

Early Risk Detection of Anorexia on Social Media

Diana Ramírez-Cifuentes, Marc Mayans, and Ana Freire

Web Science and Social Computing Research Group
Universitat Pompeu Fabra, Barcelona
Carrer Tanger, 122-140, 08018 , Barcelona, Spain
{diana.ramirez,ana.freire}@upf.edu
marc.mayans01@estudiant.upf.edu

Abstract. This paper proposes an approach for the early detection of anorexia nervosa (AN) on social media. We present a machine learning approach that processes the texts written by social media users. This method relies on a set of features based on domain-specific vocabulary, topics, psychological processes, and linguistic information extracted from the users' writings. This approach penalizes the delay in detecting positive cases in order to classify the users in risk as early as possible. Identifying anorexia early, along with an appropriate treatment, improves the speed of recovery and the likelihood of staying free of the illness. The results of this work showed that our proposal is suitable for the early detection of AN symptoms.

Keywords: Early Risk Detection, Eating Disorders, Social Media, Anorexia, Machine Learning.

1 Introduction

Eating Disorders (ED) are characterized by abnormal attitudes towards food and unusual eating habits [3]. Every 62 minutes, at least one person dies as a direct result from an eating disorder [8]. Anorexia Nervosa is an ED defined by the restriction in eating to keep a low weight [3]. With a mortality rate of 5% per decade, AN has the highest mortality rate of all mental disorders [8]. Due to the fact that the symptoms associated with mental illnesses have been proved to be observable on social media [21], different automated methods to detect them have been designed. The review made by Guntuku's et al. [16] shows that most of these methods are based on the analysis of user-generated data on online social networks, Web forums and blogs.

Early intervention for eating disorders is essential. According to the findings of Treasure et al. [29], when adolescents with AN are given family-based treatment within the first three years of the illness onset, they have a much greater likelihood of recovery. The current automated methods to detect mental illnesses, and eating disorders within them, are based on machine learning approaches that do not consider the delay in detecting positive cases. This issue is addressed by

the work of Losada et al. [19], where the proposal of a temporal-aware risk detection benchmark, complements the evaluation of the accuracy of the decisions taken by the algorithms.

We develop an approach suitable for the early detection of AN symptoms using a labeled dataset corresponding to the eRisk 2018¹ research collection [19], which contains writings posted in Reddit². This approach does not aim to be a complete system for the AN detection, but an algorithm that could be used for the development of a platform for detecting users in risk.

This paper is structured as follows: Section 2 reports the related work in detecting eating disorders on social media and the application of early risk measures. Section 3 shows our research proposal, describing mainly the feature extraction process and the learning algorithms used for both tasks. Our experimental setup is reported in Section 4, followed by our results and findings in Section 5. Finally, Section 6 summarises our conclusions.

2 Related Work

In 2003, eating disorders represented the third most common chronic illness in adolescent females worldwide. The prevalence of AN was about 0.3%, whereas Bulimia Nervosa (BN) was more common, with a prevalence of about 1% in young women and 0.1% in men [32]. In Europe, according to a more recent study conducted in 2016, AN was reported by 1-4% of women, and 0.3-0.7% of men. Among these people, only about one-third was detected by health-care [17]. Moreover, young people with ages between 15 to 24 years old with anorexia have 10 times the risk of dying compared to their same age peers [20].

If left untreated, eating disorders tend to become more severe and less receptive to treatment [20]. This can provide an insight on how important is to detect their symptoms as early as possible. In the Web, Pro-eating disorder sites usage is prevalent among adolescents with these conditions [33]. Moreover, their engagement with this type of content has recently been suggested as a screening factor for these kind of illnesses [6]. On social media, people with eating disorders, such as anorexia and bulimia, can be identified by the usage of certain keywords that characterise and promote these conditions [2, 31]. In these sense, features or variables that have been extracted from labeled user-generated data [16] are used to generate predictive models capable of doing an automated analysis of social media data. Based on the related work done for detecting mental illnesses in online social platforms, the analysed data is obtained either by diagnosing participants with the usage of surveys [14, 26, 30], or by crawling directly the data from public online sources like Reddit, Twitter or Facebook [2, 11, 24].

In order to build predictive models, the most common features extracted for analysing and predicting mental illnesses are those related to the texts written by the users, such as: topics [23, 28, 30], frequencies of words or combinations of

¹ <http://early.irlab.org/>

² <http://www.reddit.com/>

words (N-grams) [28, 30], and features obtained using dictionaries like LIWC³, which can provide an insight on the usage of self references, social words and emotions [13, 14]. Related works have studied the users' posting frequency in different periods of the day and year [2, 10, 13], and have also obtained features from the relationships between users, taking into account the number of friends, or followers [2, 14]. Additionally, some studies use features based on sentiment analysis, considering the subjectivity or polarity of a phrase [13, 14, 30].

To the extent of our knowledge, building predictive models to detect early risk of ED is not a widely explored task yet. For detecting other mental illnesses, such as depression, some works have attempted to do their analysis with data prior to the diagnosis [14, 30]. In particular, the work of Losada et al. [19] proposes a new metric to measure the effectiveness of early alert systems. This metric, known as Early Risk Detection Error (ERDE) penalizes the delay in detecting positive cases, and is suitable to evaluate our proposal.

3 Proposal

The main objective of our proposal is to detect positive cases of anorexia as soon as possible, minimizing the ERDE and maximizing the F1 Score. We use machine learning techniques that combine a set of features extracted from the concatenated writings of users on social media. An approach based on the *dynamic strategy* proposed in [19] is used. This method consists in building incrementally, writing per writing, a representation of each user, and applying a classifier, which was previously trained with all the users texts. Following this approach, depending on the algorithm used, a decision is made if the classifier outputs a confidence value above a given threshold.

Our approach modifies the dynamic strategy of [19] by defining a minimum amount of information that should be seen by the system before applying the classifier and emitting a positive decision. We include this threshold to avoid false positives when the posts contain very few words. For instance, if the first post of a user is read and it contains five words with the word *laxative* mentioned twice, our classifier might give it a high score and classify the user as anorexic having seen only one short post. The threshold -number of words- is defined by the *text length threshold TLT* (see Eq. 1). To define the TLT we first assume that each user has a fixed number of words per post, denoted as *maxPostLenght*. To calculate this number we plotted a histogram to visualise the distribution of the number of words per post of all the users with anorexia (see Figure 1). We could observe that 80% of the users wrote up to 90 words per post. Based on Pareto's principle we chose this number of words for the *maxPostLenght* value. We asumed that seeing just one post of a fixed size was not enough, we considered that exploring more posts would reduce the amount of false positives. This number of posts, was defined by a percentage *selectionPercentage* of the *average number of posts per user*, 372.6 in our case, which was denoted as *avgPosts*. For our

³ <http://liwc.wpengine.com/>

experiments we chose to work with 10% for the value of *selectionPercentage*. For the analysed dataset the TLT value resulted in 3330 words.

$$TLT = maxPostLength \times avgPosts \times selectionPercentage \quad (1)$$

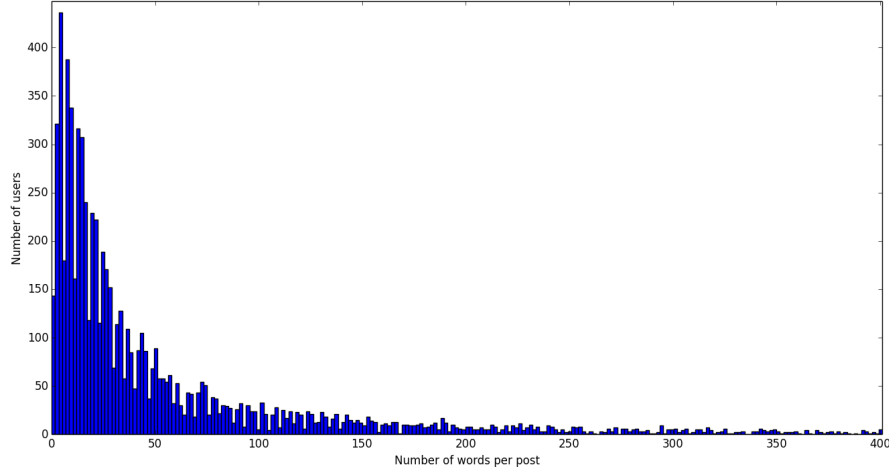


Fig. 1. Histogram of the number of words per post.

3.1 Feature Extraction

We fed our models with features that characterise the content of the writings. Further details are explained below and summarised in Table 1.

Psychological and linguistic processes: We calculate features to characterise the users’ writings. These features were calculated by taking into account the frequency of words belonging to the categories of the LIWC2007 dictionary [22], which has been previously used in detecting mental health issues [9,10]. In this sense, scores that consider linguistic and psychological processes, as well as personal concerns and spoken categories were obtained. We consider a new feature value for each category defined in the LIWC2007 dictionary. The list and description of these categories can be found in [22]. The scores were calculated normalising the frequencies of words by the total number of words in the writings of a user. Given that certain words could belong to multiple categories, the normalization value was augmented in one each time a word was part of more than one category.

Domain-related vocabulary: 9 features were defined by creating categories of words that belong to domains related with anorexia. The vocabulary for these categories was obtained from the codebook’s domains and sample keywords defined in [2]. The domains are: anorexia, body image, food and meals, eating, caloric restriction, binge, compensatory behavior, and exercises. These features were calculated in the same way as the psychological and linguistic processes features.

N-grams: They consist of sets of co-occurring words within a given window (N). Studies have extensively used them in text mining and natural language processing tasks [15]. Considering that previous works have considered them as features for detecting depression and eating disorders [28, 30], we did a $tf \cdot idf$ vectorization of the unigrams and bigrams of the training set writings. For this step, we used the *TfidfVectorizer* from the *scikit-learn* Python library⁴, with a stop-words list and the removal of the n-grams that appeared in less than 20 documents. The content of a document was defined by the concatenation of all the writings of a user.

Topic modelling: Topic Modeling consists on automatically extracting and identifying topics that are present in documents in order to obtain hidden patterns of a corpus. A known method proposed by David Blei et al. [4] is the Latent Dirichlet Allocation (LDA), which is an unsupervised generative statistical model in which the topics are represented by a set of terms or words. Many authors evidence that, in tasks of prediction and classification, the use of this method is significant and relevant. For instance, the authors in [27, 30, 34] conclude that features based on topic modeling are useful in tasks for recognising depressive, and suicidal users. Also, this technique has been used in [7], combined with other features, to quantify and predict the mental illnesses severity in online pro-eating disorder communities.

To define the topics we used English stopwords and only considered the words that appeared at least in 10% of the training documents. The 50 features used by the model are given by the probabilistic distribution of 50 topics for each analyzed text. The *LatentDirichletAllocation* module from the *scikit-learn* Python library⁴ was used to do this implementation.

Table 1. Features considered

Feature Type	Details and resources	Number of features
Linguistic and psychological processes	LIWC	64
Domain-related vocabulary	Anorexia vocabulary	9
N-grams	Unigrams	4303
N-grams	Bigrams	667
Topic modeling	Topics using LDA	50

⁴ <http://scikit-learn.org/>

3.2 Learning Algorithms

We explored four different prediction methods, i.e., Logistic Regression, Random Forest, Support Vector Machine (SVM), and Multilayer Perceptron (MLP) since they have been used previously as classifiers for similar tasks [18, 19, 23]. They are briefly explained below:

Logistic Regression (LR): it stands as a statistical method used to predict a binary outcome given a set of independent variables. This algorithm fits data to a logit function in order to predict the probability of occurrence of an event [1].

Random Forest (RF): this is a classification method that works by building many decision trees at training time. For the classification tasks, its output is the class that is the mode of the classes of the individual trees [5].

Support Vector Machine (SVM): algorithm that finds a decision plane that maximizes the distance between the different data classes to classify [12]. We have used a linear kernel and balanced class weights.

Multilayer Perceptron (MLP): this algorithm is a feed-forward artificial neural network [25]. We implemented an MLP with one hidden layer of 200 neurons and a logistic activation function. The solver for weight optimization is the Limited-memory BFGS, which is an optimization algorithm in the family of quasi-Newton methods that approximates the Broyden–Fletcher–Goldfarb–Shanno algorithm.

4 Experimental Setup

Our experiments are conducted over the eRisk 2018 research collection [19], which contains a labeled dataset with writings of a control group and of people diagnosed with anorexia as detailed in section 4.1. *Python 3.6.5*⁵ and, in particular, the *scikit-learn Python* library was used for the implementation of the proposed methods.

Using the provided data we defined training and test sets. To train our models we applied 10-fold cross validation and optimized the parameters through grid search. Each instance for our training task represented a user, and was defined by the features mentioned in section 3.1. These features were extracted from all the sequentially-concatenated writings of each user. The test set allowed us to evaluate the behavior of the dynamic method, where the classifiers were applied each time a new writing was read. Also, this set was used to define a threshold that represented the minimum probability value required by an instance to be classified as positive. After having tested different values for the LR and RF

⁵ <https://docs.python.org/3/>

classifiers, this threshold was set as 0.75 and 0.55 respectively. Moreover, a minimum number of words (3330) defined by the TLT (1), had to be processed before emitting a positive decision.

4.1 Dataset Description

Our method analysed a collection composed by chronologically ordered writings (posts or comments) from a set of Reddit users [19]. Users were labeled as anorexic and non-anorexic. The dataset statistics are detailed in Table 2.

Table 2. Main statistics of the train and test collections

	Train		Test	
	Anorexia	Control	Anorexia	Control
Num. subjects	20	132	41	279
Num. Writings	7,452	77,514	17,422	151,364
Avg num. writings	372.6	587.2	424.9	542.5
Avg num. words per submission	41.2	20.9	35.7	20.9

4.2 Evaluation

For the evaluation of the performance of our methods we report the Precision, Recall and F1-Score. In addition to this commonly known measures, we evaluate our proposal in terms of the ERDE [19], which gives a cost c to each binary decision d taken by the system at a number k of textual items seen before making a decision. This error is defined in our case as:

$$ERDE_o(d, k) = \begin{cases} c_{fp} & \text{if } d = \text{False Positive (FP)} \\ c_{fn} & \text{if } d = \text{False Negative (FN)} \\ lc_o(k) \cdot c_{tp} & \text{if } d = \text{True Positive (TP)} \\ 0 & \text{if } d = \text{True Negative (TN)} \end{cases} \quad (2)$$

where $c_{fp} = 0.13$, $c_{fn} = 1$, $c_{tp} = 1$, and the cost function $lc_o(k)$ is a sigmoid function to penalize the late emissions:

$$lc_o(k) = 1 - \frac{1}{1 + e^{k-o}} \quad (3)$$

Notice that the c_{fp} value was set according to the proportion of positive cases in the data and o is a parameter of (3) which defines the point at which the cost grows more quickly [19].

4.3 Baselines

In order to test our approach we set four different baselines based on Losada et al. [19] since they have a similar objective and a dataset extracted from Reddit as well. The baselines are described as follows:

Random: This is a naïve strategy that makes a random decision after seeing the first message. This decision could be positive (“anorexic”) or negative (“non anorexic”). This strategy is fast but we assume its effectiveness will be poor.

Minority: This is another naïve strategy that predicts a positive decision (“anorexic”) for each user right after processing the first message. Again, this strategy makes a quick decision but it is expected to have poor effectiveness.

First n : This strategy consists on concatenating the first n submissions from each user and getting the prediction provided by our model. If the user has less than n posts, all the data available for that user is used.

Dynamic: This strategy consists on analyzing and building incrementally a representation of each user. Messages are concatenated one per one, and a prediction is made by our model each time a new message is added. This is done until a positive decision is reached. The system can emit a positive decision (“anorexic”) only if the classifier outputs a confidence value above a given threshold (0.75). This method does not work with a fixed number of messages. If the stream of texts is over, the method concludes with a negative decision (“non anorexic”).

For the *Dynamic* and the *First n* baselines we have employed a quite simple model, which consisted in a Logistic Regression classifier trained and tested with features extracted from the posts. These features were based on the *tf-idf* vectorization considering unigrams.

5 Results

The evaluation of the proposed baselines can be found in Table 3. The best results are obtained by the *Dynamic* strategy with a confidence threshold of 0.75. In terms of precision, the *First n* strategy offers better results. As we expect, the *Random* and *Minority* strategies have a poor performance. Also we can see clearly that the best results of the ERDE metric are obtained by the *Dynamic* strategy, which is why we choose its results as our baseline to be compared with the models of our proposal.

In Table 4 we report the results obtained after running the learning algorithms described in Section 3.2 in combination with the features described in Section 3.1. We can see that the usage of features based in the LIWC dictionary, the TLT value and the LDA features have improved the results of the baseline, in terms of the F1 score and ERDE100, using the same learning algorithm (LR).

For the Support Vector Machine classifier, we can see that the algorithm improves significantly the baseline in terms of the F1 score (0.85), and ERDE100 (4.22%). We consider this as our best model.

In the case of the Random Forest algorithm, the results have worsened in terms of the F1 score having a low Recall value but the best precision of all our

Table 3. Baselines (Linear Regression). In bold, the best result for each evaluation metric.

	F1	P	R	ERDE5	ERDE10	ERDE50	ERDE100
Random	0.18	0.11	0.44	13.05%	12.95%	12.95%	12.95%
Minority	0.23	0.13	1.00	11.40%	11.17%	11.17%	11.17%
First 10	0.46	0.81	0.32	12.87%	11.17%	9.46%	9.46%
First 100	0.72	0.73	0.71	13.17%	13.15%	11.53%	8.80%
First 500	0.72	0.64	0.83	13.29%	13.27%	11.65%	11.42%
Dynamic	0.72	0.67	0.78	9.89%	8.76%	6.97%	5.80%

models (0.94). In terms of the ERDE it does not get better results compared to our baseline.

With the Multilayer Perceptron algorithm, the F1 score has improved, with a higher precision, but the Recall value has decreased compared to the dynamic strategy of the baseline.

Referring to our results, we can see that for all our models the ERDE5 and ERDE10 scores are high due to the TLT value that we have defined in order to decrease the number of false positives. In this sense we have tested the best algorithm (SVM) results, with the same features but without the TLT restriction. Table 5 reports these results. The most remarkable thing is the improvement in the ERDE metric, which is important considering our attempt to detect positive cases as early as possible.

Table 4. Proposed classifiers with the modified Dynamic strategy. In bold, the best result for each evaluation metric.

	F1	P	R	ERDE5	ERDE10	ERDE50	ERDE100
LR + LDA + TFIDF + LIWC	0.76	0.72	0.80	13.33%	13.31%	7.86%	5.05%
SVM + LDA + TFIDF + LIWC	0.85	0.85	0.85	13.05%	13.04%	7.26%	4.22%
RF + LDA + TFIDF + LIWC	0.53	0.94	0.37	12.85%	12.85%	11.90%	10.43%
MLP + LDA + TFIDF + LIWC	0.77	0.81	0.73	13.09%	13.09%	8.01%	5.90%

Table 5. Comparison between the best model (SVM + LDA + TFIDF + LIWC) with TLT and without TLT. In bold, the best result for each evaluation metric.

	F1	P	R	ERDE5	ERDE10	ERDE50	ERDE100
SVM + LDA + TFIDF + LIWC	0.85	0.85	0.85	13.05%	13.04%	7.26%	4.22%
SVM (NO TLT)	0.80	0.73	0.88	8.38%	7.41%	4.43%	3.80%

6 Conclusions and future work

In this paper we have proposed models for the early detection of cases of anorexia on social media. We present a temporal-aware approach, which aims to reduce the penalisation of the delay for detecting positive cases. Different machine learning models were built and tested. These models were fed with features based on linguistic information, domain-specific vocabulary, topics and psychological processes. In terms of the F1 Score and ERDE of our best models, the results have shown that the proposed approaches are suitable for the early detection of anorexia as they clearly improve our baselines.

As a future work, new features and learning algorithms will be tested. We will study in depth the introduction of voting methods, as they have been previously applied with success in similar cases [18]. Finally, we will investigate how to overcome the trade-off between avoiding the prediction of false positives and reducing the time needed to do the prediction. This means that we will explore a way to get a better precision without getting a high ERDE.

Acknowledgements

This work was supported by the Spanish Ministry of Economy and Competitiveness under the Maria de Maeztu Units of Excellence Programme (MDM-2015-0502).

References

1. Agresti, A.: Categorical Data Analysis. Wiley Series in Probability and Statistics, Wiley (2013), <https://books.google.es/books?id=UOrr47-2oisC>
2. Arseniev-Koehler, A., Lee, H., McCormick, T., Moreno, M.: #proana: Pro-eating disorder socialization on twitter **58** (04 2016)
3. Association., A.P., Association., A.P.: Diagnostic and statistical manual of mental disorders : DSM-5. American Psychiatric Association Arlington, VA, 5th ed. edn. (2013)
4. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (Mar 2003), <http://dl.acm.org/citation.cfm?id=944919.944937>
5. Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (Oct 2001). <https://doi.org/10.1023/A:1010933404324>
6. Campbell, K., Peebles, R.: Eating Disorders in Children and Adolescents: State of the Art Review . <https://doi.org/10.1542/peds.2014-0194>
7. Chancellor, S., Lin, Z., Goodman, E.L., Zerwas, S., De Choudhury, M.: Quantifying and predicting mental illness severity in online pro-eating disorder communities. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. pp. 1171–1184. CSCW '16, ACM, New York, NY, USA (2016). <https://doi.org/10.1145/2818048.2819973>
8. Coalition, E.D.: Facts about eating disorders: What the research shows (2016)
9. Coppersmith, G., Harman, C., Dredze, M.: Measuring post traumatic stress disorder in twitter pp. 579–582 (01 2014)

10. Coppersmith, G., Dredze, M., Harman, C.: Quantifying mental health signals in twitter (2014)
11. Coppersmith, G., Dredze, M., Harman, C., Hollingshead, K., Mitchell, M.: Clpsych 2015 shared task: Depression and ptsd on twitter. In: CLPsych@HLT-NAACL (2015)
12. Cortes, C., Vapnik, V.: Support-vector networks. *Machine Learning* **20**(3), 273–297 (Sep 1995). <https://doi.org/10.1007/BF00994018>
13. De Choudhury, M., Counts, S., Horvitz, E.J., Hoff, A.: Characterizing and predicting postpartum depression from shared facebook data. In: *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. pp. 626–638. CSCW '14, ACM, New York, NY, USA (2014). <https://doi.org/10.1145/2531602.2531675>
14. De Choudhury, M., Gamon, M., Counts, S., Horvitz, E.: Predicting depression via social media. *AAAI* (July 2013), <https://www.microsoft.com/en-us/research/publication/predicting-depression-via-social-media/>
15. Elberichi, Z.: Text mining using n-grams (01 2006)
16. Guntuku, S.C., Yaden, D.B., Kern, M.L., Ungar, L.H., Eichstaedt, J.C.: Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences* **18**, 43 – 49 (2017). <https://doi.org/10.1016/j.cobeha.2017.07.005>, big data in the behavioural sciences
17. Keski-Rahkonen, A., Mustelin, L.: Epidemiology of eating disorders in europe: prevalence, incidence, comorbidity, course, consequences, and risk factors **29** (09 2016)
18. Leiva, V., Freire, A.: Towards suicide prevention: Early detection of depression on social media. In: Kompatsiaris, I., Cave, J., Satsiou, A., Carle, G., Passani, A., Kontopoulos, E., Diplaris, S., McMillan, D. (eds.) *Internet Science*. pp. 428–436. Springer International Publishing, Cham (2017)
19. Losada, D.E., Crestani, F.: A test collection for research on depression and language use. In: Fuhr, N., Quaresma, P., Gonçalves, T., Larsen, B., Balog, K., Macdonald, C., Cappellato, L., Ferro, N. (eds.) *Experimental IR Meets Multilinguality, Multimodality, and Interaction*. pp. 28–39. Springer International Publishing, Cham (2016)
20. Maximilian, F.M., Norbert, Q.: Mortality in eating disorders - results of a large prospective clinical longitudinal study. *International Journal of Eating Disorders* **49**(4), 391–401. <https://doi.org/10.1002/eat.22501>
21. Park, M., Cha, C., Cha, M.: Depressive moods of users portrayed in twitter pp. 1–8 (01 2012)
22. Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., Booth, R.J.: The Development and Psychometric Properties of LIWC2007. This article is published by LIWC Inc, Austin, Texas 78703 USA in conjunction with the LIWC2007 software program., <http://www.liwc.net/LIWC2007LanguageManual.pdf>
23. PreoŢiuc-Pietro, D., Eichstaedt, J., Park, G., Sap, M., Smith, L., Tobolsky, V., Schwartz, H.A., Ungar, L.: The role of personality, age, and gender in tweeting about mental illness. In: *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. pp. 21–30. Association for Computational Linguistics (2015). <https://doi.org/10.3115/v1/W15-1203>
24. Prieto, V.M., Matos, S., Alvarez, M., Cacheda, F., Oliveira, J.L.: Twitter: A good place to detect health conditions. In: *PloS one* (2014)
25. Ramchoun, H., Amine, M., Idrissi, J., Ghanou, Y., Ettaouil, M.: Multilayer perceptron: Architecture optimization and training. *IJIMAI* **4**(1), 26–30 (2016)

26. Reece, A.G., Reagan, A.J., Lix, K.L.M., Dodds, P.S., Danforth, C.M., Langer, E.J.: Forecasting the onset and course of mental illness with twitter data. In: Scientific Reports (2017)
27. Resnik, P., Garron, A., Resnik, R.: Using topic modeling to improve prediction of neuroticism and depression in college students pp. 1348–1353 (01 2013)
28. Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Park, G., Sap, M., Stillwell, D., Kosinski, M., Ungar, L.H.: Towards assessing changes in degree of depression through facebook (2014)
29. Treasure, J., Russell, G.: The case for early intervention in anorexia nervosa: theoretical exploration of maintaining factors. *British Journal of Psychiatry* **199**(1), 5–7 (2011). <https://doi.org/10.1192/bjp.bp.110.087585>
30. Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh, Y., Ohsaki, H.: Recognizing depression from twitter activity. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 3187–3196. CHI '15, ACM, New York, NY, USA (2015). <https://doi.org/10.1145/2702123.2702280>
31. Wang, T., Brede, M., Ianni, A., Mentzakis, E.: Detecting and characterizing eating-disorder communities on social media. In: WSDM (2017)
32. Wijbrand, H.H., van Hoeken Daphne: Review of the prevalence and incidence of eating disorders. *International Journal of Eating Disorders* **34**(4), 383–396. <https://doi.org/10.1002/eat.10222>
33. Wilson, J.L., Peebles, R., Hardy, K.K., Litt, I.F.: Surfing for thinness: a pilot study of pro-eating disorder web site usage in adolescents with eating disorders. *Pediatrics* **118** 6, e1635–43 (2006)
34. Zhang, L., Huang, X., Liu, T., Chen, Z., Zhu, T.: Using linguistic features to estimate suicide probability of chinese microblog users. *CoRR* **abs/1411.0861** (2014), <http://arxiv.org/abs/1411.0861>