Using Bio-inspired Intelligence for Web Opinion Mining

George Stylios
University of Patras, Dept. of Computer Engineering and Informatics, 26500, Patras, Greece

Christos D. Katsis
Technical Educational. Institute of Ionian Islands, Dept of Business Admin, Inf. Systems Section 3100, Lefkada, Greece

Dimitris Christodoulakis
University of Patras, Dept. of Computer Engineering and Informatics, 26500, Patras, Greece

ABSTRACT

This work proposes a bio-inspired based methodology in order to extract and evaluate user's web texts / posts. To validate the methodology, a dataset is constructed using real data arising from Greek fora. The obtained results are compared with a commonly used machine learning technique (decision trees- C4.5 algorithm). The bio-inspired algorithm (namely the hybrid PSO/ACO2 algorithm) achieved average classification accuracy 90.59% in a 10 fold cross validation experiment, outperforming the C4.5 algorithm (83.66%). The proposed methodology could be easily integrated with a decision support system providing services in the fields of e-commerce or e-government in order to help merchants acquire customer satisfaction or public administrators capture common understanding.

General Terms

Web data mining, Decision support systems, Pattern recognition, Algorithms

Keywords

Artificial Intelligence, Bio-inspired Algorithms, Decision Trees, PSO/ACO2, Web texts, Web Opinion Mining

1. INTRODUCTION

Recently, there has been a shift from just *existing* on the Web to *participating* on the Web. Community applications such as collaborative wikis, blogging, photo and bookmark sharing, as well as online social networks have become very popular, both in personal/social and professional domain [1]. According to Eurostat [2] the percentage of individuals in the EU who used the internet in 2012 was 73% while the number of people using on line services (e.g e-commerce, e-government) continuously increases. A huge amount of data and information are available on the internet. Companies, governments and newspapers are motivating people to write and post their views and comments in electronic forms which are upon their sites for this purpose. For example companies want to know what their customers think about a product. Additional governments need to understand what the society thinks about a new law, decision, or policies [3].

Electronic government, or e-government, is correlated with the use of the web in the management and delivery of public services, by enhancing the efficiency of the public sector and developing more personal, customized relations between citizens and their government. E-government indicates that management services and functions are transferred onto the internet. Thus, it is a way for governments to use the most innovative information and communication technologies to offer citizens efficient access to information and services [4, 5]. The Semantic Web plays a crucial role in automatic delivery of customized e-government services. Furthermore, it extends the existing Web by providing a framework for technologies that give meaning to data and applications for automatic processing. On the other hand, it is well known that "What other people think" has always been an important piece of information for most of us during the decision making process. The Web has now made it possible to find out about the opinions and the experiences of those in the vast pool that are neither our personal acquaintances nor well known professional critics – that is, people we have never heard of. And conversely, more and more people are making their opinions available to strangers via the internet. Nowadays, the web is the most important place for expressing sentiments, evaluations, and reviews. Lots of people are tending to give their opinions in forums, blogs or wikis. However, with the rapid growth of e-commerce and e-government activity, the number of reviews and opinions about products or political ideas has increased exponentially and this source of information is becoming unworkable. For example, a customer who wants to buy a product usually searches information on the Internet looking for analyses of other customers. In fact, web sites such as Amazon, Epinions or IMDb can affect the customer decision. Nevertheless, it is becoming an impossible task to read all of these reviews and opinions in different forums or blogs. On the other hand, it is also very difficult for the companies to track this amount of evaluations about their products or services. Therefore, it is necessary to develop new methods that can improve the access to this kind of information. The automatic processing of documents to detect opinion expressed therein, as a unitary body of research, has been denominated opinion mining [6].

Opinion mining (O.M.) is a recent research area in the field of the text mining that has been designated by different terms like subjectivity analysis, sentiment analysis or sentiment orientation. According to Dave et al. [7], the ideal opinion-mining tool would "process a set of search results for a given item, generating a list of product attributes (quality, features, etc.) and aggregating opinions about each of them (poor, mixed, good)". Much of the subsequent research self identified as opinion mining fits this description in its emphasis on extracting and analyzing judgments on various aspects of given items. Although it focuses on the automatic identification and extraction of opinions from text and multimedia, the term has recently also been interpreted more broadly to include many different types of analysis of evaluative text [8]. Motivation for this component is based on providing support for decision makers to automatically track attitudes on certain topics in online media and user generated content [9]. For example, opinion detection has been proposed as a key enabling technology in e-Rulemaking, allowing the automatic analysis of the opinions that people submit about pending policy or government-regulation proposals [10, 11]. Different approaches have been applied in the field of opinion mining but two main methodologies used can be distinguished. On the one hand, there is a lot of work based on the symbolic approach, which applies manually crafted rules and lexicons. The documents in this approach are represented as collections of words. Then, the sentiment of each word can be determined by different methods, for example, using a web search [12] or consulting a dictionary like WordNet [13] . On the other hand, the use of machine learning techniques is reported for the classification of reviews according to their orientation. In these approaches, the documents are usually represented by different features for the classification task. Then, a machine learning algorithm is applied. These features may include the use of n-grams or defined grammatical roles like, for instance, adjectives. Machine learning algorithms commonly used are Support Vector Machines, Maximum Entropy or Naïve Bayes [6, 14-17].

In this work, a methodology that combines both approaches (symbolic and machine learning) is proposed. Moreover, instead of the "traditional" artificial intelligence techniques (e.g neural networks, support vector machine, decision trees, etc) one of the latest biologically inspired intelligent methods namely the hybrid PSO/ACO2 algorithm [18] is used. Furthermore, the classification results of the C4.5 algorithm (decision trees) are provided for comparison purposes. Biologically inspired techniques have recently appeared in literature. Their main concept relies on the evolution principles of real-life species. The term swarm intelligence (SI) describes the algorithms that mimic the collective behavior of decentralized, self-organized organisms, such as ants, birds, etc. PSO/ACO2 is a rule induction technique, thus it can offer an explanation of the classification outcome which in turn could result to higher acceptability by decision makers. As such, it constructs/discovers classification rules in the form IF (term1 AND term2 AND …) THEN (predicted class), where each term is in the form <feature=value>. PSO/ACO2 is a hybrid method that combines Ant Colony Optimization (ACO) for the selection of nominal terms, sequential covering for rule extraction and Particle Swarm Optimization (PSO) for the inclusion of continuous terms in the rules. More details regarding the PSO/ACO2 algorithm are given in section III.

In the following sections, initially, the related work reported in the literature is provided. Next, the proposed methodology is presented and validated using real data. Finally, limitations, future work and conclusions are given.

2. RELATED WORK

Although the area of sentiment analysis and opinion mining has recently enjoyed a huge burst of research activity, there has been a steady undercurrent of interest for quite a while. One could count early projects on beliefs as forerunners of the area [19, 20]. Opinion detection has been proposed as a key enabling technology in e-Rulemaking, allowing the automatic analysis of the opinions that people submit about pending policy or government-regulation proposals [10, 21, 22]. Marketing researchers, companies and governments have long been interested in capturing information and knowledge about the opinions of potential buyers or citizens. However, interviewing people about their opinions is time consuming and costly, and there is concern if the individual is telling the truth or telling the marketer what they want to hear. In contrast, blogs provide a readily available and opinion-based content media that provides sentiment about a range of issues. Further, that qualitative content can be matched against key performance indicators, such as sales, profits or stock price. As a result, being able to use those blogs for gathering opinion information potentially can provide a low cost source of information about those opinions and sentiment, regarding particular issues and concerns, gathered in real time [23].

Opinion mining (OM) has lately become a topic of interest trying to combine statistics, Artificial Intelligence and Data Mining technologies in a unified framework [24]. It is a recent subdiscipline at the crossroads of information retrieval and computational linguistics which is concerned not with the topic a document is about, but with the opinion it expresses. OM has a rich set of applications, ranging from tracking users' opinions about products or political candidates as expressed, to customer relationship management [25]. Users'/citizens' opinions can be used as guidelines for companies to change their strategies toward specific target groups, customers to decide on the purchase of a product or destination place for their holidays and lately for governments to improve services, launch campaigns etc. [26].

Web opinion mining aims to extract, summarize, and track various aspects of subjective information on the Web. Ku, [26] applied web mining techniques to mine positive and negative sentiment words and their weights on the basis of Chinese word structures. Xu [27] proposed a system for opinion mining using poll results on the web dealing with opinion answering question, opinion mining on a single object and opinion mining on multiple objects. Furuse [28] developed a search engine that can extract opinion sentences relevant to an open-domain query -based not only on positive or negative measurements but also on neutral opinions, requests, advice, and thoughts- from Japanese blog pages. Quan et al. [29] proposed a method for measuring the similarity between two short text snippets by comparing each of them with the probabilistic topics. The method starts by finding the distinguishing terms between two short text snippets and comparing them with a series of probabilistic topics, extracted by the Gibbs sampling algorithm. Weng et al. [30] proposed a two-level retrieval solution based on a document decomposition idea. According to their approach, a document is decomposed to a compact vector and a few document specific keywords by a dimension reduction technique. The compact vector embodies the major semantics of a document, and the document specific keywords complement the discriminative power lost in dimension reduction process. The authors adopt locality sensitive hashing (LSH) to index the compact vectors and next then they re-rank documents in this set by their document specific keywords. By this way, it is able to index large-scale corpus for efficient similarity-based document retrieval. Nallapati et al [31], proposed a model, called Link-PLSA-LDA, that combines Probabilistic Latent Semantic Analysis (PLSA) and Latent Dirichlet Allocation (LDA) into a single framework, and explicitly models the topical relationship between the linking and the linked document, providing interesting visualizations of topics and influential blogs on each topic. Xiaowen Liu [32], proposed the use of linguistic rules to deal with the problem of determining the semantic orientations of opinions expressed on product features in reviews using an opinion aggregator function providing a system called Opinion Observer. Miao proposed AMAZING [33], a sentiment mining and retrieval system which mines knowledge from consumer product reviews by utilizing data mining and information retrieval technology based on a ranking mechanism taking temporal opinion quality and relevance into account to meet customers' information needs. Zhai developed Opinion Observe [34] to compare consumer opinions of different products based on online reviews, while Sun [35] created BlogHarvest which is a blog mining and search framework that extracts the interests of the blogger, finds and recommends blogs with similar topics and provides blog-oriented search functionality. An opinion utility named Jodange was built in the Leveraging Cornell University [36-38]. Jodgane identified opinion holders on issues, organizations, or people of interest. It was able to track the impact of an issue via publication, region, opinion holder, tonality or any other measurement, uncover important sentiment trends on key issues and correlate opinions against specific outcomes. VIStology's IBlogs project [39], provided blog analysts a tool for monitoring, evaluating, and anticipating the impact of blogs by clustering posts by news event and ranking their significance by relevance, timeliness, specificity and credibility. Jin et al. [40] proposed a unified framework based on Probabilistic Latent Semantic Analysis to create models of Web users, taking into account both the navigational usage data and the Web site content information. The model was based on a set of discovered latent factors that explained the underlying relationships among page views in terms of their common usage and their semantic relationships. Based on the discovered user models, the authors proposed algorithms for characterizing Web user segments and to provide dynamic and personalized recommendations based on these segments.

The need for identifying opinions has motivated a number of automated methods for detecting opinions or other subjective text passages [41-47] and assigning them to subcategories such as positive and negative opinions [48-51]. A variety of machine learning techniques have been employed for this purpose, generally based on lexical cues associated with opinions. Current approaches share a common pattern. They focalize on an entire document [48, 49] or on complete sentences [43, 52, 53].

The above mentioned text mining / sentiment analysis approaches main scope is the extraction of users'/citizens' opinions and the classification of polarities from web texts. A limited number of work reported centers on the evaluation of text coherence making use of statistical or artificial intelligence methodologies [54-56].

In this work, a methodology based on a biologically inspired hybrid classifier namely (Particle swarm optimization -PSO) PSO/ACO2 (Ant Colony Optimization-ACO) is proposed, able to detect and analyze the public's comments (posts) towards commercial products and governmental decisions. Posts are processed in order to: i) Identify and extract users' opinions and ii) automatically weight the users' comments according to the presence or not of arguments justifying their opinion about a certain service, product or governmental decision using the PSO/ACO2 [18, 57] algorithm.

3. PROPOSED METHODOLOGY

Our methodology infers the users' web-texts by conducting a two level analysis: During the first level phrases containing user opinions are detected and extracted from their posts. Apparently, it is of great importance not only to extract someone's opinion, but also to estimate if someone supports his/hers opinion using arguments. Thus, in the second analysis level, quantitative features are extracted from the posts which are then fed into the PSO/ACO2 and C4.5 classifiers to categorize them as supported or not by arguments. These analysis levels are described in the following paragraphs.

3.1 Extraction of user's opinion from real web content textual data

In this analysis level, on line postings and users' evaluations regarding commercial products or governmental decisions are detected and extracted from fora. Initially, the content of the users' posts is downloaded and non-textual elements (etc. images, symbols) are eliminated by applying HTML. Then tokenization is applied to the postings' body in order to extract the lexical elements of the user generated text. Afterwards, the text is passed through a Part of Speech tagger, responsible for identifying tokens and annotate them to appropriate grammar categories. Also the topics being discussed as well as the user's opinion regarding those topics have to be identified. To detect such references within a post a syntactic dependency parser is employed, able to identify the proper noun to which every adjective refers when given as input text containing adjectives. To obtain the opinion phrases communicated on user postings, we rely on the syntactically depended adjective–noun pairs, which are all collected to a dataset. To identify how users evaluate commercial products or governmental decisions, the motion of word's semantic orientation is used. To obtain the semantic frame of an adjective, every adjective extracted from the harvested postings against an ontology which contains fully annotated lexical units is examined. Sentiment analysis of users' opinions refers to labeling opinion phrases with a suitable polarity tag (positive or negative) to the adjectives. The criterion under which labeling takes place, is that positive adjectives give praise to the topic, while negative adjective criticizes it.

3.2 Classification of postings

At the second analysis level of the proposed approach, a classifier is used to classify harvested postings into two categories: i) postings supported by arguments and ii) postings not supported by arguments. This analysis level is based in the underlying fact that when people weigh or consider an idea in order to express their arguments about a topic, their written language has usually a more complicated structure and organization compared with posts expressing just their opinion about the subject. Therefore, the way parts of speech and punctuations symbols are used alters [23]. Our approach takes advantage of this differentiation in order to automatically extract correlations between parts of speech and presence or absence of arguments. More specific, initially parts of speech (traced during the previous analysis level), as well as punctuation symbols, use of capital letters, word and spelling mistakes are counted per post. Then, quantitative features are extracted (per post) as presented in Table 1.

Table 1. The extracted features

Feat.#	Feat. description
1	# of words per comment
2	#of conjunctions divided by the # of words per comment
3	# of sentences divided by the # of words per comment
4	# of pronouns divided by the # of words per comment
5	# of adverbs divided by the # of words per comment.
6	# of prepositions divided by the # of the words per comment.
7	# of punctuation symbols used divided by the # of the words per comment
8	# of articles divided with the # of the words per comment
9	# of nouns by the # of the words per comment
10	# of adjectives divided by the # of the words per comment.
11	# of verbs divided by the # of the words per

	comment
12	Usage of uppercase letters or not (usage designated as 1, where lack of usage was designated as 0).
13	Usage of punctuation symbols i.e. dots, commas, interrogation marks etc (usage designated as 1, where lack of usage was designated as 0)
14	# of spelling mistakes divided by the # of the words per comment

Finally, the features are fed to the classifier in order to automatically classify users' postings into "supported by arguments" and "not supported by arguments". In this work, two classifiers have been employed: The PSO/ACO2 hybrid bio inspired algorithm and one "traditional" artificial intelligence classifier namely decision trees (C4.5 algorithm) [58] for comparison purposes. In the following paragraphs, the key functions of the aforementioned algorithms are provided.

3.3 The Decision tree classifier – C4.5 Algorithm

The construction of the decision tree is implemented using the C4.5 inductive algorithm [58]. The essence of the algorithm is to construct a decision tree from the training data. Each internal node of the tree corresponds to a principal component, while each outgoing branch corresponds to a possible range of that component. The leaf nodes represent the class to be assigned to a sample. The C4.5 algorithm applies to a set of data and generates a decision-tree, which minimizes the expected value of the number of tests for the classification of the data. Moreover, the C4.5 algorithm can solve the overfitting problem using a post-pruning method. The most important feature in the C4.5 algorithm is its ability to automatically select the feature which is appropriate at each node. The feature of each node is selected in order to divide input samples effectively. Information gain is used as a measure of effectiveness. In order to define the information gain, a measure called entropy is initially defined, which is the degree of complexity about input samples. In the case of existing c classes in a set S, the entropy of S, H(S) , is defined as:

$$H(S) \equiv \sum_{i=1}^{c} -p_i log_2 p_i \qquad (1)$$

Where $p_i$is the ratio of the class i in the set *S*. Then, the information gain is defined as the reduction of entropy. Information gain for a feature A, *Gain (S, A)*, is obtained as

$$Gain(S, A) \equiv H(S) - \sum_{u \in Values(A)} \frac{|S_u|}{|S|} H(S_u) \qquad (2)$$

Where *Values(A)* represents the range of features *A* and S_u is a subset of *S* having *u* as a result of feature *A*. In our case, some of the extracted features are continuous valued. Continuous-valued features can be incorporated into the decision tree by dynamically defining new discrete-valued features which partition the continuous feature value into a discrete set of intervals. More precisely, for a feature A that is continuous valued, the algorithm can dynamically create a new Boolean feature A_t that is true if $A \leq t$ and false otherwise. The major problem is the selection the best value for the threshold *t*. A threshold *t* must be selected which produces the greatest information gain. By sorting the training instances according to the continuous feature *A* and then identifying adjacent examples that differ in their target classification, a set of candidate thresholds is generated midway between the corresponding values of *A*. These candidate thresholds can then be evaluated by computing the information gain associated with each one. The information gain is computed for each of the candidate features, and the one with the highest information gain is selected. After the induction of the decision tree, a post pruning method [58] is applied and more specifically the pessimistic pruning, in order to avoid overfitting. This method prescribes that if the predicted error for a root node in a subtree is less than the predicted error for the subtree, then a subtree will be replaced with its root node, which becomes a new leaf in a pruned tree. Decision trees offer interpretability of classification by constructing decision rules.

Decision rules come in the form if < antecedent >, then < consequent >. The antecedent consists of the feature values from the branches taken by the particular path through the tree, while the consequent consists of the classification value for the target variable given by the particular leaf node.

3.4 The PSO/ACO2 algorithm

PSO/ACO2, proposed by Holden and Freitas [57] belongs to the family of bio inspired algorithms. It is inspired by Particle Swarm Optimization (PSO) [59] and Ant Colony Optimization (ACO) [60]. PSO is a stochastic optimization method that mimics the behavior of bird flocking or fish schooling. In PSO, the potential solutions are called particles, which fly through the problem space by following the current optimum particles. Particles keep track of their coordinates, which are associated with the best solution they have achieved so far. The Ant Colony Optimization algorithm (ACO) is a probabilistic technique for solving computational problems which can be reduced to finding good paths through graphs. PSO/ACO2 is a hybrid algorithm for classification rule generation.

The rule discovery process in PSO/ACO2 algorithm is performed into two separate phases. Specifically, in the first phase, a rule is discovered using nominal attributes only, using a combination of ACO and PSO approach. In the second phase, the rule is extended with continuous attributes. The PSO/ACO2 algorithm uses a sequential covering approach to extract one classification rule at each iteration. The sequential covering approach initializes the rule set with an empty set, and then, for each class the algorithm performs a loop, where a set is used to store the set of training examples the rules will be created from. At each iteration of this loop, PSO/ACO2 discovers a rule-based only on nominal attributes. A particle is actually an example from the training set and both ACO and PSO are applied in each dimension of this particle. A combination of pheromone updating (ACO) and particle term selection (PSO) is used and a nominal rule is extracted. Pheromone is updated according to the rule quality that is measured using a modification of the precision with Laplace correction, as:

$$(1 + TP) / (1 + TP + FP) \qquad (3)$$

The generated rule is not complete as it does not include terms with continuous values. In order to include terms with continuous values, the best rule discovered by the PSO/ACO2 algorithm is used as a base for the discovery of terms with continuous values. For the continuous part of the rule, a standard PSO algorithm is applied to continuous attributes. The vector that is optimized by the PSO consists of two dimensions for every continuous attribute, one for the lower bound and one for the upper bound. At every particle evaluation, the vector is converted to a set of terms, denoting the rule conditions, which are subsequently added to the rule produced by the PSO/ACO2 algorithm fitness evaluation. After the creation of the best rule, with both nominal and continuous terms, the rule is pruned and the training examples covered by that rule are removed from the training instances; thus the next rule is generated from the reduced training instances. This procedure is repeated until the training instances become less than a threshold of maximum uncovered cases per class. Finally, after all rules have been extracted, a second pruning procedure is applied that removes both terms from the already found rules, as well as, rules from the rule set. The pseudocode of the PSO/ACO2 algorithm for rule discovery is shown in Algorithm 1.

```
RuleSet = Ø
TE = number of training examples per class
FOR Each Class
WHILE TE> Max_uncovered_training_instances per class
FOR i=1 to Maximum Iterations
   FOR j=1 to number of particles
    Qbest=0/*Qbest is past best quality*/
    R = Empty/*Rule*/
    FOR Each Dimension
     Initialize T/*Term*/
    END FOR
Probabilistically Choose T for Nominal
Features
    Calculate Q(j)/*Quality*/
    IF Q(j)> Qbest
     THEN   Add T to R
      Qbest=Q(j)
    END IF
    FOR each dimension
     Update Pheromone
    END FOR
END FOR
END FOR
Add continuous T to R
Choose b_upper, b_lower/*upper and lower bounds*/
Optimize b_upper, b_lower using PSO
Return Best R
Prune Best R
Add Best R to RuleSet
Remove Examples covered by Best R
TE = TE – (number of Examples removed)
END WHILE
END WHILE
Prune RuleSet
```

Algorithm 1: PSO/ACO2 pseudo code

4. EXPERIMENTAL RESULTS

To evaluate our proposed methodology a set of real citizen posts downloaded from Greek forums is collected. After processing these comments as described previously, the included distinct opinion phrases are extracted. Data derived from 563 users' comments were carefully read by an experienced sociologist who annotated each one of them according to the presence or not of arguments. According to the expert 395 posts were classified as "presence of arguments" and 168 posts were classified as "absence of arguments". The expert's opinion is used as a golden standard for our classification schema. Next, quantitative features based on parts of speech, the use or not of capital letters and the use or not of punctuation symbols are extracted from the users' posts as described in Table 1 are automatically computed. A dataset has been constructed consisting of these features as well as the expert's annotation for each post.

In order to avoid model selection bias, and to estimate and compare the classification performance of the C4.5 and PSO/ACO2 algorithms, a 10 fold cross-validation (or rotation estimation) [61-63] procedure is followed. Data are first partitioned into 10 equally sized segments or folds. Subsequently 10 iterations of training and validation are performed such that within each iteration a different fold of the data is held-out for validation while the remaining folds are used for learning. Data are stratified[1] prior to being split into folds. The average classification results of the C4.5 and the PSO/ACO2 algorithms are provided in Table 2. It is obvious that the PSO/ACO2 algorithm outperforms the C4.5 classification performance (90.59% and 83.66% respectively).

Table 2. C4.5 and PSO/ACO2 classification results

Algorithm	**Classification accuracy (%)**	**Standard deviation (STD)**
C4.5	83.66	+ 5.31
PSO/ACO2	90.59	+ 3.41

In order to have an in depth view of the achieved results, the confusion matrix for the C4.5 and the PSO/ACO2 algorithms for the above classification schemas are provided in Table 3 and Table 4 respectively.

Table 3. Confusion matrix for the C4.5 algorithm

Algorithm C4.5	**Classified as "Absence of arguments"**	**Classified as "presence of arguments"**
"Absence of arguments"	119	49
"presence of arguments"	43	352

Table 4. Confusion matrix for the PSO/ACO2 algorithm

Algorithm PSO/ACO2	**Classified as "Absence of arguments"**	**Classified as "presence of arguments"**
"Absence of arguments"	148	20
"presence of arguments"	33	362

Finally, the classification results in terms of *Sensitivity*[2] and *Specificity*[3] are given in Table 5.

[1] Stratification is the process of rearranging the data as to ensure each fold is a good representative of the whole.

[2] $Sensitivity = \frac{TP}{FP+FN}$

[3] $Specificity = \frac{TN}{FP+TN}$

Table 5. Classification results in terms of sensitivity and specificity

Algorithm	Sensitivity (%)	Specificity (%)
C4.5	73.46	87,78
PSO/ACO2	81,77	94,76

5. DISCUSSION

In this work, a methodology that extracts users' opinions from text web sources (e.g blogs) and classifies them according to the presence or not of arguments is presented. To validate the proposed methodology, a database is constructed consisting of the extracted features as well as the annotation per post provided by an expert, for a total of 563 posts. The classification schema used consists of the PSO/ACO2 and C4.5 algorithms, trained and tested with the above-mentioned database. As it can be seen from Table 2 and Table 5 the achieved results accomplished from the bio-inspired algorithm PSO/ACO2 are superior in terms of *Sensitivity*, *Specificity* and *Accuracy*.

Our system has been validated using real data arising from Greek blogs, (posts are written in Greek). Regarding the proposed methodology's drawbacks it must be mentioned that in this stage it cannot be used in post's written in different languages (i.e. English) since the machine learning algorithms used have to be retrained. On the other hand, the training procedure is realized only once, therefore, our system can easily be retrained - modified in order to be used with different languages. As far for the computational time for model building is concerned, C4.5 is the less computational intensive, needing 1.2 seconds for a 10 fold cross-validation experiment, while the most computational intensive PSO/ACO2 algorithm needs more than 3 hours. In any case once model building (training) is finished, classification time for a new case is negligible. All the above figures correspond to an Intel Core 2 Duo CPU 2.80 GHz – 4 Gb RAM.

Future work comprises extended testing in different forum topics and other languages. In this way, different system's versions will be implemented providing optimized classification results regarding forums dedicated to different topics (e.g politics, customers' satisfaction, etc) or different languages. Moreover, the proposed system will be integrated with a model that harvests user-generated content from fora and applies lexico-semantic processing to the collected data. Such models have already been proposed by the authors previously and the interested reader can find more information about it in [64]. By this way, valuable information about users' opinions will be extracted automatically from fora. Also, although the obtained accuracy is quite high (PSO/ACO2 achieved classification *accuracy* 90.59%) future work might focus on the combination of classifiers on what is called "ensemble based classification" in machine learning. Ensemble classifiers combine the advantages of their component classifiers in an analogous to multi-voting schema. Further tests on even larger real datasets acquired from a variety of forums could verify the experimental results.

6. CONCLUSIONS

A methodology in order to extract users' comments as well as the presence or absence of arguments regarding the stated comments is presented. Also the applicability of a state of the art bio – inspired classifier (PSO/ACO2) has been examined. The PSO/ACO2 algorithm seems to perform satisfactory in terms of accuracy, sensitivity and specificity, presenting a low variance among different experiments. Therefore it should be considered as promising classifier for the domain. Our methodology could become a part of an integrated decision support system; such a system could provide services and lexical tools able to help merchants acquire valuable feedback (e.g. consumers' satisfaction regarding their products) and facilitate public administrations to capture the common understanding of e-Government.

7. ACKNOWLEDGMENTS

The authors would like to express their *sincere* gratitude to Dr. Sofia Stamou for her valuable suggestions and guidance.

8. REFERENCES

[1] Kolbitsch, J. and Maurer H., The transformation of the Web: How emerging communities shape the information we consume. Journal of Universal Computer Science, 2006. 12(2): p. 187-213.

[2] Seyberd, H., Internet use in households and by individuals in 2012, t.a.s. Industry, Editor. 2012, Eurostat.

[3] Xu, G., Zhang, Y., Modelling User Behaviour for Web Recommendation Using LDA Model, in in Proceedings of International Workshop on E-Commerce, Business, and Services. 2008.

[4] Breck, E. and Y. Choi, Joint extraction of entities and relations for opinion recognition, in Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. 2006: Morristown, NJ, USA.

[5] Ali, H., Macaulay, L., Zhao, L. A Collaboration Pattern Language for e-Participation":A Strategy for Reuse. in Proceedings of the 9th European Conference on e-Government. 2009.

[6] Rushdi Saleh, M., et al., Experiments with SVM to classify opinions in different domains. Expert Systems with Applications, 2011. 38(12): p. 14799-14804.

[7] Dave, K., Lawrence, S., Pennock D.M., Opinion extraction and semantic classification of product reviews, in In Proceedings of WWW. 2003. p. 519-528.

[8] Chesley, P., Vincent, B. , Xu, L., Srihari, R. Using verbs and adjectives to automatically classify blog sentiment. in Proceedings of AAAI-CAAW. 2006.

[9] Lin, W.H., Wilson, T., Wiebe, J., Hauptmann, A. Which side are you on? Identifying perspectives at the document and sentence levels. in Conference on computational natural language learning. 2006.

[10] Allen, C.T., et al., A place for emotion in attitude models. Journal of Business Research, 2005. 58(4): p. 494-499.

[11] Shulman, S., Hovy, E., Callan, J., Zavestoski, S. Language processing technologies for electronic rulemaking. in Proceedings of InterNational Conference on Digital Government Research. 2006.

[12] Hatzivassiloglou, V., Wiebe, J., Effects of adjective orientation and gradability on sentence subjectivity, in International Conference on Computational Linguistics. 2000. p. 299-305.

[13] Kamps, J., Marx, M., Mokken, R.J., Rijke, M.D., Using wordnet to measure semantic orientation of adjectives., in Conference on language resources and evaluation. 2004. p. 1115-1118.

[14] Pang, T.B., Pang, B., Lee, L. Thumbs up? sentiment classification using machine learning. in In Proceedings of EMNLP. 2002.

[15] Mullen, T., Collier, N., Sentiment analysis using support vector machines with diverse information sources, in In Proceedings of the conference on empirical methods in natural language processing. 2004. p. 412-418.

[16] Prabowo, R., Thelwall, M., Sentiment analysis: A combined approach. Journal of Informetrics, 2009. 3(2): p. 14.

[17] Stylios, G., Katsis, C.D., Glaros, C., Simaki, V., Christodoulakis, D., MiCoGo, an Integrated System That Automatically Detects the Presence of Opinion in web Texts Regarding eCommerce and eGovernment, in The 12th European Conference on e-Government (ECEG, 2012). 2012: Barcelona, Spain. p. 683-690.

[18] Holden, N.P. and A.A. Freitas, A hybrid PSO/ACO algorithm for classification, in Proceedings of the 2007 GECCO conference companion on Genetic and evolutionary computation. 2007, ACM: London, United Kingdom. p. 2745-2750.

[19] Chomsky, N., Peng, F. C. C., Studies on Semantics in Generative Grammar American Anthropologist, 1973. 75(6): p. 1918-1921.

[20] Carbonell, J., Subjective Understanding: Computer Models of Belief Systems. PhD thesis, in Yale. 1979.

[21] Kwon, N., Shulman, S. W., Hovy, E, Multidimensional text analysis for eRulemaking, in Proceedings of DG.O (Inter)National Conference on Digital Government Research. 2006.

[22] Shulman, S., Hovy, E., Callan, J., Zavestoski, S. Language processing technologies for electronic rulemaking. in Proceedings of DG.O (Inter)National Conference on Digital Government Research. 2006.

[23] O'Leary, D.E., Blog mining-review and extensions: “From each according to his opinion”. Decision Support Systems, 2011. 51(4): p. 821-830.

[24] Pang, B., Lee Li. Opinion Mining and Sentiment Analysis. Foundations and Trends in information Retrieval, 2008. 2, 1-135 DOI: DOI:10.1561/1500000001.

[25] Esuli, K., Opinion mining. 2009.

[26] Ku, L.W. and H.H. Chen, Mining opinions from the web: Beyond relevance retrieval. Journal of the American Society for Information Science and Technology, 2007. 58(12): p. 1838-1850.

[27] Xu, Z. and R. Ramnath. Mining opinion from Poll Results in Web Pages in www2009. 2009. Madrid, Spain.

[28] Furuse, O., et al. Opinion sentence search engine on open-domain blog. in Proccedings of the 20th International Joint Conference on Artificial Intelligence in Hyderabad. 2007. India.

[29] Quan, X.J., et al., Short text similarity based on probabilistic topics. Knowledge and Information Systems, 2010. 25(3): p. 473-491.

[30] Weng, L., et al., Query by document via a decomposition-based two-level retrieval approach, in Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. 2011, ACM: Beijing, China. p. 505-514.

[31] Nallapati, R., Cohen, W., A New Unsupervised Model for Topics and Influence of Blogs, in In International Conference for Weblogs and Social Media. 2008. p. 84-92.

[32] Xiaowen, D., Liu, B. The utility of linguistic rules in opinion mining. in Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval 2007. New York.

[33] Miao, Q.L., Q.D. Li, and R.W. Dai, AMAZING: A sentiment mining and retrieval system. Expert Systems with Applications, 2009. 36(3): p. 7192-7198.

[34] Zhongwu. Zhai, et al. Clustering Product Features for Opinion Mining. in Proceedings of 4th ACM International Conference on Web Search and Data Minin, . 2011. Hong Kong, China.

[35] Sun, J.T., et al. A comparative web search system. in 15th International Conference on World Wide Web (WWW). 2006. Edinburg, Scotland.

[36] Breck, E., Y. Choi, and C. Cardie. Identifying expressions of opinion in context. in In Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-2007). 2007. Hyderabad, India,.

[37] Breck, E., Y. Choi, and C. Cardie. Joint extraction of entities and relations for opinion recognition. in Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. 2006. Morristown, NJ, USA.

[38] Choi, Y. and C. Cardie. Structured local training and biased potential functions for conditional random fields with application to coreference resolution. in In Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; . 2007. Rochester, New York.

[39] Ulincy, B., et al. Uses of Ontologies in Open-Source Blog Mining in In proceedings of the Ontology for the intelligence Community. 2007. Columbia, Maryland.

[40] Jin, X., Zhou, Y., Mobasher, B. A unified approach to personalization based on probabilistic latent semantic models of web usage and content. in AAAI 2004 Workshop on Semantic Web Personalization (SWP’04). 2004. San Jose.

[41] Wiebe, J. Learning subjective adjectives from corpora. in In Proceedings of the 17th National Conference on Artificial Intelligence (AAAI -2000). 2000.

[42] Wiebe, J., et al., Learning subjective language. 2002, University of Pittsburgh: Pennsylvania. .

[43] Hatzivassiloglou, V. and J.M. Wiebe, Effects of adjective orientation and gradability on sentence subjectivity, in Proceedings of the 18th conference on Computational linguistics - Volume 1. 2000, Association for Computational Linguistics: Saarbr\&\#252;cken, Germany. p. 299-305.

[44] Roussinov, D. and J.L. Zhao, Automatic discovery of similarity relationships through Web mining. Decision Support Systems, 2003. 35(1): p. 149-166.

[45] Wong, T.-L. and W. Lam, An unsupervised method for joint information extraction and feature mining across different Web sites. Data &amp; Knowledge Engineering, 2009. 68(1): p. 107-125.

[46] Yang, H.-C. and C.-H. Lee, A text mining approach on automatic generation of web directories and hierarchies. Expert Systems with Applications, 2004. 27(4): p. 645-663.

[47] Velásquez, J.D., L.E. Dujovne, and G. L'Huillier, Extracting significant Website Key Objects: A Semantic Web mining approach. Engineering Applications of Artificial Intelligence, 2011. 24(8): p. 1532-1541.

[48] Turney, P. Thumps up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. in In Proceedings of the 40th Annual Meeting of the association for Computational linguistics. 2002.

[49] Pang, B., L. Lee, and S. Vaithyanathan. Thumps up? Sentiment classification using machine learning techniques. in In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP-02). 2002.

[50] Yu, H. and V. Hatzivassiloglou. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. in In Proceedings of the Conference on empirical Methods in Natural language Processing 2003.

[51] Stylios, G., et al. Deciphering the Public Stance Towards Govermental Decisions. in 10th European Conference on E-Government. Ireland.

[52] Wiebe, J., R. Bruce, and T. O'Hara. Development and use of a gold standard data set for subjectivy classifications. . in In Proceedings of the 37th Annual Meeting of the Association for Computational Lingyistics (ACL-99). 1999.

[53] Stylios G., et al., Public Opinion Mining for Governmental Decisions. Online Journal for Government., 2011.

[54] Eneva, E., Hoberman, R., Lita, L. , Learning Within-Sentence Semantic Coherence, in Empirical Methods in Natural Language Processing. 2001.

[55] Lapata, M., Barzilay, R. , Automatic evaluation of text coherence: Models and representations, in Proceedings of the 19th International Joint Conference on Artificial Intelligence. 2005.

[56] Barzilay, R. and M. Lapata, Modeling local coherence: An entity-based approach. Computational Linguistics, 2008. 34(1): p. 1-34.

[57] Holden, N. and A.A. Freitas, A Hybrid PSO/ACO Algorithm for Discovering Classification Rules in Data Mining. Journal of Artificial Evolution and Applications, 2008. 2008.

[58] Quinlan, R.J., Programs for Machine Learning. 1993: Morgan Kauffman.

[59] Kennedy, J. and R. Eberhart. Particle swarm optimization. in Neural Networks, 1995. Proceedings., IEEE International Conference on. 1995.

[60] Dorigo, M., V. Maniezzo, and A. Colorni, Ant system: Optimization by a colony of cooperating agents. Ieee Transactions on Systems Man and Cybernetics Part B-Cybernetics, 1996. 26(1): p. 29-41.

[61] Geisser, S., Predictive Inference (Chapman & Hall/CRC Monographs on Statistics & Applied Probability). 1993: Published by Chapman and Hall/CRC

[62] Kohavi, R., "A study of cross-validation and bootstrap for accuracy estimation and model selection", in Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence. 1995, Morgan Kaufmann. p. 1137-1143.

[63] Devijver, P.A., Pattern Recognition: A Statistical Approach. 1982, London, GB: Prentice-Hall.

[64] Stylios, G., Christodoulakis, D., Besharat, J., Vonitsanou, M., Kotrotsos, I., Koumpouri, A. Stamou, S, Public Opinion Mining for Governmental Decisions. Electronic Journal of e-Governmen, 2010. 8(2): p. 14.