Grounded Emotions

Vicki Liu, Carmen Banea and Rada Mihalcea
University of Michigan
Ann Arbor, MI
{*vickiliu,carmennb,mihalcea*}*@umich.edu*

***Abstract*—Emotions are grounded in contextual experience. While natural language processing tools typically look at textual content to find clues pertaining to an author's emotional state, factors occurring throughout the day, such as weather or news exposure, may prime one toward a particular emotional response. In this paper, we explore five types of external factors and through extensive analyses show their impact and correlation with a user's emotional state. Ultimately, we show that when combining all extrinsic features, we are able to predict with an accuracy of 67% the emotional state of a user.**

1. Introduction

Extensive research has gone into predicting emotions of tweets based on the Twitter text itself using traditional natural language processing techniques. While these classifiers have found success, they do not account for the environmental factors in which these tweets were composed, and thus do not account for the effect that real world phenomena have on emotions. As such, this paper explores the concept of *grounded emotions*, focusing on how external factors, ranging from weather, news exposure, social network emotion charge, timing and mood predisposition may have a bearing on one's emotion level throughout the day. By testing the correlation between certain external factors and Twitter sentiment, we explore which of them are most significant in grounding emotions, and therefore gain a deeper understanding of the connections that exist between contextual factors and one's internal emotional state.

Psychologists have posited that emotions are personal reactions which give validity to interactions with real world phenomena [1]. There have been several studies that have found interesting patterns of human emotion in regards to daily [2], seasonal [3], and weather related factors [4]. In 2014, Facebook researchers conducted a study [5] that sought to quantify the opposite, namely how altering external factors could cause a change in emotional response. As such, they purposefully altered the amount of positive or negative posts incoming from the social networks pertaining to approximately 700,000 of its users and gauged how this would modify the emotional response of the user, a concept they called "emotional contagion". They were able to see that users exposed to more negative posts from their friends contributed fewer positive posts and more negative posts, while those exposed to more positive posts increased their authorship of positive posts, and decreased the contribution of negative posts. A recent publication in The Guardian [6] links the Brexit phenomenon to the ability of private companies to influence the emotional response of users and to sway them to be for or against a given issue. It is clear that the more we publicly study what primes people toward a particular emotional response, the more we will be able to avoid and inoculate against such unscrupulous tactics.

To the best of our knowledge, ours is the first study that looks at the ability to predict a user's emotion without relying on the content of the posted tweet. We exploit self identified user emotional state through tags like #happy or #sad, thus avoiding crowd-sourced studies where a third party would be asked to infer what was the emotional state of the author, and also avoiding automatic emotion classifiers with their inherent pitfalls. As we show through experiments, relying on state-of-the-art sentiment classifier output based on a given tweet content, we only achieve an accuracy of at most 64.8%, and a Pearson correlation of 0.24 with user specified emotion, while using signals external to the tweet allow us to achieve an accuracy of 66.9%, and a correlation of 0.3. While other works [7], [8] have also linked external factors to sentiment expressed in the Twitter network, their research was conducted at an aggregate location level, not user level, using inferred sentiment, not expressed sentiment, and using a narrower set of external influencers.

2. Related Work

Sentiment classification of tweets has been a well researched area in natural language processing [9], [10], [11], [12], [13], [14], and it has been leveraged to draw correlations between the sentiment in a social network and the stock market [15], [16], movie success [17], consumer insights [18], and others.

Related to our work, Hannak et al. [8] sought to determine if there is any influence on user mood caused by weather or time. They focused on a large number of tweets collected in 2009, and geolocated using the location field in the tweet. The unigrams extracted from the tweets were used to create a word list, where each entry was scored based on its co-occurrence with positive or negative emoticons,

978-1-5386-0563-9/17/$31.00

and were then employed in a machine learning framework. However, unlike our work, where we are looking at the impact of external factors on one particular tweet and its author at a time, they considered the aggregate response over all tweets occurring in hour-long buckets in different areas with respect to season, location, time and weather, and obtained the best performance by considering the entire set of dimensions.

Park et al. [7] focused their exploration on the effect of weather parameters on sentiment carried in Twitter data. Similarly to [8], they also considered aggregate correlation between positive and negative sentiment in the Twitter users from a given state and aspects such as temperature, humidity, precipitation, atmospheric pressure and wind. However, all the weather parameters were considered globally at a state scale, which can be problematic, as large states such as Texas or California can have markedly different weather at less than 100 miles away; our focus on metropolitan areas allows us to circumvent this problem. In addition, they only looked at various weather aspects, while we explore several other facets that we believe to be key to emotion prediction such as social network sentiment, or current news events. Furthermore, the aggregate positive or negative sentiment of a given state was derived using the Linguistic Inquiry and Word Count (LIWC) tool [19] applied over the content of the published tweets in that state, thus not relying on a human generated gold standard, as is possible in our case with the self-identified emotion by the tweet author. Unlike [7], [8], the analyses that we perform are not with respect to an emotion that is inferred, but rather to an emotion that the author of the tweet experiences at that precise moment. Thus, while the set of tweets we collected is more limited in size, it is of a higher quality, and allows direct correlations to be observed.

3. Data

Twitter is a social networking and content sharing platform, where users primarily interact by posting, responding to, liking, and re-posting tweets. Tweets are small statuses or microblogs limited to 140 characters that can include emojis, links to external sites, and of course text. Tweets also often include hashtags that can be embedded in the tweet context, allowing the tweet to be searchable by the hashtagged keyword, as well as provide further contextual meaning and emphasis to the tweet content itself. By default, tweets are publicly visible to everyone and Twitter users may follow or subscribe to other Twitter users in order to see followed accounts' tweets on their dashboard. The short nature of Tweets encourages users to be more "in the moment" and post updates as they happen, making Twitter a very fast paced, live streaming social platform.

For our study, we collected a set of microblogs published between January 18, 2017 and April 14, 2017 via the Twitter platform, which were self-tagged by their author with a #happy or #sad hashtag.[1] For each tweet, we consider the hashtag to represent the label, capturing the instantaneous emotional state of its author, and we collected the tweet's remaining content, as well as metadata, such as the time it was published, its author, and its location. The set was filtered based on location, such that the collected tweets originated from 20 large US metropolitan areas, making sure that no more than three cities were located in the same state, allowing us to obtain a large representative sample. Table 1 shows several examples of emotion tagged tweets.

The final set contains 2,557 labeled instances, consisting of 1,525 happy tweets and 1,032 sad tweets written by 1,369 unique users, with 30.9% of the users from large cities in California, 24.4% from large cities in New York, 11.5% from large cities in Texas, and the rest of them living in cities such as Miami, Atlanta, Portland, Kansas City, etc. Figure 1 shows the distribution of happy and sad tweets for the 10 metropolitan areas with the highest tweet volume of the 20 areas we monitored.

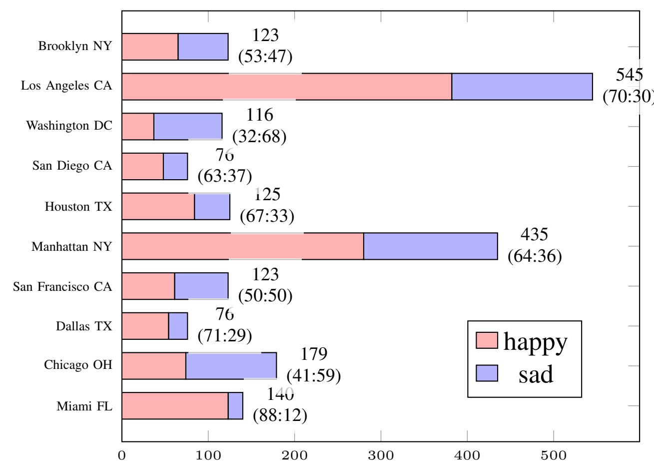

Figure 1. Happy / sad tweet distribution for 10 of the metropolitan areas covered by the data. X-axis: number of tweets; Y-axis: city.[3]

This data served as a core to be augmented with several external factors enumerated below.

Atmospheric conditions (weather). Similar to [7], we obtained local weather data from the Weather Underground API.[4] The site gives access to historical weather data by city with an hourly granularity, including information such as temperature, humidity, precipitation, etc., as well as a descriptive phrase summarizing atmospheric conditions, such as "heavy thunderstorms and rain." From this data, we extracted binary features for atmospheric conditions such as fog, rain, tornado, hail, snow and thunder, encoded as 0 if they did not occur or 1 if they did occur in the prior 12 hours to an emotion tagged tweet. Additionally, we used the 96 descriptive phrases provided by Weather Underground, but due to their very fine distinctions, such

1. https://dev.twitter.com/rest/reference/get/search/tweets

3. For a given data point, the first number represents the overall tweet count, while the numbers in parentheses delineate the percentage split between happy and sad-tagged tweets.

4. https://www.wunderground.com/weather/api/

TABLE 1. SAMPLE TWEETS TAGGED WITH #HAPPY OR #SAD HASHTAG.

Happy tweets	**Sad tweets**
Wow, this ice cream is delicious! #happy #nycblond #newyork #nyc #ny #ilovenewyork #delicious	You know what the hardest part about being ME is? The people who should know me best don't & the ones I loathe think they do know me. #Sad
Even far away you put a smile on my face ! I think friendship can turn into love #mymarine #happy #helovesme	I didn't eat any cake today #sad
Weekend aka where passion projects begin. #FridayFeeling #inspired & #happy	When did my world of the people become so sarcastic and twisted? #sad

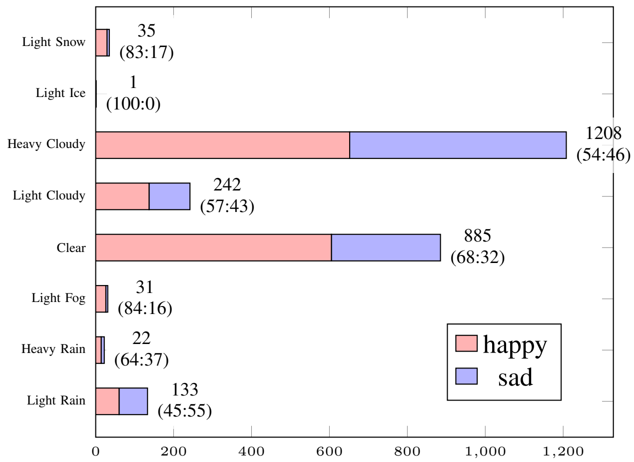

Figure 2. Happy / sad tweet distribution for various weather attributes. X-axis: number of tweets; Y-axis: weather attribute.[3]

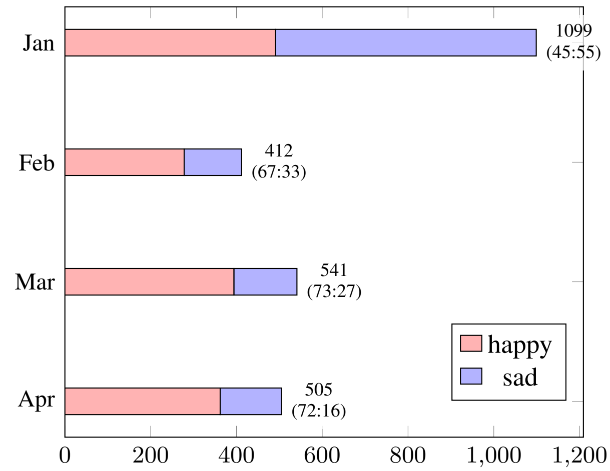

Figure 3. Happy / sad tweet distribution by month. X-axis: number of tweets; Y-axis: months.[3]

as for cloudy weather being described as "overcast," "partly cloudy," "mostly cloudy" or "scattered clouds," we defined more coarse groupings. For example, "overcast" and "mostly cloudy" were grouped into "heavy clouds," while the remaining were mapped to "light clouds." This allowed us to reduce the categories to 18 phrases, while still retaining the intensity of a given meteorological aspect. From the prior 12 hours with respect to an emotion tagged tweet, we retained the description that was most frequent. The API also gave us access to hourly temperature and humidity information. From this we extracted the most extreme (highest or lowest) temperature and humidity that occurred within the 12 previous hours of an emotion tagged tweet. We decided to use the most extreme temperature and humidity on the basis that people's emotions are more affected by abnormal weather conditions rather than common ones. Overall, this resulted in 8 weather features which we used to train and test our data.

News exposure (news). We obtained national news information using the New York Times API[5] by querying for the stories that appeared on the front page of the New York Times within the prior 24 hours of a user's tweet. The New York Times gives access to historical data going back to 1851, and the metadata for a given story includes not only the date when the story was published and its authors, but also whether it was front page, or part of the sports, entertainment, or other sections of the print version of the newspaper. For each front page story we extracted two items: the headline and the short snippet introducing the story. Each of these were separately analyzed using the Stanford CoreNLP sentiment annotator [20] in order to obtain a sentiment prediction of positive, negative, and neutral. Since neutral stories may not influence the emotional state of a reader, we focused only on the positive and negative headlines and snippets, and generated four scores that seek to represent the emotional exposure of a reader to the main positive and negative stories of the day: positive headline exposure, negative headline exposure, positive snippet exposure and negative snippet exposure. Each of these represents the number of items tagged with a particular sentiment normalized by the total number of such items on the front page.

Tweet participation timing (timing). Based on the tweet timestamp, we were able to determine the season, month, day of the week, and hour for when a tweet was authored; we also mapped the hours to several time intervals ranging from early morning, morning, afternoon, late afternoon, evening, late evening and night. Figures 3 and 4 show the distribution of happy and sad tweets by percent for each month and day of the week during the data collection period. We notice that while in January the happy to sad

5. https://developer.nytimes.com/

TABLE 2. SAMPLE TWEET AUGMENTED WITH EXTERNAL FACTORS

Tweet	Weather	User History	SN History	News
Great Idea but big waste of time! #Trumpublicans will NEVER give up Power for sake of America's well being. #SAD **Timing** Thu Mar 09 17:58:36 +0000 2017	Mostly Cloudy 72°F 25% Humidity No Rain No Fog No Thunder No Tornado No Hail No Snow	How are we supposed to show We Are #ProtestingTrump if our elected officials won't take our calls? #ClearYourVM America has a lot to say!	RT[6]@HillaryClinton: On #International-WomensDay I'm thinking about this young girl, & all the others like her out there.	A Brooklyn Charter School Looks Past 'No Excuses'. With Rebels Gone, Colombia Jumps Into the Pot Industry.

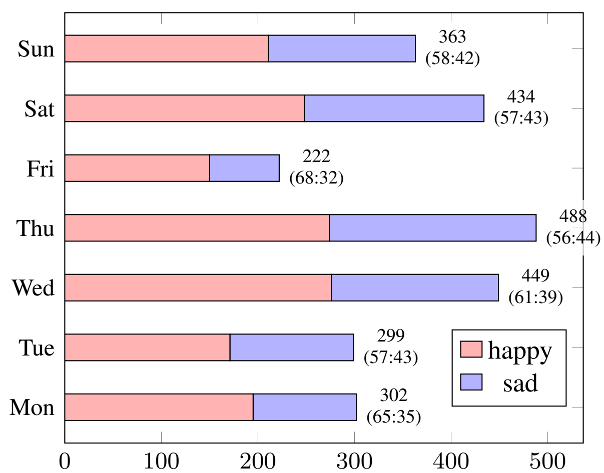

Figure 4. Happy / sad tweet distribution by day. X-axis: number of tweets; Y-axis: day of the week.[3]

distribution was more balanced, a strong preference toward a happy emotional state starts to emerge in February, showing an increase from 45% to 67% of total tweets being happy. The trend continues in March and April, but the distribution stabilizes at approximately 73% happy tweets. In terms of day of the week emotional tweet distribution, we notice that Friday is the most positive Twitter day, with 68% of the emotional tagged tweets being happy. Interestingly enough, the second most positive day is Monday, with 65% happy tweets. One explanation for this may be that people tweet on Monday about things they have done in the weekend, however additional research into this trend is needed.

Social network sentiment charge (social network). For each tweet, we queried all followees, or people that the given user follows, gathering tweets posted in the 24 hours prior to the original tweet's timestamp. Each followee tweet is tagged as positive or negative using the Stanford CoreNLP sentiment annotator [20]. In order to gauge the amount of exposure to positivity or negativity in a Twitter's user social network, we compute the proportion of positive and negative incoming tweets. As such, we compute the positive score, which is a summation over the incoming positive tweets, normalized by the total number of incoming tweets, and similarly we compute the negative score.

User sentiment predisposition (predisposition). We posit that users have a relatively consistent emotional state, where fast changes between happiness and sadness are unlikely. For this reason, we also collected the tweets that the user posted within 12 hours prior to the target tweet. Similar to the calculation done for social network sentiment charge, these were tagged as positive or negative by the Stanford CoreNLP sentiment annotator and aggregated into a positive and a negative sentiment score representing the user's predisposition toward happiness and sadness.

To better exemplify, Table 2 summarizes the type of information we extracted for a given tweet, which was then used to derive the features mentioned above.

4. Analyses of External Factors

We conduct 10-fold cross validation evaluations over the emotion tagged dataset we collected, ensuring that all the tweets pertaining to a given user are either present in the train set or the test set, but never in both. This insures that the particulars of a given user will not be leveraged in order to produce a grounded emotion prediction. We evaluate each feature type using the random forests [21] supervised learning algorithm distributed with the Weka machine learning toolkit [22]. We also employed several other algorithms tested on the first fold, such as support vector machines [23], Naive Bayes and decision trees (C4.5 [24]), however, the performance of random forests was significantly better, and we are employing it throughout all our experiments. Due to the skew in the happy / sad class distribution, the majority class baseline is 59.56%. Table 3 shows the results obtained by predicting grounded emotions for each feature type.

Analyzing each feature type individually, we notice that the sentiment extracted from a user's tweet history is the best predictor for an upcoming happy emotion, while the social network sentiment charge follows closely. However, sadness does not propagate the same way, and social network negativity does not translate in an emotional negative

6. RT indicates a "retweet" or reposting of a tweet authored by somebody else (whose username is specified by the @ handle).

TABLE 3. INDIVIDUAL FEATURE RESULTS.

Features	F-happy	F-sad	F	Acc	r
Features independent of tweet content					
Weather	0.68	0.43	0.55	59.9%	0.119
Predisposition	**0.75**	0.38	0.57	**64.6%**	**0.225**
Social network	0.74	0.10	0.42	59.2%	0.022
News	0.69	**0.50**	**0.60**	61.4%	0.185
Timing	0.71	0.45	0.58	61.8%	0.176
Features based on tweet content					
Unigrams	0.87	0.81	0.84	84.3%	0.677
Sentiment RB	0.21	0.53	0.37	47.2%	0.047
Sentiment ML	0.73	0.50	0.65	64.8%	0.243

impact for the user, having the lowest impact among all feature types we explored at 10% F-measure. For a given user, the sentiment extracted from historical tweets (in the prior 12 hours to a tagged tweet) shows a 0.24 Pearson correlation with the self-specified emotion tag, while the sentiment expressed in a user's network exhibits an only 0.02 correlation, indicating that it is a very poor predictor of user emotion. We should note that while Facebook's study [5] has shown that exposure to positive / negative content for its users triggers them to contribute more posts with a matching sentiment type, and produce less posts with an opposing sentiment type, Facebook-based social networks are more compact and contain more people that the user personally knows compared to the network of people Twitter users follow, which has fewer social friendships. It may be the case that the social distance between a user and the people he / she follows plays a role in how the latter's group emotion influences the user.

Sadness seems to be best predicted by news exposure, and actually news derived sentiment is the most balanced feature type we tested, performing at 69% F-measure for happiness, and at 50% F-measure for sadness, for an overall accuracy of 61%. In addition, news-derived sentiment emotion prediction exhibits a high correlation, being the top performing feature that is not user dependent, second only to user emotion predisposition at 0.22. We should note that even if not all Twitter platform users read the New York Times, the fact that we consider the sentiment derived from front page stories, which are most likely to appear in other national publications or online news venues as well, does provide us with a snapshot of the news stories that have primed the user for a given day. News stories have the highest overall F-measure among all single features.

From the timing-derived features we extracted, the month and day of the week are most robust, while the hourly information or mapping the hours to more generic time slots in a day do not carry additional information. Timing based emotion prediction displays a correlation of 0.18 with the gold standard. We originally thought that the hour of the day in which the tweet was posted would have a bearing on the tweet's emotion; however, our experiments have shown that including the hour in a 24 hour format as a feature provided no additional information, leading to a slightly lower performance that was also not statistically significant.

We also tried to represent the tweets in terms of the season when they occurred, grouping January through February tweets into the winter group, and the March through April tweets into the spring group, however this did not have a positive outcome on accuracy. It is however possible that since we only collected data over winter and spring, that more data and further analyses would be needed to better validate this observation.

Unlike findings in prior research linking weather derived features to mood [7], [8], the weather derived features are not particularly good influencers of personal emotional state, displaying an accuracy less than half a point above the baseline. We should note, however, that we are not looking at aggregate emotion influenced by weather in a large slice of the population, but rather at an individual level, and at this level, weather does not seem to play an important role. Actually, at an individual level, the F-measure for happiness is the lowest among all the feature types we tested, while for sadness the F-measure is 43%, ranking third out of 5 feature types explored. In addition, weather-based emotion predictions achieve the second lowest correlation with the gold standard.

We are aware that classification using tweet content derived features can achieve better performance in predicting tweet emotion. For example, using unigram features extracted from tweet content stripped of hashtags, we are able to achieve a prediction accuracy of 84.27%. Nonetheless, what we seek to model in this paper are the aspects that can influence a given tweet's emotion that are external to the tweet itself. We should point out that using the Stanford CoreNLP sentiment classifier on the text of a tweet to label each tweet with a score from 0 to 4, where 0 is highly negative, 1 is negative, 2 is neutral, 3 is positive, and 4 is highly positive, in a rule-based approach (referred to in Table 3 as Sentiment RB) mapping all negative predictions to sad and all positive predictions to happy (neutral predictions are considered wrong) only achieves an F-measure for happy of 20.5%, an F-measure for sad of 53.3%, and an overall accuracy of 47.2%. While the sadness F-measure is higher than any of those obtained for feature types independent of tweet content, the happiness F-measure lags considerably, indicating that the sentiment classifier is better tuned to identifying negative sentiment, and that even directly mapping sentiment predictions does not amount to surpassing an unsupervised 50/50 binary class baseline. Since using a rule-based classifier in our setup is not conducive to mapping the neutral score to a given emotion class, we also conducted an additional experiment, where the scores predicted by the Stanford sentiment annotator were used to train a random forest classifier in a similar way to our earlier experiments (listed in Table 3 as Sentiment ML). This classifier achieves a 64.8% accuracy, which falls in the vicinity of performances that we can achieve from external factors not reliant on the tweet content.

5. Grounded Emotion Prediction

TABLE 4. COMBINED FEATURES PREDICTION RESULTS.

Features	F-happy	F-sad	F	Acc	r
Pred.[7]+News	**0.75**	0.51	0.63	**66.8%**	0.286
Pred.+Timing	0.73	0.48	0.61	65.1%	0.203
Pred.+News+Timing	0.73	0.53	0.63	65.7%	0.267
All Features-Pred.	0.71	0.51	0.61	63.8%	0.230
All Features	0.74	**0.56**	**0.65**	**66.9%**	**0.299**

While in the previous section we explored the impact that each factor alone has on prediction accuracy, in this section we look at the joint modeling capabilities of the feature sets. As such, we combine the features exhibiting the highest correlation with the gold standard labels, testing user predisposition & news-based features, and user predisposition & timing-based features. The first combination, involving the sentiment extracted from a given user's tweet history prior to the target tweet and the sentiment inferred from the New York Times front page, displays the strongest performance on both the happiness and sadness classes, with an overall accuracy of 66.84%, and a correlation of 0.29. The second variation achieves lower results. While the accuracy obtained by news and timing features by themselves was relatively higher for timing features, the fact that news features were able to improve by 5% for sadness F-measure, ultimately allowed the combination involving it to aid in deriving a more robust feature-mix.

Learning from the best three features improves the sadness F-measure by 2 points (from 51% to 53%), but lowers the happiness F-measure by the same amount, resulting in an overall lower accuracy. Considering all the five proposed feature types however, results in the strongest performing grounded emotions system, with a happyness F-measure of 74%, a sadness F-measure of 56% and an overall accuracy of 66.9%, all 3 metrics being the highest we obtained on this dataset. It is quite surprising to think that the emotion of a tweet can be predicted with a 66.9% accuracy from information external to the tweet itself, and it lends additional credibility to the fact that people can be primed toward a particular emotional response. This variation reaches the highest correlation with the gold standard, at 0.3. Comparing this classifier with the performance of the classifier trained on the sentiment score derived from the content of the tweet itself (variation Sentiment ML in Table 3), we note that we are able to outperform it across all metrics: happiness F-measure (from 73% to 74%), sadness F-measure (50% to 56%), overall accuracy (64.8% to 66.9%), and correlation (from 0.24 to 0.3). The latter aspect shows that extrinsic features to the tweet exhibit a higher correlation with author emotion than sentiment intrinsic to the tweet.

In order to better understand the behavior of extrinsic features, we also evaluate the scenario where we do not take any user sentiment predisposition based on historical tweets into consideration (Table 4 combination all features-predisposition). This variation achieves a 71% F-measure on the happiness class, a 56% F-measure on the sadness class, and an overall accuracy of 63.75%, further lending evidence that attributes completely unrelated to a user can have a significant impact on his/her emotional response.

We should note that all combinations involving predisposition-based signals (rows 1, 2, 3, and 5 in Table 4) are statistically significant compared to predisposition-only based predictions (second row in Table 3), at p equal 0.05.

Ultimately, the various experiments we performed show that external factors do have an impact and participate in priming users toward a particular emotional response, allowing us to predict their emotion with 66.9% accuracy without analyzing the actual content of the tweet they posted.

7. Predisposition was abbreviated as *pred*.

6. Conclusion

While traditionally researchers have looked at textual content to derive an author's emotion, this work focuses on grounding emotions through the various aspects of our day-to-day living, and seeks to identify the elements that prime us toward a particular emotional response. We explored five feature types, namely weather, news exposure, social media participation timing, social network sentiment charge, and sentiment predisposition based on the tweet history of the user. At an individual feature type level, we show that a user's prior textual content exhibits a high correlation with an emotional response experienced twelve hours later, showing that users are consistent in their emotional states. We also show that the cumulative sentiment expressed in news is the second best predictor of user emotion. By combining all grounding signals together, we are able to obtain an emotional predictive accuracy of 66.9%, surpassing the majority class baseline of 59%, as well as a system using a state-of-the-art sentiment detection leveraging the content of tweets at 64.8%. This study not only shows that external factors do prime us toward emotional responses, but also that the performance of such external features can surpass the predictive accuracy of sentiment annotating systems with access to tweet content.

The data set is available for download at http://lit.eecs.umich.edu/downloads.html.

Acknowledgements

This material is based in part upon work supported by the Michigan Institute for Data Science, by the National Science Foundation (grant #1344257), and by the John Templeton Foundation (grant #48503). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the Michigan Institute for Data Science, the National Science Foundation, or the John Templeton Foundation.

References

[1] L. Barrett and E. Bliss-Moreau, *Affect as a Psychological Primitiver*. Elsevier, 2009, pp. 167–218. [Online]. Available: https://doi.org/10.1016/S0065-2601(08)00404-8

[2] S. Golder and M. Macy, “Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures,” *Science*, vol. 33, no. 6051, pp. 1878–1881, 2011.

[3] K. J. Rohan and S. T. Sigmon, “Seasonal mood patterns in a northeastern college sample,” *Journal of Affective Disorders*, no. 59, pp. 85–96.

[4] J. J. A. Denissen, L. Butalid, L. Penke, and M. A. G. van Aken, “The Effects of Weather on Daily Mood: A Multilevel Approach,” *Emotion*, vol. 8, no. 5, pp. 662–667, 2008.

[5] A. D. I. Kramera, J. E. Guilloryb, , and J. T. Hancockb, “Experimental evidence of massive-scale emotional contagion through social networks,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 29, pp. 10 779–10 779, 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24994898{\%}5Cnhttp://www.pnas.org/cgi/doi/10.1073/pnas.1412469111

[6] C. Cadwalladr, “The great British Brexit robbery: how our democracy was hijacked,” *The Guardian*, May 2017. [Online]. Available: https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy#comments

[7] K. Park, S. Lee, E. Kim, M. Park, J. Park, and M. Cha, “Mood and weather: Feeling the heat?” in *International Conference on Weblogs and Social Media (ICWSM 2013)*, 2013, pp. 709–712.

[8] A. Hannak, E. Anderson, L. F. Barrett, S. Lehmann, A. Mislove, and M. Riedewald, “Tweetin ’ in the rain: Exploring societal-scale effects of weather on mood,” in *International Conference on Web and Social Media (ICWSM 2012)*, Dublin, Ireland, 2012, pp. 479–482.

[9] A. Go, R. Bhayani, and L. Huang, “Twitter sentiment classification using distant supervision,” Tech. Rep., 2009.

[10] A. Bermingham and A. F. Smeaton, “Classifying sentiment in microblogs: Is brevity an advantage?” in *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, ser. CIKM 2010. New York, NY, USA: ACM, 2010, pp. 1833–1836.

[11] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, “Sentiment analysis of Twitter data,” in *Workshop on Language in Social Media (LSM 2011)*, no. June, Portland, OR, USA, 2011, pp. 30–38.

[12] H. Saif, Y. He, and H. Alani, *Semantic Sentiment Analysis of Twitter*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 508–524. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-35176-1{_}32

[13] M. Ghiassi, J. Skinner, and D. Zimbra, “Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network,” *Expert Systems with Applications*, vol. 40, no. 16, pp. 6266–6282, 2013.

[14] E. Martinez-Camara, M. T. Martin-Valdivia, L. A. Urena-Lopez, and A. Montejo-Raez, “Sentiment analysis in twitter,” *Natural Language Engineering*, vol. 20, no. 1, p. 128, 2014.

[15] V. S. Pagolu, K. N. R. Challa, G. Panda, and B. Majhi, “Sentiment analysis of Twitter data for predicting stock market movements,” in *Proceedings of the International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES 2016)*, 2016, p. 6. [Online]. Available: http://arxiv.org/abs/1610.09225

[16] J. Smailović, M. Grčar, N. Lavrač, and M. Žnidaršič, *Predictive sentiment analysis of Tweets: A stock market application*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 77–88. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-39146-0{_}8

[17] S. Asur and B. A. Huberman, “Predicting the future with social media,” in *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01*, ser. WI-IAT 2010. Washington, DC, USA: IEEE Computer Society, 2010, pp. 492–499. [Online]. Available: http://dx.doi.org/10.1109/WI-IAT.2010.63

[18] W. Chamlertwat and P. Bhattarakosol, “Discovering consumer insight from Twitter via sentiment analysis.” *Journal of Universal Computer Science*, vol. 18, no. 8, pp. 973–992, 2012.

[19] 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.

[20] C. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky, “The Stanford CoreNLP Natural Language Processing Toolkit,” in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 2014, pp. 1–6.

[21] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[22] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software: An update,” *SIGKDD Explorations*, vol. 11, no. 1, 2009.

[23] J. C. Platt, “Fast training of support vector machine using sequential minimal optimization,” in *Advances in Kernel Methods - Support Vector Learning*. MIT Press, January 1998.

[24] R. Quinlan, *C4.5: Programs for machine learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.