

# Sentiment Transitions in Online Medical Forums

Authors

Affiliation / Address line 1

email@domain

No Institute Given

**Abstract.** This work studies sentiment transitions on an online medical forum where users correspond in English. We work with discussions dedicated to reproductive technologies, an emotionally-charged issue. We analyze authors' activity on the forum, communication flow in discussions and sentiment flow in consecutive messages. We show how sentiment transitions can be successfully used to disambiguate sentiments. In several learning problems, we demonstrate that multi-class sentiment classification significantly improves when messages are represented by combination of affective terms and sentiment transitions (paired t-test,  $P=0.0011$ ).

## 1 Introduction

Online public forums discuss personal experience and often convey the sentiments and emotions of the forum participants. Personal sentiments expressed in the posted messages<sup>1</sup> set interaction patterns among the members of online communities and have a strong influence on the public mood [7, 20].

In health-related online communities, emotional factors may prevail over the informative content. Surveys of medical forum participants revealed that personal testimonials attract attention of up to 49% of participants, whereas only 25% of participants are motivated by scientific and practical content [3]. In a manual survey of online infertility support groups, empathy and shared personal experience constituted 45.5% of content, gratitude – 12.5%, recognized friendship with other members – 9.9%, whereas the provision of information and advice and requests for information or advice accounted for 15.9% and 6.8% respectively [15]. Shared online emotions can improve personal well-being and empower patients in their battle against an illness [8, 14].

Sentiment transition has become a popular topic in sentiment analysis following maturity of online communities and availability of relevant data [19]. It has been shown that accuracy of sentiment classification in consecutive messages strongly depends on the topic of debates and can vary from 53% to 69%,

---

<sup>1</sup> Terms “messages” and “posts” are equivalent in this work.

with 60% accuracy obtained on health care discussions [2]. Most of such studies work with only positive and negative polarity and analyze propagation of positive and negative sentiments [12, 10].

In this study, conducted on forums where users communicate in English, we work with the categories of *encouragement*, *gratitude*, *confusion*, *facts*, *endorsement* and analyze how sentiments change in consecutive messages. Below is an example of the data we study:

im spotting more brown today... its darker brown and more of it. I'm terrified and so sad. Had to work all day and I was just fighting back the tears all day. I don't know how to do this anymore and put on that fake smile every day and act like I'm just a happy person when my heart is breaking on the inside. **confusion**

Have you tested yet? I know a lot of people get this brown spotting, and even red, and are still pregnant. It's so hard going through this, I hope it works out for the best for you. **encouragement**

The paper is organized as follows: in section 2 we describe related work, followed by data analysis in section 3, then we present our sentiment transition and sentiment classification results in sections 4 and 5, and our discussion and conclusions in sections 6 and 7.

## 2 Related Work

Sentiment analysis of health-related messages has been stipulated by the demands of policy-makers and the emergence of infoveillance and infodemiology [6, 17]. Health-related affects and emotions are often analyzed on Twitter data [6, 5] and public forums dedicated to health [1, 11, 15]. Further, we limit the discussion to research areas directly relevant to the current project.

The studies of sentiment propagation in social networks is an emerging area of sentiment analysis [7, 19]. Results obtained for three online communities have shown that the original polarity is preserved in the follow-up posts: negative posts mainly trigger negative follow-ups, positive posts - positive follow-ups, and neutrality brings forth neutrality [7]. Sentiment propagation between two Twitter interlocutors has been studied in several works, [13] being the most cited. In it, the authors selected sequences of tweets (aka chains) based on positive and negative emoticons. These chains were analyzed for the propagation of 6 emotions: *love*, *joy*, *surprise*, *anger*, *sadness*, and *fear*. Our current study is different in 3 aspects: we study discussions carried out by many participants, the text of forum messages is quite different from tweets, and we analyze another set of sentiment categories.

In sentiment analysis of health-related messages, some studies only manually analyze the data [15] or use a keyword search without discussing sentiment

annotation [6, 9]. We, on the other hand, present results of sentiment annotation and compare them with machine learning results.

General and domain-specific affective lexicons were used by [11] and [18]. Both studies showed a preference for domain-specific lexicons over general lexicons. [11] applied SentiWordNet, Subjectivity Lexicon and a polarized domain lexicon to classify drug reviews into positive, negative and neutral. Combination of the three lexicons produced the best classification of positive reviews ( $F\text{-score}_{pos} = 0.62$ ) and the best classification of negative reviews ( $F\text{-score}_{neg} = 0.32$ ); the authors did not report the overall classification results. [18] adapted Pointwise Mutual Information to build a domain lexicon HealthAffect. In 6-class sentiment classification, HealthAffect produced an  $F\text{-score} = 0.52$  whereas WordNetAffect produced an  $F\text{-score} = 0.30$ . [2] automatically classified difference in opinions within discussion threads. They represented text by unigrams, Linguistic Inquiry and Word Count (LIWC), and the MPQA dictionary. For HealthCare discussions the best accuracy 60.64% was achieved when the text was represented by LIWC. At the same time, classification of health-related sentiments is mainly constrained to individual messages. Our work on sentiment transitions goes beyond individual messages.

Few publications discussed manual annotation of sentiments in online messages. Topic-specific opinions in blogs were evaluated in [16]. Agreement among seven manual annotators was computed for five classification categories, including positive, negative, mixed opinions and non-opinionated and non-relevant categories. [5] asked multiple annotators to categorize tweets into positive and negative sentiments and neutral tweets. Manual annotation was applied in multi-class sentiment classification in [18]. The cited works did not consider how ambiguous annotations should be re-solved. Our study, on other hand, suggests a disambiguation procedure to resolve annotation conflicts.

### 3 Discussion Data and Sentiment Categories

We work with data obtained from the In Vitro Fertilization (IVF) website<sup>2</sup>. The data was introduced in [18] and became available on demand. On the IVF forum, 95% of the participants are women. Amidst high emotional tension, the participants discuss IVF treatment, fertility clinics, medications and expected outcomes, as shown in the example in Section 1. We analyze discussions from the IVF Ages 35+ sub-forum<sup>3</sup>. Discussions with  $< 10$  posts usually consisted of a question and factual replies. Discussions with  $> 20$  posts usually ventured outside the initial topic. As a result, we kept 80 discussions, with 10 – 20 posts

<sup>2</sup> [www.ivf.ca/forums](http://www.ivf.ca/forums)

<sup>3</sup> <http://ivf.ca/forums/forum/166-ivf-ages-35>

each, and 1322 posts for analysis; the average was 17 posts per discussion. In total, 359 participants contributed to these discussions.

We implemented a two-step process of sentiment annotation. In the first step, annotators were asked to mark messages in an unsupervised manner (i.e., we did not provide them with pre-set sentiment categories). The result was a selection of 35 types of sentiments, including worry, concern, doubt, impatience, support, hope, happiness, enthusiasm, thankfulness. In the next step, we gradually generalized them into three groups:

**confusion** including *worry, concern, doubt, impatience, uncertainty, sadness, anger, embarrassment, hopelessness, dissatisfaction, dislike;*

**encouragement** including *cheering, support, hope, happiness, enthusiasm, excitement, optimism;*

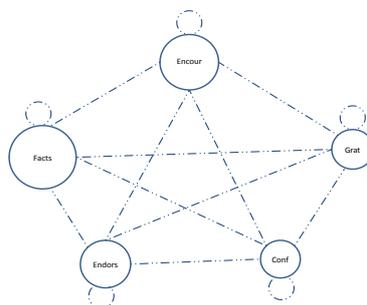
**gratitude** including *thankfulness.*

Each post was independently assigned by two annotators with one dominant sentiment. The annotators reached a strong agreement with Fleiss Kappa = 0.737. If annotators assigned a message to different categories, the message was labeled *ambiguous*.

Note that the three sentiment categories did not have explicitly negative sentiments; instead, we considered *confusion* as a non-positive label whereas *encouragement* and *gratitude* were considered as positive labels. A special group of sentiments was presented by expressions of *compassion, sorrow, and pity*. According to the WordNetAffect classification, these sentiments should be considered negative. However, in the context of health discussions, these emotional expressions appeared in conjunction with moral support and encouragement. Hence, we treated them as a part of *encouragement*.

As expected, not all posts were predominantly sentimental. Participants speak about treatment outcomes, share details about medical procedures and clinics, and discuss symptoms and conditions. Messages presenting only factual information were marked as *facts*. Some posts contained factual information followed by emotional utterances. Those utterances almost always conveyed support and implicit approval (“hope, this helps”, “I wish you all the best”, “good luck”). We called this category of messages *endorsement*. *Facts* were found to be the most frequent category, with 433 messages, followed by *encouragement* – 310 messages, *endorsement* – 162, *gratitude* – 124, *confusion*, with 117 messages, was found to be the least frequent category. There were 176 *ambiguous* messages on which annotators did not agree on the label.

Edges of the graph  $G$  in Figure 1 show all possible transitions between the sentiments, whereas the node sizes are proportional to sentiments’ volumes. Further, we aim to answer two questions:



**Fig. 1.** Possible transitions among the sentiments in discussions.

1. What sentiment transitions do occur in discussions? It is quite possible that not every sentiment is transitioned in to every other sentiment or itself.
2. How probable are those transitions?

#### 4 Sentiment Transition in the Forum Discussions

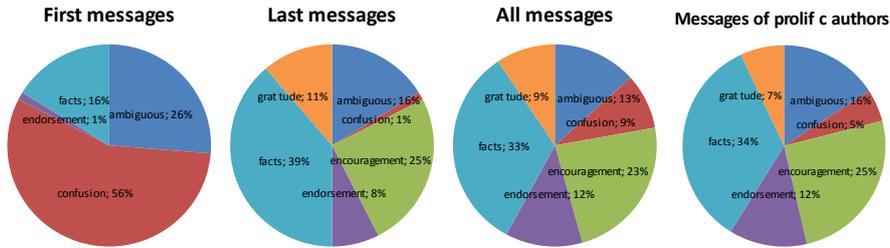
Forum discussions are intrinsically heterogenous. A large number of participants, many of whom added only one or two messages and the time delay (the last message can be posted weeks or even months after the first one) might weaken relations between messages and the expressed sentiments. We observed that the discussion could start by one participant expressing her doubts and concerns, continued by describing a treatment and come to a conclusion by the announcement of the results. All these cornerstone messages received corresponding replies. Within discussions messages were related: every posted message answered to one or several previous messages, and in many cases did not diverge from the discussed topic. At the same time, new participants were bringing new ideas and emotions. It is natural to expect that in discussions the first messages and last messages will have a pronounced difference, including assigned sentiments.

In online communities, the most active participants significantly influence the forum discussions [19, 20]. In our data, the 15 most active authors posted 15–50 messages each. Together this comes to 387 messages, or 29.3 % of the 1322 analyzed messages. We call these authors “prolific”. Their posts mostly contained factual information and encouragement. Typically, 86% messages would be equally divided between *facts* and *encouragement* and 7% – between *confusion* and *gratitude*.

Some prolific authors were adding posts with information and moral support to various discussions. However, most started at least one discussion describing their own problems and expressing concern and worry. In messages posted by

the prolific authors, *confusion* appeared less than other sentiments: 20 posts in total, or 5.2% of their messages. However, *gratitude* was the least frequent sentiment among the authors: 7 authors had no posts marked as *gratitude*. In contrast, only 4 prolific participants had no posts marked as *confusion*.

Among the authors, 10 authors were less emotional and 5 were more emotional. The 10 authors who requested information in a less emotional manner obtained similar neutral factual responses. The 5 authors with a comparatively large number of posts marked *confusion* (16 posts combined) had also a considerable number of posts marked as *gratitude* (17 posts combined). Their messages contained much more emotions and received similar replies which expressed compassion and moral support. Appreciation of this support was shown by responses of thankfulness and gratitude. The pie charts below display the proportions of sentiments found in first messages, last messages, all messages and in messages of the prolific authors.



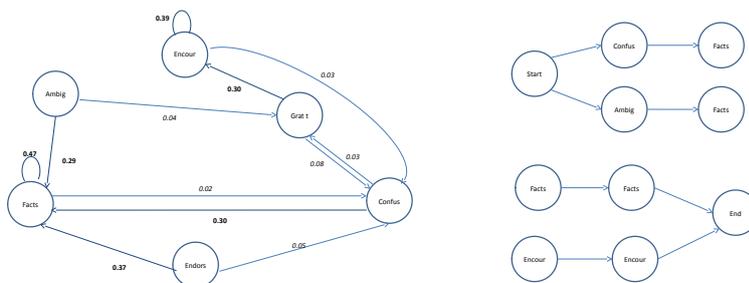
We collected information about changes of sentiments during discussions. We built a matrix  $A$  where an element  $a_{ij}$  shows occurrence of a message with sentiment  $i$  (rows) being followed by a message with sentiment  $j$  (columns) (the left side of Table 1). The row “start” indicates sentiments of the first messages of the discussions; the column “end” indicates sentiments of the last messages of the discussions. Note that barring 0s for encouragement and gratitude in the start row, the other elements of matrix  $A$  are non-zero. This means that there exists at least one transition between every pairs of categories. Relating matrix  $A$  with the graph  $G$  (Figure 1), we show that the sentiment transitions form a strongly connected balanced digraph, where for each node its indegree (i.e., the number of head points) is equal to its outdegree (i.e., the number of tail points).

Based on the transition matrix  $A$ , we built a stochastic matrix  $P$  by dividing each value by the row total. This matrix of probabilities forms a first-order Markov model, where the probability of each category of message depends on the nature of the previous message (the right side of Table 1). For each category, we mark in **bold** the greatest probability of the next category and in *italics* – the

prev	The transition matrix of sentiments							The matrix of sentiment transition probabilities						
	ambig	confus	encour	endors	facts	gratit	end	ambig	confus	encour	endors	facts	gratit	end
start	21	45	0	1	13	0	0	0.27	0.57	0.00	0.02	0.17	0.00	0.00
ambig	39	8	44	19	51	7	13	0.22	0.05	0.25	0.11	<b>0.29</b>	<i>0.04</i>	0.01
confus	14	23	31	20	39	3	1	0.11	0.18	0.24	0.16	<b>0.30</b>	<i>0.03</i>	0.01
encour	34	7	112	31	47	41	20	0.12	<i>0.03</i>	<b>0.39</b>	0.11	0.17	0.15	0.07
endors	9	7	28	33	58	16	6	0.06	<i>0.05</i>	0.18	0.22	<b>0.37</b>	0.11	0.04
facts	45	7	59	46	197	40	31	0.11	<i>0.02</i>	0.14	0.11	<b>0.47</b>	0.10	0.08
gratit	13	9	36	12	28	17	9	0.11	<i>0.08</i>	<b>0.30</b>	0.10	0.23	0.14	0.08

**Table 1.** Transitions of sentiments in discussions. For each category, the greatest transition probability is in **bold**, the smallest is in *italics*.

smallest probability of the next category; these probabilities are also shown in Figure 2. Figure 3 shows the most likely sentiments at the beginning and the end of discussions.



**Fig. 2.** Greatest and smallest transition probabilities for each category

**Fig. 3.** Most likely beginning and end of discussions

Sentiment transitions show that two categories tend to reinforce themselves: *encouragement* is most frequently followed by *encouragement*, and *facts* – by *facts*. On other hand, *confusion* and *gratitude* trigger an emotional change in discussions: *confusion* is most often followed by *facts*, and *gratitude* – by *encouragement*. A message of *confusion* is very rarely followed by one of *gratitude* or appears at the end of the discussion.

## 5 Sentiment Classification

In this section, we test how accurately Machine Learning algorithms can recognize the sentiments introduced in Section 3. We solve four multi-class classification problems:

**6-class** classification where 1322 posts are classified into *confusion*, *encouragement*, *endorsement*, *gratitude*, *facts*, *ambiguous*

**5-class** classification where the *ambiguous* class is removed and the remaining 1146 posts are classified into the other 5 classes.

**4-class** classification where 1322 posts are grouped as follow: *facts* and *endorsement* classes make up a (*factual*) class, *encouragement* and *gratitude* classes become a *positive* class, and the *confusion* and *ambiguous* classes remain.

**3-class** classification where 176 *ambiguous* messages are removed, and the remaining 1146 messages are classified as *positive*, *confusion* or *factual* as in 4-class classification.

Note that these problems treat relations between the categories differently:

- 6- and 5-class classifications ignore relations between categories and treat all correct and erroneous outputs in the same way;
- 3- and 4-class classifications acknowledge that some categories are closer to each other than to the others and recognize these relations in the output results.

We assessed the results by computing multi-class Precision (Pr), Recall (R), and F-score (F). Labeling all examples as the majority class gave the baseline evaluation: 6-class – F=0.162, 5-class – F=0.207, 4-class – F=0.281, and 3-class – F=0.356. For the benchmark comparison with other sentiment analysis methods, we represented our data by SentiWordNet; this representation had 4 039 attributes. We compared Decision Trees, K-Nearest Neighbor, Naive Bayes and Support Vector Machines (SVM) and applied 10-fold cross-validation to select the best classifier. SVM outperformed other classifiers. Here we report the results obtained by SVM.

We tested three types of message representation:

1. sentiment transition, where messages are represented by two attributes whose values are categories assigned by each annotator to the previous message, two attributes whose values are categories assigned by each annotator to the next message, and three attributes showing whether the previous, current and next messages are first, middle, or last ones – 7 attributes altogether;
2. affective words, where we represent messages through a domain-specific lexicon HealthAffect (Section 2); this representation had 1 193 attributes;
3. combination of the two representations, with 1 200 attributes.

Our experiments show that *encouragement*, *gratitude*, *confusion*, *facts*, and *endorsement* can be successfully recognized by machine learning algorithms.

Representation	# attr.	6-class			5-class			4-class			3-class		
		P	R	F	P	R	F	P	R	F	P	R	F
SentiWordNet	4039	0.385	0.405	0.387	0.485	0.485	0.478	0.502	0.525	0.509	0.632	0.635	0.631
Sentiment transit.	7	0.395	0.408	0.356	0.441	0.458	0.414	0.475	0.528	0.489	0.620	0.620	0.605
HealthAffect	1193	0.402	0.416	0.405	0.496	0.494	0.491	0.501	0.517	0.505	0.628	0.620	0.620
HealthAffect and sentiment transit.	1200	<b>0.431</b>	<b>0.430</b>	<b>0.423</b>	<b>0.512</b>	<b>0.513</b>	<b>0.511</b>	<b>0.541</b>	<b>0.549</b>	<b>0.549</b>	<b>0.663</b>	<b>0.653</b>	<b>0.660</b>

**Table 2.** Results of multi-class sentiment classifications. For each problem, the best P, R and F are in **bold**. The baseline: 6-class – F=0.162, 5-class – F=0.207, 4-class – F=0.281, and 3-class – F=0.356.

On the three representations and in solution of all the four problems, we obtained better results than the baseline and the benchmark. Our results are comparable or better than those reported in multi-classification studies [1, 2, 18]. Table 2 reports the results. It is not surprising that in all representations 6 classes were classified with the lowest F-score. Another anticipated result was achieved for 3 classes which were classified with the highest F-score on each representation. For 6- and 5- class classification, multi-categorical HealthAffect performed better than polarity-oriented SentiWordNet. On 4- and 3-class classification where the emphasis was on differentiation between positive and negative sentiments, SentiWordNet outperformed HealthAffect.

The obtained results show that information on sentiment transition is critical for reliable sentiment classification. Although sentiment transitions were represented by 7 attributes, they provided for the results which were not statistically different from the results obtained on the 1 193 affective terms (paired t-test,  $P=0.1152$ ). On the other hand, the combination of sentiment transition attributes and affective terms provided for more significant results than affective terms (paired t-test,  $P=0.0011$ ). This analysis shows that a Markov model of sentiment transition can be used to find the most likely class of message when the classifier was unable to do this definitively. Conditional Random Fields are another type of classifier which takes into account that the classes of different members of a sequence or pattern might depend on each other [4].

## 6 Discussion

In this study we encountered challenges that do not usually arise when sentiment classification is simply viewed through a prism of positive and negative polarity. Our first challenge came from the wide range of emotions found in the messages. Working with the data, the annotators perceived worry, concern, doubt, impatience, uncertainty, sadness, anger, embarrassment, hopelessness, dissatisfaction, compassion, sorry, hope, happiness, enthusiasm, excitement, optimism,

gratitude, etc. We wanted to generalize the found sentiments in a meaningful way and reduce their number. The final set of sentiments had to satisfy two requirements: it had to be adequate to cover the emotional characteristics of the messages; it should be small enough to allow the automated learning of sentiments. Our set of categories *encouragement, gratitude, confusion, facts, endorsement* was a trade-off between the two requirements.

Another challenge was intrinsic to medical forums where most participants are dealing with health-adverse issues. Their messages provide a health-related, factual content within a highly-emotional context. A message could start with doubts, then list some pragmatic details and end on a much brighter note. This multi-emotionality mixed with the facts was the main source of ambiguous manual annotations. To overcome this challenge, we suggest possible solutions, each having its *pros* and *cons*:

1. use a sentence-level annotation where every sentence in the message is labeled with one sentiment; *pros* – the annotation unit becomes smaller and more succinct, thus it should be easier for annotators to identify emotions; *cons* – the number of annotation units raises exponentially, which will cause a considerable increase in labor and annotation time;

2. keep a message-level annotation and use multiple labels for a message (e.g., facts and confusion, gratitude and facts); *pros* – while working with messages, annotators use a larger content to identify sentiments; *cons* – multi-labeled learning requires exponentially more training data than classification.

Another challenge is the portability of our study to a larger data set. When predicting sentiment of a message in a thread, we used the manual annotations of sentiments directly as features. Due to manual labeling costs, this is unrealistic in a real world setting. To reduce the costs, we hypothesize that classifying sentiments of adjacent messages can be done either through active learning or semi-supervised learning. The choice will depend on the accuracy of the sentiment classification.

## 7 Conclusions and Future Work

We have focused on sentiment transition in public discussions on online medical forums. We have identified a set of sentiments that represents emotions expressed by the forum participants. We have shown that there is at least one transition between each pair of sentiments. Solving a series of multi-class sentiment classification problems, we demonstrated that 7 attributes of sentiments transitions provided for the results which were not statistically different from the results obtained on the 1 193 affective terms (paired t-test,  $P=0.1152$ ). At the same time, adding sentiment transitions to affective terms in message rep-

resentations statistically improved the results (paired t-test,  $P=0.0011$ ). We also analyzed the sentiment distribution in the forum discussions and worked closely with the messages written by prolific authors.

Our results on sentiment distribution and sentiment transition can be used in an emerging field of social mining, especially to domains where participants post emotionally-charged messages (e.g., ecology, climate change). A Markov model can help in disambiguation when the annotators disagree on the message sentiments. This can be done by analogy with the CLAWS part-of-speech disambiguator. Although the annotators reached a strong agreement in assigning sentiments to the messages, 176 messages, or 13.3% of data, were assigned different sentiments, thus labeled as *ambiguous*. To resolve the disagreement problem, we can evoke probabilities from Table 1. Consider the following example:

For thread A, the annotators disagree on the first post – one decided *confusion*, the other *facts*. Which one was most probably correct? We know that this message was at the very start of the thread, and we also know that the second message was classed as *facts* by both annotators. Assuming that at least one annotator was correct in each case, we must decide which sequence is the more likely to start the discussion: *confusion*  $\rightarrow$  *facts*, or *facts*  $\rightarrow$  *facts*. For both sequences, we find the product of the probabilities of each step in the path. The probability of the first path is 0.17 (calculated above). For the other path, the product of the transition probabilities is  $0.17 \times 0.47 = 0.08$ . Thus the relative likelihood of the two paths is in the ratio 0.17:0.08, meaning that the sequence *confusion*  $\rightarrow$  *facts* is the more likely (and thus the first message is more likely to be *confusion* than *facts*). The same method can be used when automatic classifiers cannot give a definitive class decision. In cases where there are many different annotations for a long sequence of messages there will be too many possible sequences to find the most probable by trying them all out in turn, so a “short cut” like the Viterbi algorithm would be used, which would only consider certain “most promising” paths.

Although study of affective lexicons was not the main goal of this work, we have noticed that a smaller number of HealthAffect terms provided for comparable or better results than a considerably large number of SentiWordNet terms. We suggest further exploration of what set of the affective terms can better represent messages in learning experiments.

## References

1. Ali, T., D. Schramm, M. Sokolova and D. Inkpen. 2013 Can I hear you? Sentiment Analysis on Medical Forums, *Proceedings of IJCNLP*, p.p. 667 – 673

2. Anand, P., M. Walker, R. Abbott, J. E. Fox Tree, R. Bowmani, and M. Minor 2011 Cats Rule and Dogs Drool!: Classifying Stance in Online Debate *the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, ACL-HLT 2011*, pp. 1–9.
3. Balicco, L. and C. Paganelli. 2011. Access to health information: going from professional to public practices, *Information Systems and Economic Intelligence: 4th International Conference*
4. Blunsom, P. and T. Cohn. 2006. Discriminative word alignments with Conditional Random Fields. *Proceedings of COLING/ACL 2006*, pp. 65–72.
5. Bobicev, V., M. Sokolova, Y. Jaffer, D. Schramm. 2012. Learning Sentiments from Tweets with Personal Health Information. *Advances in Artificial Intelligence 25*, p.p. 37–48, Springer.
6. Chew, C. and G. Eysenbach. 2010. Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak. *PLoS one*.
7. Chmiel, A., J. Sienkiewicz, M. Thelwall, G. Paltoglou, K. Buckley, A. Kappas, and J. Holyst. 2011. Collective Emotions Online and Their Influence on Community Life. *PLoS one*.
8. Coulson, N. 2005. Receiving social support online: an analysis of a computer-mediated support group for individuals living with irritable bowel syndrome. *CyberPsychology and Behavior*, 8(6):580–584.
9. Culotta, A. 2013. Lightweight methods to estimate influenza rates and alcohol sales volume from Twitter messages. *Language Resources and Evaluation*, 47:217– 238.
10. Dodds, P., K. Harris, I. Kloumann, C. Bliss, and C. Danforth. 2011. Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. *PLoS one*.
11. Goeuriot, L., J. Na, W. Kyaing, C. Khoo, Y. Chang, Y. Theng and J. Kim. 2012. Sentiment lexicons for health-related opinion mining. *the 2nd ACM SIGHIT International Health Informatics Symposium*, p.p. 219 – 225.
12. Hassan, A., A. Abu-Jbara, D. Radev. 2012. Detecting subgroups in online discussions by modeling positive and negative relations among participants. *Proceedings of EMNLP*, pp. 59–70.
13. Kim, S., Bak, J., Jo, Y., Oh, A. 2012 Do You Feel What I Feel? Social Aspects of Emotions in Twitter Conversations *Proceedings of ICWSM*, p.p. 495-498.
14. Kummervold, P. E., D. Gammon, S. Bergvik, J. A. K. Johnsen, T. Hasvold, and J. H. Rosenvinge. 2002. Social support in a wired world: use of online mental health forums in Norway. *Nordic journal of psychiatry*, 56(1):59–65.
15. Malik S. and N. Coulson. 2010. Coping with infertility online: an examination of self-help mechanisms in an online infertility support group. *Patient Educ Couns*, 81(2):315-318
16. Osman, D., J. Yearwood, and P. Vamplew. 2010. Automated opinion detection: Implications of the level of agreement between human raters. *Information Processing and Management*, 46: 331–342.
17. Paul, M. and M. Dredze. 2011. You Are What You Tweet: Analyzing Twitter for Public Health. *International Conference on WSM*
18. Sokolova, M. and V. Bobicev. 2013. What Sentiments Can Be Found in Medical Forums? *Recent Advances in Natural Language Processing*, pp. 633–639.
19. Tan, C., L. Lee, J. Tang, L. Jiang, M. Zhou, P. Li. 2011. User-level sentiment analysis incorporating social networks *the 17th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*.
20. Zafarani, R., W. Cole, and H. Liu. 2010 Sentiment Propagation in Social Networks: A Case Study in LiveJournal. *Advances in Social Computing (SBP 2010)*, pp. 413–420, Springer