ARTICLE IN PRESS

Computers in Human Behavior xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

# Sentiment detection in social networks and in collaborative learning environments

Francesco Colace [a], Luca Casaburi [a], Massimo De Santo [a], Luca Greco [b,*]

[a] DIEM, University of Salerno, Fisciano, Italy
[b] DIIN, University of Salerno, Fisciano, Italy

ARTICLE INFO

*Article history:*
Available online xxxx

*Keywords:*
Information extraction
Sentiment analysis
Latent Dirichlet allocation

ABSTRACT

Daily millions of messages appear on the web, which is becoming a rich source of data for opinion mining and sentiment analysis. The computational study of opinions, feelings and emotions expressed in a text often relates to the identification of agreement or disagreement with statements, contained in comments or reviews, that convey positive or negative feelings. The detection and analysis of sentiment in textual communication is a topic attracting attention also in the context of collaborative learning in social networks, being learners actively engaged in presenting and defending ideas and opinions, as well as exchanging moods about courses with peers. In this paper, we investigate the adoption of a probabilistic approach based on the Latent Dirichlet Allocation (LDA) as Sentiment Grabber. Through this approach, for a set of documents belonging to a same knowledge domain, a graph, the Mixed Graph of Terms, can be automatically extracted. The paper shows how this graph contains a set of weighted word pairs, which are discriminative for sentiment classification. The proposed method has been tested in different context: a standard dataset containing movie reviews; a real-time analysis of social networks posts; a collaborative learning scenario. The experimental evaluation shows how the proposed approach is effective and satisfactory.



## 1. Introduction

Thanks to blogs, microblogs, social networks or reviews collector sites, millions of messages appear daily on the web. In general, this textual information can be divided in two main categories: facts and opinions. Facts are objective statements while opinions reflect people's sentiments about products, other people and events; the latter appear to be extremely important, being able to influence the decisional process (Dascalu, Bodea, Lytras, de Pablos, & Burlacu, 2014; Lytras & de Pablos, 2008, 2011; Lytras, Sicilia, Naeve, de Pablos, & Lytras, 2008; Sebastiani, 2002). The interest, that potential customers show in opinions and reviews about products, is something that vendors are gradually paying more and more attention to. Companies are interested in what customers say about their products as politicians are interested in how different news media are portraying them. The detection and analysis of sentiment in textual communication is a topic attracting attention also in the context of collaborative learning in social networks, being learners actively engaged in presenting and defending ideas and opinions, as well as exchanging diverse beliefs with peers. Many institutions are adopting e-learning and collaborative learning both to improve their traditional courses and increase the potential audience since it allows more flexibility and quality in general. Anyway, e-classrooms are often composed by students inattentive or appearing bored and wondered. So the main question for the teacher becomes: why am I not able to reach these students and catch their attention? why are they not excited about the material although my efforts to present it in an organized and coherent manner? This sense of frustration increases when he faces students' poor performance on tests. Recent studies showed that emotions can affect the e-learning experience. What are emotions? A general definition for emotions is the following: emotions are complex psychophysical processes that evoke positive or negative psychological responses (or both) and physical expressions, often involuntary. Emotions are often related to feelings, perceptions or beliefs about elements, objects or relations between them, in reality or in the imagination. They typically arise spontaneously, rather than through conscious effort. An emotion (reaction or state) is often differentiated from a feeling (sensation or impression), although the word "feeling" is used as a synonym for "emotion" in some contexts. In fact, emotion has to do with how one feels. This feeling, if positive is believed to have a productive effect on the individual; otherwise it seems to impact

* Corresponding author.
*E-mail address:* lgreco@unisa.it (L. Greco).

http://dx.doi.org/10.1016/j.chb.2014.11.090
0747-5632/© 2014 Elsevier Ltd. All rights reserved.

negatively on the individual's learning experience. Obviously, the topic of emotions goes far beyond this simple definition and it is especially hard to detect in an e-learning environment. In a face-to-face class instructors can detect facial expressions of students but, in an online environment, students need to establish an online presence and the instructors need to be able to pick up on this. In this scenario, a promising approach is the sentiment analysis: the computational study of opinions, sentiments and emotions expressed in a text (Liu, 2010; Wang et al., 2013). In literature, there are many approaches to the sentiment analysis. A very broad overview of the existing work was presented in Pang and Lee (2008). The authors describe in detail the main techniques and approaches for an opinion oriented information retrieval. Early work in this area was focused on determining the semantic orientation of documents. In particular some approaches attempt to learn a positive–negative classifier at a document level. Turney (2002) introduces the results of review classification by considering the algebraic sum of the orientation of terms as respective of the orientation of the documents. Starting from this approach other techniques have been developed by focusing on some specific tasks as finding the sentiment of words (Wilson, Wiebe, & Hwa, 2004). Baroni and Vegnaduzzo (2004) proposed to rank a large list of adjectives according to a subjectivity score by employing a small set of manually selected subjective adjectives and computing the mutual information of pairs of adjectives using frequency and co-occurrence frequency counts on the web. The work of Turney and Littman (2002) proposes an approach to measure the semantic orientation of a given word based on the strength of its association with a set of context insensitive positive words minus the strength of its association with a set of negative words. By this approach sentiment lexicons can be built and a sentiment polarity score can be assigned to each word (Gamon & Aue, 2005; Neviarouskaya, Prendinger, & Ishizuka, 2011). Sentiment polarity score means the strength or degree of sentiment in a defined sentence pattern. Artificial Intelligence and probabilistic approaches have also been adopted for sentiment mining. In Pang, Lee, and Vaithyanathan (2002) three machine learning approaches (Naïve Bayes, Maximum Entropy and Support Vector Machines) has been adopted to label the polarity of a movie reviews datasets. A promising approach is presented in Prabowo and Thelwall (2009) where a novel methodology has been obtained by the combination of rule based classification, supervised learning and machine learning. In Shein (2009a) a SVM based technique has been introduced for classifying the sentiment in a collection of documents. Other approaches are inferring the sentiment orientation of social media content and estimate sentiment orientations of a collection of documents as a text classification problem (Colbaugh & Glass, 2010). More in general, sentiment related information can be encoded within the actual words of the sentence through changes in attitudinal shades of word meaning using suffixes as discussed in Esuli and Sebastiani (2006). This has been investigated in Neviarouskaya, Prendinger, and Ishizuka (2011) where a lexicon for sentiment analysis has been obtained. In Yu, Liu, Huang, and An (2012) a probabilistic approach to sentiment mining is adopted. In particular this paper uses a probabilistic model called Sentiment Probabilistic Latent Semantic Analysis (S-PLSA) in which a review, and more in general a document, can be considered as generated under the influence of a number of hidden sentiment factors (Hofmann, 1999). The S-PLSA is an extension of the PLSA where it is assumed that there are a set of hidden semantic factors or aspects in the documents related to documents and words under a probabilistic framework. In this paper, we investigate the adoption of a similar approach based on the Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003). In LDA, each document may be viewed as composed by a mixture of various topics. This is similar to probabilistic latent semantic analysis, except that in LDA the topic distribution is assumed to have a Dirichlet prior. By the use of the LDA approach on a set of documents belonging to a same knowledge domain, a Mixed Graph of Terms can be automatically extracted (Clarizia, Colace, De Santo, Greco, & Napoletano, 2011; Clarizia, Colace, Greco, De Santo, & Napoletano, 2011; Colace, De Santo, Greco, & Napoletano, 2014; Napoletano, Colace, De Santo, & Greco, 2012). Such a graph contains a set of weighted word pairs (Colace, De Santo, Greco, & Napoletano, 2015), which we demonstrate to be discriminative for sentiment classification.

The main reason of such discriminative power is that LDA-based topic modeling is essentially an effective conceptual clustering process and it helps discover semantically rich concepts describing the respective "sentimental" relationships. By means of applying these semantically rich concepts, that contain more useful relationship indicators to identify the sentiment from messages and by using a terminological Ontology Builder which allows to identify the kind of semantic relationship between word pairs in mGT, the proposed system can accurately discover more latent relationships and make less errors in its predictions.

The rationale of this paper is the following: Section 2 discusses the extraction of a Mixed Graphs of Terms from a document corpus; Section 3 introduces the proposed approach for the sentiment extraction. The Section 4 discusses the experimental results. Finally, conclusions and further works are discussed.

## 2. Extracting a Mixed Graph of Terms

In this paper we explain how a complex structure, that we call a Mixed Graph of Terms (mGT), allows to capture and represent the information contained in a set of documents that belong to a well-defined knowledge domain. Such a graph can be automatically extracted from a document corpus and can be effectively used as a filter to employ in document classification as well as in sentiment extraction problems. Formally, a Mixed Graph of Terms can be defined as a graph $\mathbf{g} = \langle N, E \rangle$ where:

- $N = \{R, W\}$ is a finite set of nodes, covered by the set $R = \{r_1, \ldots, r_H\}$ whose elements are the *aggregate roots* and by the set $W = \{w_1, \ldots, w_M\}$ containing the *aggregates*. Aggregate roots can be defined as the words whose occurrence is most implied from the occurrence of all other words in the training corpus. Aggregates are defined as the words most related to aggregate roots from a probabilistic point of view.
- $E = \{E_{RR}, E_{RW}\}$ is a set of edges, covered by the set $E_{RR} = \{e_{r_1 r_2}, \ldots, e_{r_{H-1} r_H}\}$ whose elements are links between aggregate roots and by the set $E_{RW} = \{e_{r_1 w_1}, \ldots, e_{r_H w_M}\}$ whose elements are links between aggregate roots and aggregates.
  As better explained further, two aggregate roots are linked if strongly correlated (in a probabilistic sense):

$$e_{r_i r_j} = \begin{cases} 1 & \text{if } \psi_{ij} \geqslant \tau \\ 0 & \text{otherwise} \end{cases}$$

  Aggregate roots can be also linked to aggregates if a relevant probabilistic correlation is present:

$$e_{r_i w_s} = \begin{cases} 1 & \text{if } \rho_{is} \geqslant \mu_i \\ 0 & \text{otherwise} \end{cases}$$

Details about mGT building and thresholds $\tau$ and $\mu_i$ will be now discussed. The Feature Extraction module (FE) is represented in Fig. 1. The input of the system is the set of documents:

$$\Omega_r = (\mathbf{d}_1, \ldots, \mathbf{d}_M)$$

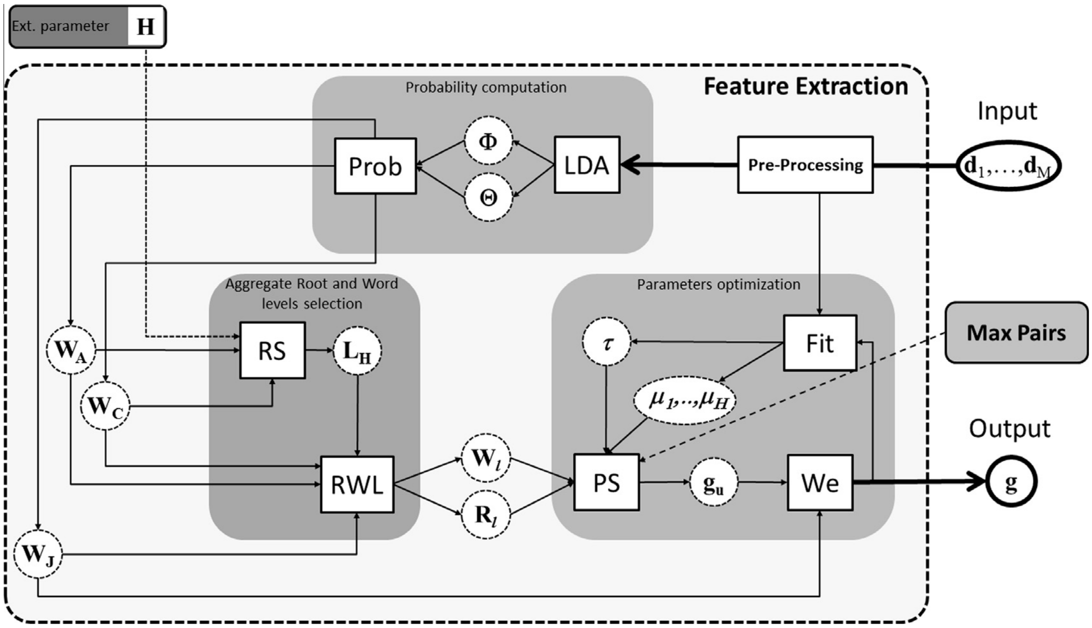


**Fig. 1.** Proposed feature extraction method. A *Mixed Graph of Terms* **g** structure is extracted from a corpus of training documents.

After the pre-processing phase, which involves tokenization, stopwords filtering and stemming, a Term-Document Matrix is built to feed the Latent Dirichlet Allocation (LDA) (Blei et al., 2003) module. The LDA algorithm, assuming that each document is a mixture of a small number of latent topics and each word's creation is attributable to one of the document's topics, provides as output two matrices – $\Theta$ and $\Phi$ – which express probabilistic relations between topic-document and word-topic respectively. Under particular assumptions (Colace, De Santo, & Greco, 2013), LDA module's results can be used to determine: the probability for each word $v_i$ to occur in the corpus $W_A = \{P(v_i)\}$; the conditional probability between word pairs $W_C = \{P(v_i|v_s)\}$; the joint probability between word pairs $W_J = \{P(v_i, v_j)\}$. Details on LDA and probability computation are discussed in Blei et al. (2003), Colace et al. (2014), Colace et al. (2013).

Defining *Aggregate roots* (AR) as the words whose occurrence is most implied by the occurrence of other words of the corpus, a set of $H$ aggregate roots $\mathbf{r} = (r_1, \ldots, r_H)$ can be determined from $W_C$:

$$r_i = \mathrm{argmax}_{v_i} \prod_{j \neg = i} P(v_i|v_j) \tag{1}$$

This phase is referred as Root Selection (RS) in Fig. 1. A weight $\psi_{ij}$ can be defined as a degree of probabilistic correlation between AR pairs: $\psi_{ij} = P(r_i, r_j)$. We define an *aggregate* as a word $v_s$ having a high probabilistic dependency with an aggregate root $r_i$. Such a dependency can be expressed through the probabilistic weight $\rho_{is} = P(r_i|v_s)$. Therefore, for each aggregate root, a set of aggregates can be selected according to higher $\rho_{is}$ values. As a result of the Root-Word level selection (RWL), an initial Mixed Graph of Terms structure, composed by $H$ aggregate roots ($R_l$) linked to all possible aggregates ($W_l$), is obtained. An optimization phase allows to neglect weakly related pairs according to a fitness function discussed in Colace et al. (2013). Our algorithm, given the number of aggregate roots $H$ and the desired max number of pairs as constraints, chooses the best parameter settings $\tau$ and $\mu = (\mu_1, \ldots, \mu_H)$ defined as follows:

1. $\tau$: the threshold that establishes the number of *aggregate root/aggregate root* pairs. A relationship between the aggregate root $v_i$ and aggregate root $r_j$ is relevant if $\psi_{ij} \geqslant \tau$.

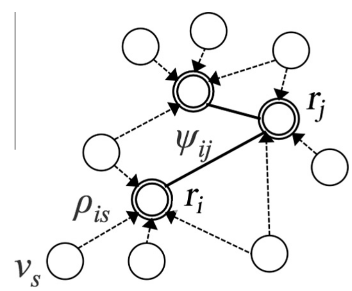


**Fig. 2.** Graphical representation of a *Mixed Graph of Terms* structure.

2. $\mu_i$: the threshold that establishes, for each aggregate root $i$, the number of *aggregate root/word* pairs. A relationship between the word $v_s$ and the aggregate root $r_i$ is relevant if $\rho_{is} \geqslant \mu_i$.

Note that a Mixed Graph of Terms structure which can be suitably represented as a *graph* **g** of terms (Fig. 2). Such a graph is made of several clusters, each containing a set of words $v_s$ (*aggregates*) related to an *aggregate root* ($r_i$), the centroid of the cluster. *Aggregate roots* can be also linked together building a centroids subgraph.

## 3. Searching the sentiment by the use of the Mixed Graph of Terms

The proposed method adopts the Mixed Graph of Terms for building a sentiment detector able to label a document according its sentiment. We propose an architecture composed by the following modules:

- *Mixed Graph of Terms building module*: this module builds a Mixed Graph of Terms starting from a set of documents belonging to a well-defined knowledge domain and previously labeled according the sentiment expressed in them. In this way the obtained Mixed Graph of Terms contains information about the words and their co-occurrences so representing a certain sentiment in a well-defined knowledge domain. Thanks to the LDA approach such a graph can be obtained by the use of a set of few documents. The output of this module is a Mixed Graph of Terms representing the documents and their sentiment. By feeding this module with positive or negative training sets, it will be possible to build mixed graphs of terms for documents that express positive or negative sentiment in a well-defined domain.

- *Sentiment Mining Module*: this module extracts the sentiment from a document thanks to the use of the Mixed Graph of Term as a sentiment filter. The input of this module is a generic document, the Mixed Graph of Terms representing positive and negative sentiment in a knowledge domain and the output is the sentiment detected in the input document.

The sentiment extraction is obtained by a comparison between document and the Mixed Graph of Terms according to the following algorithm:

- *Input of the algorithm*:
  – A set of comments, reviews about items or social posts.
  – The sentiment oriented mixed graphs of terms $mGT^{+}$ and $mGT^{-}$ obtained analyzing the (positively and negatively) training comments.
  – An annotated lexicon *L*.
- *Output of the algorithm*:
  – The average probabilities $P^{+}$ and $P^{-}$ which express the probability that a sentiment, extracted from the set of comments or posts, is "positive" or "negative".
- *Description of the main steps*:
  1. For each word in the $mGT^{+}$ and the $mGT^{-}$ their synonymous are retrieved through the annotated lexicon *L*. In this case, *Wordnet*[1] been selected as lexicon.
  2. For each comment the probabilities $P^{+}_{f_i}$ and $P^{-}_{f_i}$ are determined as:

$$\frac{P^{+/-}_{f_i} = (A + B + C + D)}{4}$$

  *A* being the ratio between the sum of occurrences in the comment of words that are Aggregate Root Nodes and the total number of the Aggregate Root Nodes in the (positive/negative) mGT; *B* the ratio between the sum of occurrences in the document of words that are Aggregates Nodes and the total number of the Aggregates Nodes in the (positive/negative) mGT; *C* the ratio between the sum of the co-occurence probabilities of Aggregate Root Nodes pairs that are in the document and the sum of all the co-occurence probabilities of Aggregate Root Nodes pairs in the (positive/negative) mGT; *D* the ratio between the sum of the co-occurence probabilities of Aggregate Nodes pairs that are in the document and the sum of all the co-occurence probabilities of Aggregate Nodes pairs that are in the (positive/negative) mGT;
  3. For each item the probabilities $P^{+}$ and $P^{-}$ are determined as:

$$P^{+} = \sum_{i} \frac{P^{+}_{f_i}}{num\_of\_comments}$$

$$P^{-} = \sum_{i} \frac{P^{-}_{f_i}}{num\_of\_comments}$$

The proposed approach is effective in an asynchronous sentiment classification, but can work also in a synchronous way. In Fig. 3 the synchronous sentiment real time classificatory architecture is depicted. For real time working two new modules have been introduced:

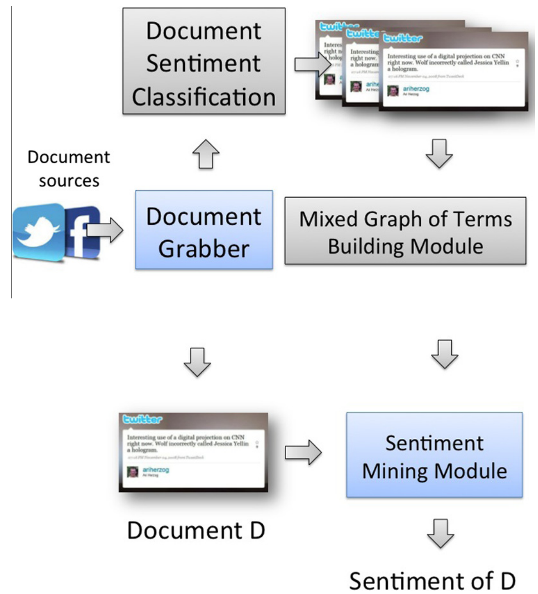


**Fig. 3.** System architecture for synchronous classification.

- *Document Grabber*. This module aims to collect documents from web sources (social networks, blogs and so on). These documents can be collected both for updating the training set and for their classification according to the sentiment. The training set update is an important feature of the proposed approach. In this way, in fact, the various mGTs can be continuously updated and improve their discriminating power introducing new words and relations and deleting inconsistent ones.
- *Document Sentiment Classification*. The new documents inserted into the training set have to be classified by the support of an expert. The aim of this module is to provide a user friendly environment for the classification, according to their sentiment, of the retrieved documents.

## 4. Experimental results

In order to evaluate the performance of the proposed approach, three experimental campaign have been conducted. The first one has been carried out using a standard dataset: the *Movie Reviews Dataset* provided by Pang et al. (2002). This dataset consist of 1000 positive and 1000 negative reviews from the Internet Movie Database. Positive labels were assigned to reviews that had a rating above 3.5 stars and negative labels were assigned to the rest. The first step of the experimental campaign was aimed to find the best size for the training set. For achieving this task nine training sets have been built selecting in a random way from the 10% to 90% of the positive and negative comments that are in the full dataset. By the use of these training sets, the positive and negative mixed Graphs of Terms have been built and the sentiment classification on the remaining comments has been conducted. The process of training sets and Mixed Graph of Terms building and documents classification has been conducted ten times. The obtained results, in terms of average accuracy, are depicted in Fig. 4.

As depicted in the figure the value of accuracy improves with the increase of the training set but the change is very low after the adoption of a training set composed from the 50% of comments contained in the dataset. After this phase a comparison with the results obtained on the same dataset by other approaches in literature has been made (Table 1).

The proposed approach shows the best results in comparison with the other ones when the 50% of dataset is used as training

[1] http://wordnetweb.princeton.edu/perl/webwn.

Please cite this article in press as: Colace, F., et al. Sentiment detection in social networks and in collaborative learning environments. *Computers in Human Behavior* (2014), http://dx.doi.org/10.1016/j.chb.2014.11.090

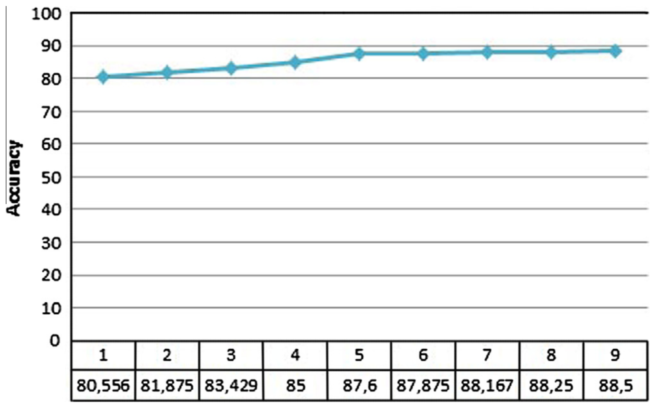


**Fig. 4.** The variation of the accuracy compared to the size of the training set (on the *x*-axis 1 means 10% of training set and so on).

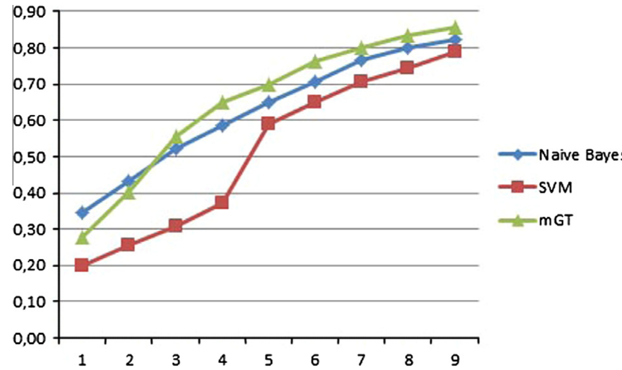


**Fig. 5.** PD_Tweets: the variation of the accuracy compared to the size of the training set (on the *x*-axis 1 means 10% of training set and so on).

**Table 1**
The accuracy obtained by the various methods on the standard dataset.

| Reference paper | Methodology | Accuracy (%) |
|---|---|---|
| Pang et al. (2002) | Support vector machines | 82.90 |
| | Naïve Bayes | 81.50 |
| | Maximum Entropy | 81.00 |
| Kennedy and Inkpen (2006) | Support vector machines | 86.20 |
| Chaovalit and Zhou (2005) | Ontology supported polarity mining | 72.20 |
| Melville et al. (2009) | Bayesian classification with the support of lexicons | 81.42 |
| Shein (2009b) | Formal concept analysis | 77.75 |
| mGT Approach | LDA | 88.50 |

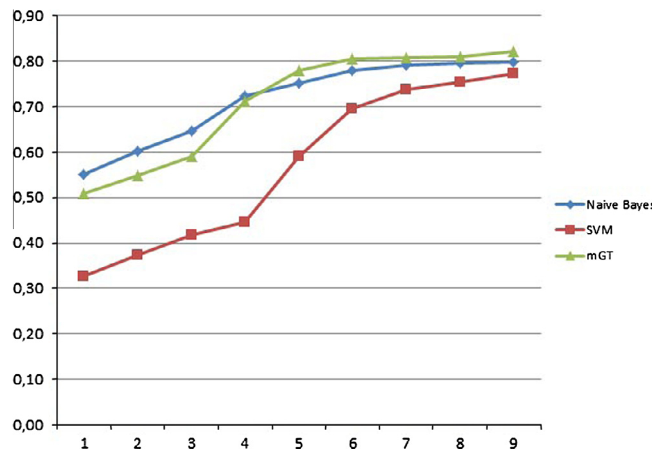


**Fig. 6.** Mobile_Phone_Producer: the variation of the accuracy compared to the size of the training set (on the *x*-axis 1 means 10% of training set and so on).

**Table 2**
Considered datasets.

| Dataset | Source | Positive | Negative | Neutral |
|---|---|---|---|---|
| Mobile phone producer | Facebook | 864 | 759 | 877 |
| PD_Tweets | Twitter | 4783 | 3498 | 1719 |

**Table 3**
The obtained results.

| Perc. (%) | PD_Tweets_Dataset | | | Mobile_Phone_Producer Dataset | | |
|---|---|---|---|---|---|---|
| | Naïve Bayes | SVM | mGT | Naïve Bayes | SVM | mGT |
| 10 | 0.55 | 0.33 | 0.51 | 0.35 | 0.2 | 0.28 |
| 20 | 0.6 | 0.37 | 0.55 | 0.43 | 0.25 | 0.4 |
| 30 | 0.65 | 0.42 | 0.59 | 0.52 | 0.31 | 0.56 |
| 40 | 0.72 | 0.45 | 0.71 | 0.58 | 0.37 | 0.65 |
| 50 | 0.75 | 0.59 | 0.78 | 0.65 | 0.59 | 0.7 |
| 60 | 0.78 | 0.7 | 0.81 | 0.71 | 0.65 | 0.76 |
| 70 | 0.79 | 0.74 | 0.81 | 0.77 | 0.7 | 0.8 |
| 80 | 0.8 | 0.75 | 0.81 | 0.8 | 0.74 | 0.83 |
| 90 | 0.8 | 0.77 | 0.82 | 0.82 | 0.79 | 0.85 |

set. The other approaches usually adopt a larger dataset and in real cases the training phase could be critical and time consuming. The proposed approach also shows good execution time. The classification phase takes about a couple of minutes using a Linux Ubuntu platform running on a 8 GB RAM single CPU, while the training phase takes among five minute to ten minutes depending on the size of the training set.

As previously said our approach gives effective results also in real time scenarios. To demonstrate this aspect, an experimental campaign on posts coming from social networks has been conducted. In particular, 2500 posts from Facebook and 10,000 tweets from twitter have been collected. Facebook's posts have been collected from an official page of a well-known mobile phone producer while the tweets have been collected from the *hashtags* related to two Italian politics, Bersani and Renzi, during the campaign for the primaries of the Italian "Partito Democratico". A group of five experts labeled the posts and the tweets according their sentiments using a majority vote rule and deleting the neutral comments (Table 2).

The methodologies based on the Naïve Bayes and Support Vector Machine introduced in the paper (Kennedy & Inkpen, 2006; Pang et al., 2002) have been implemented and applied to the collected data in order to compare results with the ones obtained by the proposed method. The first step was aimed at building the training set. As previously said, for achieving this task nine training sets have been built selecting in a random way from the 10% to 90% of the positive and negative comments contained in the full dataset. At the end of the training phase, the selected approaches have been tested obtaining the results depicted in Fig. 5 (dataset PD_Tweets), Fig. 6 (Mobile_Phone_Producer) and Table 3.

In the case of the dataset coming from twitter, the experimental results show how the proposed method offers the best performance starting from a training set composed by the 30% of the collected dataset. In general, performances start to be interesting with a training set composed by the 70% of the dataset: it is an expected

**Table 4**
Average sentiment of the classroom.

| Topic | Comments | Positive | | Negative | | Neutral | | #Days |
|---|---|---|---|---|---|---|---|---|
| | | Start | End | Start | End | Start | End | |
| XML Language and DTD | 1032 | 0.30 | 0.54 | 0.53 | 0.15 | 0.17 | 0.31 | 15 |
| XHTML/CSS Language | 803 | 0.29 | 0.52 | 0.43 | 0.21 | 0.28 | 0.26 | 15 |
| Javascript | 714 | 0.32 | 0.54 | 0.46 | 0.26 | 0.21 | 0.19 | 10 |
| PHP Language and Ajax | 1264 | 0.15 | 0.63 | 0.41 | 0.23 | 0.45 | 0.14 | 25 |
| JQuery and Bootstrap | 732 | 0.25 | 0.57 | 0.53 | 0.18 | 0.22 | 0.25 | 10 |
| Network security | 1204 | 0.16 | 0.66 | 0.32 | 0.14 | 0.52 | 0.20 | 15 |

result because tweets are typically composed by a short number of words and so systems based on Naïve Bayes and mGT have to learn from a great number of examples. When compared to SVM, our approach shows acceptable performances if employs at least the 50% of the dataset. The poor results of the SVM, compared to other methods, are due to the difficult to find a well-defined pattern for the correct classification and this task is very difficult in the case of the tweets (where there is not a well-defined structure). In the case of the dataset collected from Facebook, the same approach has been adopted for the classification of the posts. Also in this case the proposed approach shows the best results starting from a training set composed by the 30% of the dataset. In this case the performance of our system improved faster than the twitter case because Facebook's posts contain a greater number of words so that the built mGTs are more effective.

For evaluating the performance of the proposed method in the context of collaborative learning another experimental campaign has been conducted. The experimental scenario involved the analysis of posts collected from the popular e-learning platform Moodle of Software Technology for the Web 2014. In particular, the course has been held by the use of a blended approach and has been organized in the following topics:

- XML Language and DTD.
- XHTML/CSS Language.
- Javascript.
- PHP Language and Ajax.
- JQuery and Bootstrap.
- Network security.

For each topic a final test has been submitted to the students. The traditional lectures have been supported by the use of additional learning contents distributed by the use of Moodle. Chat and forum enhanced the collaborative approach of the course. About 120 students attended the lectures and used Moodle for share comments each other. The contents exchanged by the use of forum and chat have been set not visible for the teacher and this policy was known by students. A Sentiment Analysis Module has been used for grabbing the mood of students during the various lectures related to the various topics. The real time analysis of the comments provided a mood-meter of classroom regarding the various topics. In Table 4 the number of the posts collected from the chat and the forum for each topic has been reported. The average sentiment has also been reported in terms of positive, negative and neutral mood at the beginning (only the first day considered) and at the end of the observation period. The observation period (shown as the number of days for each course) expresses the length of the course part dedicated to a certain topic.

It is interesting to notice the positive trend of the positive sentiment during the observation time (Fig. 7). The reason of this trend is almost clear: at the beginning of each topic students showed a natural disorientation that is greater for topics related to the programming sections of the course. After these first phases teacher updated his teaching style according to the sentiment of

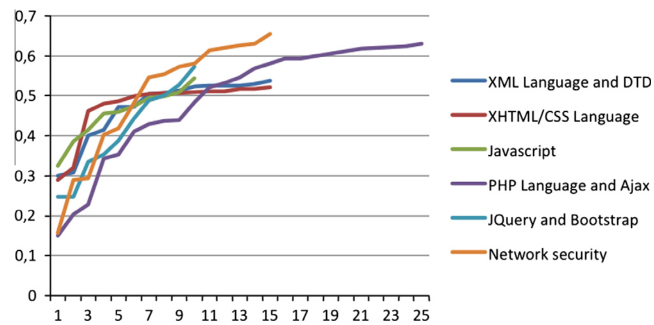


**Fig. 7.** Trend of the positive sentiment during the observation period.

the students giving them more contents or introducing more examples or exercises. In general, teacher appreciated the Sentiment Grabber tool above all for the opportunity to manage the mood of the class without having to consider explicitly the relationship teacher-student.

## 5. Conclusion

This paper proposes the use of the Mixed Graph of Terms (MGTs) structures for the sentiment classification of textual documents. Once reference mGTs have been learnt from training documents having a given sentiment orientation, the classification of a new document can be performed by using such reference mGTs as sentiment filters. Such a method has the strength to be language independent, relying mainly on a probabilistic technique. Our approach was compared to the state-of-art methods in literature using standard and real datasets. In particular, we had a chance to test our method in a social collaborative learning platform used for a university course. Results are encouraging and show that the proposed method can contribute in a interesting way to the research in the field of modern collaborative learning platforms. As future works, we plan to improve the method by expanding the mGT with additional synonyms from annotated lexicon, as WordNet and SentiWordnet. This could be also useful for a better evaluation of the words and the sentence structures.

## References

Baroni, M., & Vegnaduzzo, S. (2004). Identifying subjective adjectives through web-based mutual information. In *Proceedings of the 7th Konferenz zur Verarbeitung Natürlicher Sprache (German conference on natural language processing – KONVENS'04)* (pp. 613–619).

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *The Journal of Machine Learning Research, 3*, 993–1022.

Chaovalit, P., & Zhou, L. (2005). Movie review mining: a comparison between supervised and unsupervised classification approaches. In *Proceedings of the proceedings of the 38th annual Hawaii international conference on system sciences (HICSS'05) – Track 4, HICSS '05* (Vol. 04, pp. 112.3). Washington, DC, USA: IEEE Computer Society.

Clarizia, F., Colace, F., De Santo, M., Greco, L., & Napoletano, P. (2011). Mixed graph of terms for query expansion. In *11th International conference on intelligent systems design and applications (ISDA)* (pp. 581–586).

Clarizia, F., Colace, F., Greco, L., De Santo, M., & Napoletano, P. (2011). Improving text retrieval accuracy using a graph of terms. In *DMS* (pp. 42–47). Knowledge Systems Institute.

Colace, F., De Santo, M., & Greco, L. (2013). Weighted word pairs for text retrieval. In *Proceedings of the 3nd Italian information retrieval (IIR), CEUR workshop proceedings* (Vol. 964).

Colace, F., De Santo, M., Greco, L., & Napoletano, P. (2014). Text classification using a few labeled examples. *Computers in Human Behavior, 30*, 689–697.

Colace, F., De Santo, M., Greco, L., & Napoletano, P. (2015). Weighted word pairs for query expansion. *Information Processing & Management, 51*(1), 179–193.

Colbaugh, R., & Glass, K. (2010). Estimating sentiment orientation in social media for intelligence monitoring and analysis. In *IEEE international conference on intelligence and security informatics (ISI), may 2010* (pp. 135–137).

Dascalu, M.-I., Bodea, C.-N., Lytras, M., de Pablos, P. O., & Burlacu, A. (2014). Improving e-learning communities through optimal composition of multidisciplinary learning groups. *Computers in Human Behavior, 30*(0), 362–371.

Esuli, A., & Sebastiani, F. (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th conference on language resources and evaluation (LREC'06)* (pp. 417–422).

Gamon, M., & Aue, A. (2005). Automatic identification of sentiment vocabulary: Exploiting low association with known sentiment terms. In *Proceedings of the ACL workshop on feature engineering for machine learning in natural language processing, June 2005* (pp. 57–64). Ann Arbor, Michigan: Association for Computational Linguistics.

Hofmann, T. (1999). Probabilistic latent semantic analysis. In *Proc. of uncertainty in artificial intelligence, UAI'99* (pp. 289–296).

Kennedy, A., & Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence, 22*, 2006.

Liu, B. (2010). Sentiment analysis and subjectivity. In *Handbook of natural language processing*. Boca: Taylor and Francis Group.

Lytras, M. D., & de Pablos, P. O. (2008). The role of a "make" or internal human resource management system in spanish manufacturing companies: Empirical evidence. *Human Factors and Ergonomics in Manufacturing & Service Industries, 18*(4), 464–479.

Lytras, M., & de Pablos, P. O. (2011). Software technologies in knowledge society. *Journal of Universal Computer Science, 17*(9), 1219–1221.

Lytras, M., Sicilia, M. A., Naeve, A., de Pablos, P. O., & Lytras, M. D. (2008). Competencies and human resource management: Implications for organizational competitive advantage. *Journal of Knowledge Management, 12*(6), 48–55.

Melville, P., Gryc, W., & Lawrence, R. D. (2009). Sentiment analysis of blogs by combining lexical knowledge with text classification. In *Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining, KDD '09* (pp. 1275–1284). New York, NY, USA: ACM.

Napoletano, P., Colace, F., De Santo, M., & Greco, L. (2012). Text classification using a graph of terms. In *Sixth international conference on complex, intelligent and software intensive systems (CISIS), July* (pp. 1030–1035).

Neviarouskaya, A., Prendinger, H., & Ishizuka, M. (2011). Sentiful: A lexicon for sentiment analysis. *IEEE Transactions on Affective Computing, 2*(1), 22–36.

Neviarouskaya, A., Prendinger, H., & Ishizuka, M. (2011). Sentiful: A lexicon for sentiment analysis. *IEEE Transactions on Affective Computing, 2*(1), 22–36.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of EMNLP* (pp. 79–86).

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1-2), 1–135.

Prabowo, R., & Thelwall, M. (2009). Sentiment analysis: A combined approach. *Journal of Informetrics, 3*, 143–157.

Sebastiani, F. (2002). Machine learning in text categorization. *ACM Computing Surveys, 34*, 1–47.

Shein, K.P.P. (2009a). Ontology based combined approach for sentiment classification. In *Proceedings of the 3rd international conference on communications and information technology, CIT'09* (pp. 112–115). Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS).

Shein, K.P.P. (2009b). Ontology based combined approach for sentiment classification. In *Proceedings of the 3rd international conference on communications and information technology, CIT'09* (pp. 112–115). Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS).

Turney, P.D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics, ACL '02* (pp. 417–424). Stroudsburg, PA, USA: Association for Computational Linguistics.

Turney, P., & Littman, M. (2002). *Unsupervised learning of semantic orientation from a hundred-billion-word corpus. Technical report NRC technical report ERB-1094*. Institute for Information Technology, National Research Council Canada.

Wang, C., Xiao, Z., Liu, Y., Xu, Y., Zhou, A., & Zhang, K. (2013). Sentiview: Sentiment analysis and visualization for internet popular topics. *IEEE Transactions on Human-Machine Systems, 43*(6), 620–630.

Wilson, T., Wiebe, J., & Hwa, R. (2004). Just how mad are you? finding strong and weak opinion clauses. In *Proceedings of the 19th national conference on artificial intelligence, AAAI'04* (pp. 761–767). AAAI Press.

Yu, X., Liu, Y., Huang, X., & An, A. (2012). Mining online reviews for predicting sales performance: A case study in the movie domain. *IEEE Transactions on Knowledge and Data Engineering, 24*(4), 720–734.

Please cite this article in press as: Colace, F., et al. Sentiment detection in social networks and in collaborative learning environments. *Computers in Human Behavior* (2014), http://dx.doi.org/10.1016/j.chb.2014.11.090