

# Flat and Hierarchical Classifiers for Detecting Emotion in Tweets

Giulio Angiani, Stefano Cagnoni, Natalia Chuzhikova, Paolo Fornacciari,  
Monica Mordonini, and Michele Tomaiuolo<sup>(✉)</sup>

Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Parma,  
Parco Area delle Scienze 181/A, 43124 Parma, Italy  
{angiani,cagnoni,fornacciari,monica,tomamic}@ce.unipr.it,  
michele.tomaiuolo@unipr.it

**Abstract.** Social media are more and more frequently used by people to express their feelings in the form of short messages. This has raised interest in emotion detection, with a wide range of applications among which the assessment of users' moods in a community is perhaps the most relevant. This paper proposes a comparison between two approaches to emotion classification in tweets, taking into account six basic emotions. Additionally, it proposes a completely automated way of creating a reliable training set, usually a tedious task performed manually by humans. In this work, training datasets have been first collected from the web and then automatically filtered to exclude ambiguous cases, using an iterative procedure. Test datasets have been similarly collected from the web, but annotated manually. Two approaches have then been compared. The first is based on a direct application of a single “flat” seven-output classifier, which directly assigns one of the emotions to the input tweet, or classifies it as “objective”, when it appears not to express any emotion. The other approach is based on a three-level hierarchy of four specialized classifiers, which reflect a priori relationships between the target emotions. In the first level, a binary classifier isolates subjective (expressing emotions) from objective tweets. In the second, another binary classifier labels subjective tweets as positive or negative. Finally, in the third, one ternary classifier labels positive tweets as expressing joy, love, or surprise, while another classifies negative tweets as expressing anger, fear, or sadness. Our tests show that the a priori domain knowledge embedded into the hierarchical classifier makes it significantly more accurate than the flat classifier.

**Keywords:** Social media · Emotions detection · Hierarchical classification

## 1 Introduction

In recent years, people's interest in public opinion has dramatically increased.

Opinions have become key influencers of our behavior; because of this, people do not just ask friends or acquaintances for advice when they need to take

a decision, for example, about buying a new smartphone. They rather rely on the many reviews from other people they can find on the Internet. This does not apply only to individuals but also to companies and large corporations. Sentiment analysis and opinion mining have spread from computer science to management and social sciences, due to their importance to business and society as a whole. These studies are possible thanks to the availability of a huge amount of text documents and messages expressing users' opinions on a particular issue. All this valuable information is scattered over the Web throughout social networks such as Twitter and Facebook, as well as over forums, reviews and blogs.

Automatic classification is a generic task, which can be adapted to various kinds of media [1,2]. Our research aims, firstly, at recognizing the emotions expressed in the texts classified as subjective, going beyond our previous approaches to the analysis of Twitter posts (tweets), where a simple hierarchy of classifiers, mimicking the hierarchy of sentiment classes, discriminated between objectivity and subjectivity and, in this latter case, determined the polarity (positive/negative attitude) of a tweet [3]. The general goal of our study is to use sentiment detection to understand users' moods and, possibly, discover potential correlations between their moods and the structure of the communities to which they belong within a social network [4]. In view of this, polarity is not enough to capture the opinion dynamics of a social network: we need more nuances of the emotions expressed therein. To do this, emotion detection is performed based on Parrott's tree-structured model of emotions [5]. Parrott identifies six primary emotions (anger, fear, sadness, joy, love, and surprise); then, for each of them, a set of secondary emotions and, finally, in the last level, a set of tertiary emotions. For instance, a (partial) hierarchy for *Sadness* could include secondary emotions like *Suffering* and *Disappointment*, as well as tertiary emotions like *Agony*, *Anguish*, *Hurt* for *Suffering* and *Dismay*, *Displeasure* for *Disappointment*.

Secondly, our research aims to significantly increase the size of the training sets for our classifiers using a data collection method which does not require any manual tagging of data, limiting the need for reference manually-tagged data only to the collection of an appropriate test set.

Accordingly, in the work described in this paper, we have first collected large training datasets from Twitter using a robust automatic labeling procedure, and a smaller test set, manually tagged, to guarantee the highest possible reliability of data used to assess the performance of the classifiers derived from those training sets. Then, we have compared the results obtained by a single "flat" seven-output classifier to those obtained by a three-level hierarchy of four classifiers, that reflects apriori knowledge on the domain. In the first level of our hierarchy, a binary classifier isolates subjective from objective tweets; in the second, another binary classifier labels subjective tweets as positive or negative and, in the third, one ternary classifier labels positive tweets as expressing joy, love, or surprise, while another classifies negative tweets as expressing anger, fear, or sadness.

The paper is structured as follows. Section 2 offers a brief overview of related work. Section 3 describes the methodology used in this work. Section 4 describes the experimental setup and summarizes and compares the results obtained by the flat and the hierarchical classifiers. Section 5, concludes the paper discussing and analyzing the results achieved.

## 2 Related Work

Although widely studied, sentiment analysis still offers several challenges to computer scientists. Recent and comprehensive surveys of sentiment analysis and of the main related data analysis techniques can be found in [6, 7]. Emotional states like joy, fear, anger, and surprise are encountered in everyday life and social media are more and more frequently used to express one's feelings. Thus, one of the main and most frequently tackled challenges is the study of the mood of a network of users and of its components (see, for example, [8–10]). In particular, emotion detection in social media is becoming increasingly important in business and social life [11–15].

As for the tools used, hierarchical classifiers are widely applied to large and heterogeneous data collections [16–18]. Essentially, the use of a hierarchy tries to decompose a classification problem into sub-problems, each of which is smaller than the original one, to obtain efficient learning and representation [19, 20]. Moreover, a hierarchical approach has the advantage of being modular and customizable, with respect to single classifiers, without any loss of representation power: Mitchell [21] has proved that the same feature sets can be used to represent data in both approaches.

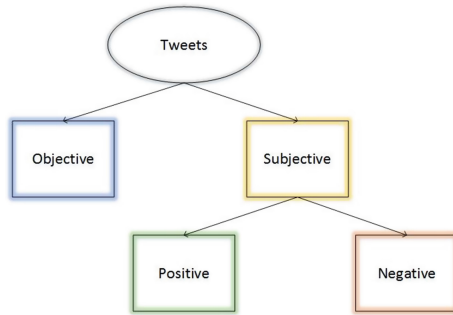
In emotion detection from text, the hierarchical classification considers the existing relationship between subjectivity, polarity and the emotion expressed by a text. In [22], the authors show that a hierarchical classifier performs better on highly imbalanced data from a training set composed of web postings which has been manually annotated. They report an accuracy of 65.5% for a three-level classifier versus 62.2% for a flat one. In our experiments, we performed a similar comparison on short messages coming directly from Twitter channels, without relying on any kind of manual tagging.

Emotions in tweets are detected according to a different approach in [23]. In that work, polarity and emotion are concurrently detected (using, respectively, SentiWordNet [24] and NRC Hashtag Emotion Lexicon [25]). The result is expressed as a combination of the two partial scores and improves the whole accuracy from 37.3% and 39.2% obtained, respectively by independent sentiment analysis and emotion analysis, to 52.6% for the combined approach. However this approach does not embed the a-priori knowledge on the problem as effectively as a hierarchical approach, while limiting the chances to build a modular customizable system.

### 3 Methodology

A common approach to sentiment analysis includes two main classification stages, represented in Fig. 1:

1. Subdivision of texts according to the principles of objectivity/subjectivity. An objective assertion only shows some truth and facts about the world, while a subjective proposition expresses the author’s attitude toward the subject of the discussion.
2. Determination of the polarity of the text. If a text is classified as subjective, it is regarded as expressing feelings of a certain polarity (positive or negative).



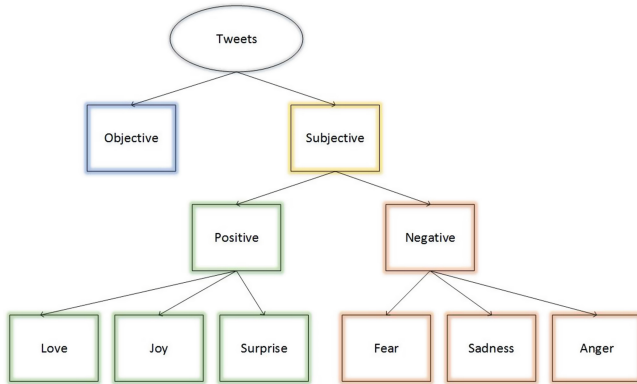
**Fig. 1.** Basic classification.

As mentioned earlier, the purpose of this paper is to improve the existing basic classification of tweets. Within this context, improving the classification should be considered as an extension of the basic model in the direction of specifying the emotions which characterize subjective tweets, based on Parrott’s socio-psychological model. According to it, all human feelings are divided into six major states (three positive and three negative):

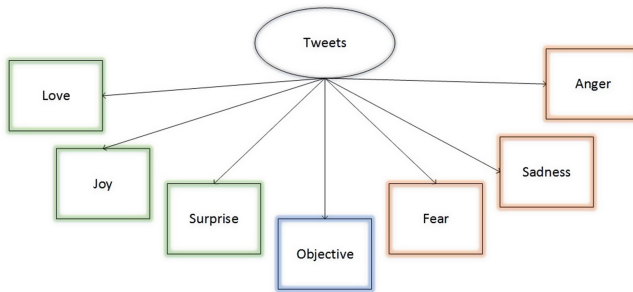
- positive feelings of love, joy, surprise
- negative feelings of fear, sadness, anger

We take into consideration flat and hierarchical classifiers, which are shown in Figs. 2 and 3, respectively. Hierarchical classification is based on the consistent application of multiple classifiers, organized in a tree-like structure. In our case, a first step uses a binary classifier that determines the subjectivity/objectivity of a tweet. The second step further processes all instances that have been identified as subjective. It uses another binary classifier that determines the polarity (positivity/negativity) of a tweet. Depending on the polarity assessed at the previous level, the third step classifies the specific emotion expressed in the text (love, joy or surprise for positive tweets; fear, sadness or anger for negative tweets).

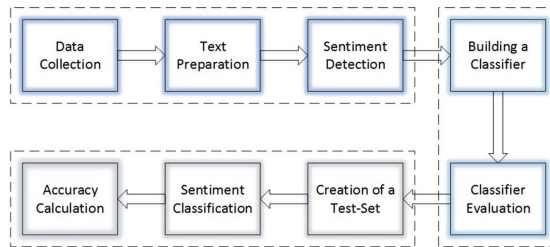
To limit the need for human intervention in the definition of the training data, and thus allow for the collection of larger datasets, we devised a strategy



**Fig. 2.** Hierarchical classification.



**Fig. 3.** Flat classification.



**Fig. 4.** Application structure.

to completely automate the collection of one training set for the construction of the flat multi-classifier, and other four for training the classifiers comprised in our hierarchical model.

We used manual labeling only in building the test sets, since the reliability of such data is absolutely critical for the evaluation of results of the classifiers taken into consideration.

The rest of this section briefly describes the modules we developed to implement our method, as well as the data and the procedure adopted to create the training sets.

The project has been developed using Java within the Eclipse IDE. It is structured into three main modules, as shown in Fig. 4.

### 3.1 Collecting Training Data

The main requirement for constructing an emotion classifier based on a machine learning approach is to download a sufficient amount of posts for the training phase. Tweets must be pre-processed, clearing them from the elements which have no emotional meaning, such as hashtags and user references. It is also important to correct spelling mistakes, and to encode special characters and emoticons appropriately as text tokens. Each sample of the training set represents a tweet and is composed by the processed text and an emotion class used as a label.

### 3.2 Training Sets and Classification

Our classifiers were trained using the “Naive Bayes Multinomial” algorithm provided by Weka. They have been trained using training data collected automatically and systematically into the training sets that contain, in each line, one tweet and the label of the class to which it belongs.

For the feature selection, we first used a Weka filter (*StringToWordVector*) to turn a string into a set of attributes representing word occurrences. After that, an optimal set of attributes (N-grams) was selected using the *Information Gain* algorithm provided by Weka, which estimates the worth of a feature by measuring the information gain with respect to the class.

For the hierarchical classifier, four training sets have been created, with data labeled according to the task of each of the four classifiers: “OBJ” or “SUB” for the objectivity classifier, “POS” or “NEG” for the polarity classifier. For the lower-level emotion classifiers (one for the tweets labeled as positive, one for those labeled as negative by the higher-level classifiers) the following labels/classes have been considered:

- Classes of positive polarity: LOVE, JOY, SURPRISE
- Classes of negative polarity: ANGER, SADNESS, FEAR

For the flat classifier, a multiclass file was created, with each sample labeled according to one of seven classes, six representing the emotions taken into account and the seventh for the objective tweets. Namely: LOVE, JOY, SURPRISE, ANGER, SADNESS, FEAR, OBJ.

The channels used to obtain emotive tweets (according to the corresponding hashtag) involve emotions that Parrott identified as either primary, secondary, or tertiary. Parrot’s taxonomy of basic human feelings, including only the primary and secondary level, is presented in Table 1.

**Table 1.** Primary and secondary emotions of Parrott’s socio-psychological model.

Primary emotion	Secondary emotion
Love	Affection, Lust/Sexual desire, Longing
Joy	Cheerfulness, Zest, Contentment, Pride, Optimism, Relief
Surprise	Surprise
Anger	Irritability, Exasperation, Rage, Disgust, Envy, Torment
Sadness	Suffering, Sadness, Disappointment, Shame, Neglect
Fear	Horror, Nervousness

The training set defining the objectivity or subjectivity of a tweet has been downloaded from the SemEval3 public repository.<sup>1</sup>

### 3.3 Classifying Data

We used the function library provided by Weka to develop a Java application for assessing the quality of our classifiers. The application supports both classification models taken into consideration for processing a test set using the Weka classifier models trained with the data described above, labels data and assesses the classifiers’ accuracy by comparing the labels assigned to the test data by the classifiers to the actual ones, reported in the test set.

## 4 Results

In this section we present the results of our research. We first describe the experimental setup and, in particular, the procedure we followed to collect the data sets for training and testing the classifiers, as well as the preliminary tests we made to evaluate the quality of data and to determine the optimal number of features as well as the size of the N-grams used as features.

Finally, we compare the flat and the hierarchical classifiers on the basis of the accuracy they could achieve on the test set.

### 4.1 Collecting Data

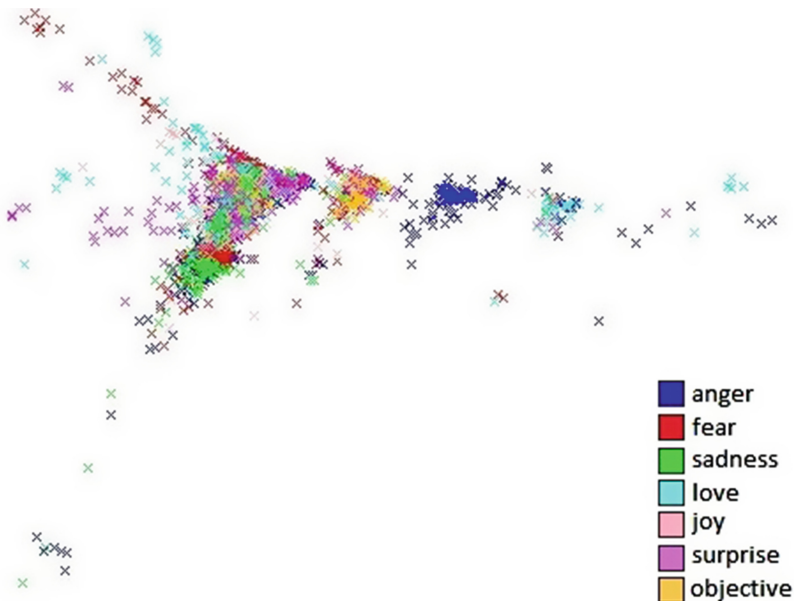
Our training sets were built in a completely automated way, without human intervention.

- **Raw training set (Training Set 1).** Our raw training set (in the following called TS1) consists of about 10,000 tweets: we collected about 1500 tweets for each emotion and as many objective tweets. For the six nuances of emotions, we gathered data coming from several Twitter channels, following Parrott’s classifications. Thus, the selection of channels was made methodically, without

<sup>1</sup> <https://www.cs.york.ac.uk/semEval-2013/task2/index.php?id=data.html>.

human evaluation. For each emotion, we used all the three levels of Parrott’s model: for example, to extract tweets expressing *sadness* we downloaded data from the channel related to the primary emotion, *#Sadness*, but also from those related to secondary (*#Suffering*, *#Disappointment*, *#Shame*, ...) and tertiary emotions (*#Agony*, *#Anguish*, *#Hurt* for *Suffering*; *#Dismay*, *#Displeasure* for *Disappointment*, and so on). The objective (neutral) tweets were selected from the data set used for the SemEval competition<sup>2</sup>.

- **Refined training set (Training Set 2).** Since the raw training set contains tweets obtained directly from Twitter channels, it may certainly contain spurious data. Thus, we adopted an automatic process to select only the most appropriate tweets. We filtered TS1 to remove the most ambiguous cases, and obtained a second training set (in the following called TS2) of about 1000 tweets for each of the six primary emotions. The filtering process was based on six binary classifiers, one for each emotion. The training set for each of them was balanced and considered two classes: the “positive” class included all raw tweets automatically downloaded from sources related to the emotion associated to the classifier; the “negative” class included tweets coming, in equal parts, from the other five emotions and from the set of objective tweets. Finally, TS2 included only the tweets which could be classified correctly by the binary classifier, in order for the tweets we used for training the main classifiers (i.e., those in TS2) to be as prototypical as possible.



**Fig. 5.** Visualization of the first two PCA components for the test set.

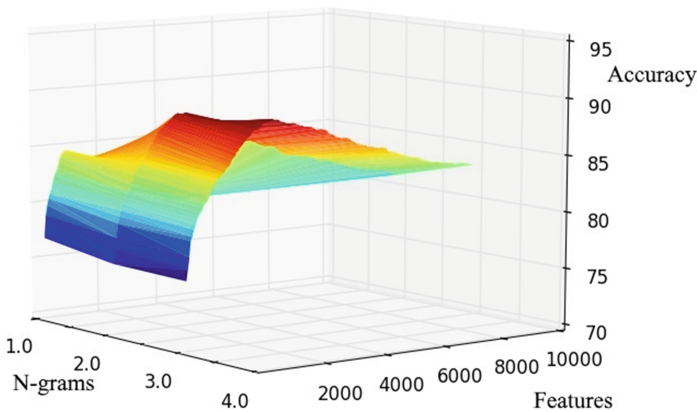
<sup>2</sup> <https://en.wikipedia.org/wiki/SemEval>.



- **Test Set.** Tweets for the test set were downloaded in the same way as those for the training set, but they were manually annotated. They consist of 700 tweets, 100 for each of the six emotions in addition to 100 objective tweets. Even if, obviously, a representation sufficiently relevant for a correct classification would require a much larger number of features, we plot their first two components obtained by Principal Component Analysis (PCA) in Fig. 5 to give a first rough idea of the distribution of the tweets in the feature space. Objective tweets (yellow) and tweets related with sadness (green) are quite clearly separated from the others even in this minimal representation. Instead, other emotions are much closer and significantly overlapped, especially those related with surprise (violet). This could actually be justified considering that secondary and tertiary emotions can play a very significant role in recognizing this emotion, since it can be equally associated with both positive and negative events.

## 4.2 Optimizing the Parameters of Classifiers

For each classifier (four for the hierarchical and one for the flat approach), a systematic preliminary analysis was performed to optimize some relevant parameters that affect the training phase. We selected a grid of configurations and then used cross-validation to estimate the quality of classifiers configured according to it. In particular, we searched for the optimal length of N-grams to be used as features. Figure 6 shows the case of the flat classifier, but the other cases are similar. It can be observed that accuracy nearly peaks at N-gram = 2. Longer sequences increase the complexity of the training phase, without producing any significant improvement of the results. We also analysed the dependence of the performance on the number of features selected, using Weka’s Information Gain algorithm. In Fig. 6 one can observe that its increase does not provide a



**Fig. 6.** Optimization of the system parameters.

**Table 2.** Parameter optimization results.

Classifier	N-Gram (max)	Features
Flat	2	1500
Sub/Obj	2	2500
Pos/Neg	2	2500
Anger/Sadness/Fear	2	1550
Love/Joy/Surprise	2	1500

monotonic improvement of the classifier quality. Instead, it has a peak at around 1500 features.

Table 2 shows the results of the parameter optimization step. In particular, the N-Gram (max) value is 2 for all our classifiers (unigram and bigram are considered). The last column shows the number of features that optimizes the performance of the classifiers.

### 4.3 Accuracy

Tables 3 and 4 report the results obtained on the test set by the seven-output flat classifier and by the hierarchical one. They display the accuracy of each approach, when trained on TS1 or on TS2, in order to assess the effect of the refinement step. These tables confirm the advantage of hierarchical classification, which intrinsically exploits a priori domain knowledge embedded into the whole classifier structure. They also show that some emotions (e.g., sadness) are classified rather well, while others are harder to classify (e.g., anger). For each class the best results in terms of precision, recall and F-measure have been emphasized. These results show that the best results have been obtained using the filtered training set (TS2).

Table 5 reports in detail the partial results of the three classification levels of the hierarchical classifiers: the first and the second level of classification, i.e., subjectivity and polarity, have an accuracy of around 90% and 75%, respectively. Aggregating the results of the flat classifier to provide the same partial responses provides systematically worse results. This is not surprising, since a seven-output classification is a harder task in general, and for ambiguous (and often mixed) emotions in particular. On the other hand, the cascaded structure of a hierarchical classifiers has a higher risk of propagating errors from the higher levels to the lower ones. From this point of view, the results show that the structure we adopted minimizes that effect, since, still not surprisingly, the accuracy of classifiers increases with their level in the hierarchy.

Finally, Table 6 shows the confusion matrix of the hierarchical classifier trained using TS2 which is the best performing approach in of our research. Notably, fear and anger are often misclassified as sadness and love is often misclassified as joy or surprise

**Table 3.** Accuracy of flat classifier, using alternatively the two training sets.

	Flat					
	TS1			TS2		
	Prec	Rec	F_M	Prec	Rec	F_M
Objective	<b>0.70</b>	0.69	0.69	0.58	0.73	0.65
Anger	<b>0.36</b>	0.18	0.24	<b>0.36</b>	0.23	0.28
Fear	0.39	0.22	0.28	<b>0.50</b>	0.13	0.21
Sadness	0.29	<b>0.68</b>	0.41	0.31	0.66	0.42
Love	0.45	0.29	0.35	<b>0.50</b>	0.32	<b>0.39</b>
Joy	0.32	0.34	0.33	0.36	<b>0.42</b>	<b>0.39</b>
Surprise	0.46	0.39	0.42	<b>0.50</b>	0.42	0.46
Total accuracy	40,48 %			42,33 %		

**Table 4.** Accuracy of hierarchical classifier, using alternatively the two training sets.

	Hierarchical					
	TS1			TS2		
	Prec	Rec	F_M	Prec	Rec	F_M
Objective	0.66	0.87	<b>0.75</b>	0.60	<b>0.88</b>	0.71
Anger	0.31	0.27	0.28	0.34	<b>0.30</b>	<b>0.32</b>
Fear	0.40	0.22	0.28	0.46	<b>0.23</b>	<b>0.31</b>
Sadness	0.34	0.56	0.43	<b>0.38</b>	0.60	<b>0.46</b>
Love	0.40	<b>0.34</b>	0.37	0.43	0.28	0.34
Joy	<b>0.39</b>	0.34	0.37	0.39	0.38	0.39
Surprise	0.47	0.41	0.44	0.50	<b>0.45</b>	<b>0.48</b>
Total accuracy	43,61 %			<b>45,17 %</b>		

**Table 5.** Accuracy of the intermediate results of hierarchical classification, based on TS1 and TS2.

	TS1	TS2
Objective/Subjective	91,90 %	90,06 %
Positive/Negative	73,36 %	75,54 %
Final classification	43,61 %	45,17 %

**Table 6.** Confusion matrix of the hierarchical classification based on TS2.

→	Objective	Fear	Anger	Sadness	Love	Joy	Surprise
Objective	<b>88</b>	3	2	3	2	1	1
Fear	15	<b>22</b>	15	23	1	13	5
Anger	15	7	<b>30</b>	38	4	5	1
Sadness	8	1	13	<b>63</b>	1	10	8
Love	5	8	8	11	<b>29</b>	20	21
Joy	8	4	7	16	17	<b>39</b>	10
Surprise	7	2	12	12	12	11	<b>47</b>

## 5 Conclusion

Emotion detection in social media is becoming a task of increasing importance, since it can provide a richer view of a user’s opinion and feelings about a certain topic. It can also pave the way to detecting provocatorial and anti-social behaviors. In this work, we have analyzed the problem of automatic classification of tweets, according to their emotional value. We referred to Parrott’s model of six primary emotions: anger, fear, sadness, joy, love, and surprise. In particular, we compared a flat classifier to a hierarchical one. Our tests show that the domain knowledge embedded into the hierarchical classifier makes it more accurate than the flat classifier. Also, our results prove that the process of automatic construction of training sets is viable, at least for sentiment analysis and emotion classification, since our automatic filtering of training data makes it possible to create training sets that improve the quality of the final classifier with respect to a “blind” collection of raw data based only on the hashtags. The results we have obtained are comparable with those found in similar works [26].

## References

1. Ugolotti, R., Sassi, F., Mordonini, M., Cagnoni, S.: Multi-sensor system for detection and classification of human activities. *J. Ambient Intell. Humaniz. Comput.* **4**, 27–41 (2013)
2. Matrella, G., Parada, G., Mordonini, M., Cagnoni, S.: A video-based fall detector sensor well suited for a data-fusion approach. *Assistive Technol. Adap. Equipment Inclusive Environ. Assistive Technol. Res. Ser.* **25**, 327–331 (2009)
3. Fornacciari, P., Mordonini, M., Tomaiuolo, M.: A case-study for sentiment analysis on Twitter. In: *Workshop dagli Oggetti agli Agenti, WOA 2015* (2015)
4. Fornacciari, P., Mordonini, M., Tomaiuolo, M.: Social network and sentiment analysis on Twitter: towards a combined approach. In: *1st International Workshop on Knowledge Discovery on the WEB, KDWeb 2015* (2015)
5. Parrott, W.G.: *Emotions in Social Psychology: Essential Readings*. Psychology Press, Philadelphia (2001)
6. Liu, B.: *Sentiment Analysis and Opinion Mining: Synthesis Lectures on Human Language Technologies*, vol. 5, pp. 1–167 (2012)

7. Mohammad, S.M.: Sentiment analysis: detecting valence, emotions, and other affectual states from text. *Emotion Measurement* (2015)
8. Mislove, A., Lehmann, S., Ahn, Y.Y., Onnela, J.P., Rosenquist, J.N.: Pulse of the nation: US mood throughout the day inferred from Twitter. Northeastern University (2010)
9. Healey, C., Ramaswamy, S.: Visualizing Twitter sentiment (2010). Accessed 17 June 2016
10. Allisio, L., Mussa, V., Bosco, C., Patti, V., Ruffo, G.: Felicità: Visualizing and estimating happiness in Italian cities from geotagged tweets. In: *ESSEM@ AI\* IA*, pp. 95–106 (2013)
11. Kao, E.C., Liu, C.C., Yang, T.H., Hsieh, C.T., Soo, V.W.: Towards text-based emotion detection a survey and possible improvements. In: *2009 International Conference on Information Management and Engineering, ICIME 2009*, pp. 70–74. IEEE (2009)
12. Strapparava, C., Valitutti, A., et al.: WordNet affect: an affective extension of WordNet. In: *LREC*, vol. 4, pp. 1083–1086 (2004)
13. Strapparava, C., Mihalcea, R.: Learning to identify emotions in text. In: *Proceedings of the 2008 ACM Symposium on Applied Computing*, pp. 1556–1560. ACM (2008)
14. Poria, S., Cambria, E., Winterstein, G., Huang, G.B.: Sentic patterns: dependency-based rules for concept-level sentiment analysis. *Knowl. Based Syst.* **69**, 45–63 (2014)
15. Franchi, E., Poggi, A., Tomaiuolo, M.: Social media for online collaboration in firms and organizations. *Int. J. Inf. Syst. Model. Des. (IJISMD)* **7**, 18–31 (2016)
16. Dumais, S., Chen, H.: Hierarchical classification of web content. In: *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 256–263. ACM (2000)
17. Silla Jr., C.N., Freitas, A.A.: A survey of hierarchical classification across different application domains. *Data Min. Knowl. Disc.* **22**, 31–72 (2011)
18. Addis, A., Armano, G., Vargiu, E.: A progressive filtering approach to hierarchical text categorization. *Commun. SIWN* **5**, 28–32 (2008)
19. Koller, D., Sahami, M.: Hierarchically classifying documents using very few words. In: *Proceedings of the 14th International Conference on Machine Learning* (1997)
20. Armano, G., Mascia, F., Vargiu, E.: Using taxonomic domain knowledge in text categorization tasks. *International Journal of Intelligent Control and Systems* (2007). Special Issue on Distributed Intelligent Systems
21. Mitchell, T.: Conditions for the equivalence of hierarchical and flat Bayesian classifier. Technical report, Center for Automated Learning and Discovery, Carnegie-Mellon University (1998)
22. Ghazi, D., Inkpen, D., Szapkowicz, S.: Hierarchical approach to emotion recognition and classification in texts. In: Farzindar, A., Kešelj, V. (eds.) *AI 2010. LNCS (LNAI)*, vol. 6085, pp. 40–50. Springer, Heidelberg (2010). doi:[10.1007/978-3-642-13059-5\\_7](https://doi.org/10.1007/978-3-642-13059-5_7)
23. Al-Hajjar, D., Syed, A.Z.: Applying sentiment and emotion analysis on brand tweets for digital marketing. In: *2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, pp. 1–6. IEEE (2015)

24. Esuli, A., Sebastiani, F.: SentiWordNet: a publicly available lexical resource for opinion mining. In: Proceedings of LREC, vol. 6, pp. 417–422. Citeseer (2006)
25. Mohammad, S.M., Kiritchenko, S.: Using hashtags to capture fine emotion categories from tweets. *Comput. Intell.* **31**, 301–326 (2015)
26. Ghazi, D., Inkpen, D., Szpakowicz, S.: Hierarchical versus flat classification of emotions in text. In: Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pp. 140–146. Association for Computational Linguistics (2010)