting emotions on Twitter." in *Proceedings of LREC*, 2012.

[47] S. Kim, J. Bak, and A. H. Oh, "Do you feel what I feel? Social aspects of emotions in Twitter conversations." in *Proceedings of ICWSM*, 2012.

[48] D. Rao, D. Yarowsky, A. Shreevats, and M. Gupta, "Classifying latent user attributes in Twitter," in *Proceedings of SMUC*, 2010, pp. 37–44.

[49] J. D. Burger, J. Henderson, G. Kim, and G. Zarrella, "Discriminating gender on Twitter," in *Proceedings of EMNLP*, 2011, pp. 1301–1309.

[50] B. Van Durme, "Streaming analysis of discourse participants," in *Proceedings of EMNLP*, 2012, pp. 48–58.

[51] M. Ciot, M. Sonderegger, and D. Ruths, "Gender inference of Twitter users in non-english contexts." in *Proceedings of EMNLP*, 2013, pp. 1136–1145.

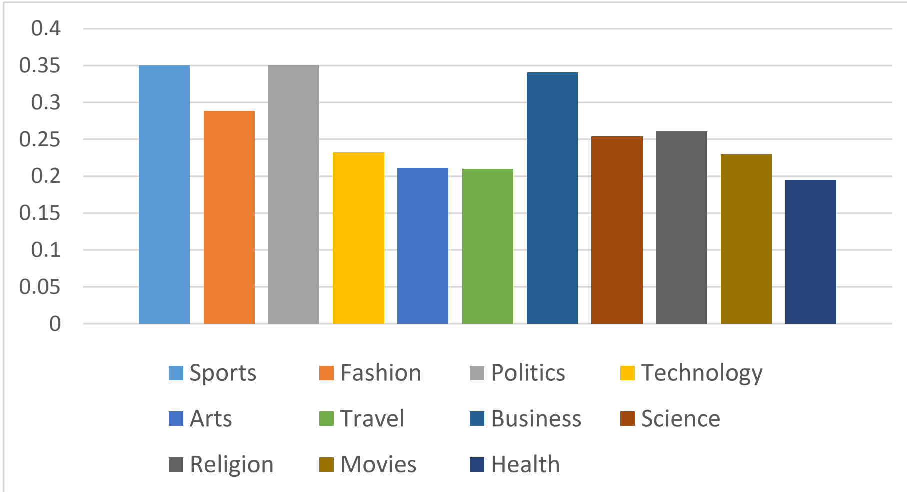


Fig. 5: Prediction quality, in R values, for the interest degree regression models M_i^r

[52] S. Mohammad and T. Yang, "Tracking sentiment in mail: How genders differ on emotional axes," in *Proceedings of WASSA*, 2011, pp. 70–79.

[53] D. Nguyen, R. Gravel, D. Trieschnigg, and T. Meder, ""How old do you think I am?" A study of language and age in Twitter," in *Proceedings of ICWSM*, 2013, pp. 439–448.

[54] Y. Bachrach, T. Graepel, P. Kohli, M. Kosinski, and D. Stillwell, "Your digital image: factors behind demographic and psychometric predictions from social network profiles," in *AAMAS*, 2014, pp. 1649–1650.

[55] M. Pennacchiotti and A. M. Popescu, "A machine learning approach to Twitter user classification," in *Proceedings of ICWSM*, 2011, pp. 281–288.

[56] R. Cohen and D. Ruths, "Classifying Political Orientation on Twitter: It's Not Easy!" in *Proceedings of ICWSM*, 2013.

[57] T. A. Mooradian and J. M. Olver, "Shopping motives and the five factor model: An integration and preliminary study," *Psychological Reports*, vol. 78, no. 2, pp. 579–592, 1996.

[58] J.-H. Huang and Y.-C. Yang, "The relationship between personality traits and online shopping motivations," *Social Behavior and Personality: an international journal*, vol. 38, no. 5, pp. 673–679, 2010.

[59] M. Kosinski, D. Stillwell, P. Kohli, Y. Bachrach, and T. Graepel, "Personality and website choice," 2012.

[60] M. Kosinski, Y. Bachrach, P. Kohli, D. Stillwell, and T. Graepel, "Manifestations of user personality in website choice and behaviour on online social networks," *Machine Learning*, vol. 95, no. 3, pp. 357–380, 2014.

[61] M. D. Conover, J. Ratkiewicz, M. Francisco, B. Gonc, A. Flammini, and F. Menczer, "Political polarization on Twitter," in *Proceedings of ICWSM*, 2011, pp. 89–96.

[62] F. A. Zamal, W. Liu, and D. Ruths, "Homophily and latent attribute inference: Inferring latent attributes of Twitter users from neighbors," in *Proceedings of ICWSM*, 2012, pp. 387–390.

[63] S. Volkova, G. Coppersmith, and B. Van Durme, "Inferring user political preferences from streaming communications," in *Proceedings of ACL*, 2014, pp. 186–196.

[64] S. Volkova, B. Van Durme, D. Yarowsky, and Y. Bachrach, "Social media predictive analytics."

[65] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[66] N. A. Smith, "Log-linear models," 2004.

[67] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in *Proceedings of EMNLP*, 2002, pp. 79–86.

[68] H. Zheng, D. Li, and W. Hou, "Task design, motivation, and participation in crowdsourcing contests," *International Journal of Electronic Commerce*, vol. 15, no. 4, pp. 57–88, 2011.

[69] X. A. Gao, Y. Bachrach, P. Key, and T. Graepel, "Quality expectation-variance tradeoffs in crowdsourcing contests." in *AAAI*, 2012.

[70] J. Witkowski, Y. Bachrach, P. Key, and D. C. Parkes, "Dwelling on the negative: Incentivizing effort in peer prediction," in *HCOMP*, 2013.

[71] P. Welinder, S. Branson, P. Perona, and S. J. Belongie, "The multidimensional wisdom of crowds," in *NIPS*, 2010, pp. 2424–2432.

[72] Y. Bachrach, T. Graepel, T. Minka, and J. Guiver, "How to grade a test without knowing the answers—a bayesian graphical model for adaptive crowdsourcing and aptitude testing," *ICML*, 2012.

[73] M. Salek, Y. Bachrach, and P. Key, "Hotspotting-a probabilistic graphical model for image object localization through crowdsourcing." in *AAAI*, 2013.

[74] B. Shalem, Y. Bachrach, J. Guiver, and C. M. Bishop, "Students, teachers, exams and moocs: Predicting and optimizing attainment in web-based education using a probabilistic graphical model," in *ECML/PKDD*. Springer, 2014, pp. 82–97.

[75] Y. Bachrach, T. Graepel, G. Kasneci, M. Kosinski, and J. Van Gael, "Crowd iq: aggregating opinions to boost performance," in *AAMAS*, 2012, pp. 535–542.

[76] M. Kosinski, Y. Bachrach, G. Kasneci, J. Van-Gael, and T. Graepel, "Crowd iq: Measuring the intelligence of crowdsourcing platforms," in *WebSci*. ACM, 2012, pp. 151–160.