

Classification of opinions in cross domains involving emotive values

Rafael Guzmán-Cabrera^{a,*}, Belém Priego Sánchez^b, T. Prasad Mukhopadhyay^a,
J.M. Lozano García^a and T. Cordova-Fraga^c

^a*DICIS, Universidad de Guanajuato, GTO, Mexico*

^b*Department of Systems, Universidad Autónoma Metropolitana, Unidad Azcapotzalco, CDMX, Mexico*

^c*DCI, Universidad de Guanajuato Campus León, GTO, Mexico*

Abstract. It is increasingly common for internet users to have access to blogs and social networks, and common for them to express opinions on such sites. This research work is framed within the scope of opinion mining. Opinions allow us to measure people's perception of a specific topic or product. Knowing the opinion that a person has towards a product or service is of great help for decision making, since it allows, between other things, that potential consumers to verify the quality of the product or service before using it. This research work is framed within the scope of opinion mining.

When the number of opinions is very large the analysis gets more complicated and generally resort to tools that allow this task to be performed automatically are sought out. The present work performs an automatic categorization of textual opinions corresponding to four products: books, DVDs, kitchens, and electronics. Both negative and positive opinions are considered for the experiment. Further categorization experiments are performed using different domains of learning. The basic idea is to investigate if we can undertake classification of opinions, positive and negative, of any given domain using instances of training from a different domain. Results obtained from different methods of learning are presented. The results obtained allow us to examine the feasibility of the proposed methodology.

Keywords: Cross Domain, emotive classification, opinion classification

1. Introduction

The Internet has a profound impact on people's engagement with their work or leisure and their search for knowledge worldwide. One of the most successful service engines on the Internet has been the World Wide Web (better known as www or simply "the web"). Owing to the web, millions of people have easy and immediate access to a vast and diverse amount of information on-line.

Some companies and persons use the blogs available on-line in order to know and share opinions

regarding any topic or specific product related to their jobs, employers, colleagues, family thus creating a large amount of information and data [5]. These massive amounts of opinions issued through the web allow us to measure the perception of users but the overwhelming amount of information contained in these statements makes analysis or evaluation complicated [6]. For this reason, it is necessary to resort to automatic tools to perform such tasks [10].

Performing automatic processing of positively or negatively valanced opinions can be potentially useful for sustained decision-making processes in any company [18]. They could benefit from having a direct feedback from the users of the product or service which they consume [20, 27, 28]. They can also be used to examine media perception of a certain

*Corresponding author. Rafael Guzmán-Cabrera, DICIS, Universidad de Guanajuato Campus, GTO, Mexico. E-mail: guzmanc@ugto.mx.

figure in public life: for example, the perception that citizens have of the government or of a certain actor [38].

Subjective expressions are none other than words and phrases used to express mental and emotional states. Such states include speculation, evaluation, sentiment or belief. These are called private states: that is to say, internal states which cannot be directly observed by others. On the contrary, polarity refers to positive or negative associations of a word or sense. Whereas there is a dependency in that most subjective senses have a relatively clear polarity, polarity can be attached to objective words or senses as well. Su and Markert [35] give the example of the word tuberculosis: it does not describe a private state, is objectively verifiable and would not cause a sentence containing it to carry an opinion, but it does carry negative associations for the vast majority of people.

The problem of making a computer understand human natural language is very difficult. Instead of trying to tackle that complex problem directly, we consider the small subproblem of determining the sentiment of a given piece of text in natural language. For example, if someone writes a review of a particular movie, then the model would be given the task of automatically determining whether or not the review is positive or negative in sentiment. In order to make this determination, a sentiment mining model would, to some degree, have to understand the language being used by the reviewer. Sentiment mining models with a high level of classification accuracy would be useful across many different domains.

In the case of categorization of type of documents that are not thematic in nature, issues of attribution include identification of the author, where the style of writing would be associated with a given author [9]. Also, it is possible to find applications associated with the categorization of documents where the authors of texts are categorized in accordance with the prevalence of certain linguistic structures [31].

Currently, one of the most popular multi-category on-line shops is Amazon.com. With many products in many different departments, it has become a hugely popular option for on-line shoppers. This popularity increases the number of customer reviews which in turn adds to the site's utility. Aside from a "star rating" from 1 to 5, customers can also submit textual feedback and product accounts, made available on the 1 product's page on Amazon.com. Next to each review are three simple user-interface elements: a label, "Was this review helpful to you?", and two buttons, "Yes" and "No". It is this mechanism that allows

users to vote up or down the helpfulness of a product review. The website then allows customers to sort reviews by their voted helpfulness (the site's default review ordering) or temporally. While providing an excellent option for customers to filter out "good" and "bad" reviews, the problem with this system is the necessity of participation from review readers and the possibility that reviews that were not voted on or were authored so long ago they are not high up in the ordering of reviews. This means that helpful reviews would likely not be seen by customers unless they enumerated through a potentially very large set of other reviews.

2. Opinion classification

The analysis of feelings of opinions is a task of *clasificación* Texts combining text mining techniques and Natural language Processing (NLP), whose objective is to detect the polarity (positive, negative or neutral), of an opinion given by a certain User. Feeling analysis is not recent is not a recent development in computational sciences [3, 39], but the technique experienced a great boom after 2000 with the arrival of the Web and the popularity of forums and social networking sites. From this moment on, classification of opinions became more important because companies wanted to understand and/or study both positive and negative opinions that people were expressing about their products. Another important search was related to what customers wanted to compare before purchasing products [21].

Users generally search for web pages to consider what is discussed about the products concerned, and then extract the information that contains an opinion on the product. Immediately, a classification of opinions is carried out (positive, negative) to deduce the popularity of a given product. Within the same genre, we have the work of [36] that classifies the comments according to two categories: Recommended or Not Recommended. In [40] the authors propose the classification of opinion, according to the polarity, and the intensity of the opinion expressed.

Additionally, there are works that present applications associated with the analysis of feelings and in which it is possible to determine the polarity of certain text with different characteristics, both linguistics as statistics, and certain combinations of words [33]; taking into account the context that surrounds the text and incorporating characteristics of available lexical resources [32].

Knowing the opinion that a person has towards a product or service is a great help for taking of decisions, since it allows, among other things, potential consumers to verify the quality of the product or service before using it. This investigation is framed within the scope of opinion mining. In studies on the classification of opinions, i.e., opinions issued by users, we observe that a huge amount of work has focused on the resolution of this task by means of analysis of linguistic features [14, 17, 36, 41], and also statistics features [26, 37] or combination of both [16, 23, 29, 40].

Within the studies focused on the classification of opinions, positive or negative, is the one presented in [26], one of the first investigations on the analysis of feelings in which they are used data of movie critics found on the Web; These are used in three classification algorithms, surpassing the baselines produced manually for a human. They published their work on the classification of documents based on the sentiment expressed in these. They analyzed reviews about movies, they found that the techniques of Machine Learning improve the performance of generated baselines by the human experts. They used three algorithms of Machine Learning: Naïve Bayes (NB), Maximum Entropy (ME) and Support Vector Machines (SVM).

One of the pioneers who introduced the term Sentiment Analysis is [22]. In this paper the authors define this task as finding expressions of feelings for a given subject and determine the polarity of them. In the previous research, the analysis has made on the general polarity of a document. However, in this approach is about identifying the opinion of each subject mentioned in the text.

The objectives of [12] was to find the characteristics of the references in the criticisms, and identify the statements expressing opinions as well as their polarity and finally summarize the results. Creation of a small list of manually labeled “seed” adjectives was proposed, depending on if they expressed positive or negative feeling. Subsequently, this list was augmented using WordNet [19].

Opinions, made through a conversation or comment, it is possible to consult them through of forums, blogs or social networks, the latter being the novelty and having the highest current boom. A job based on a collection of blog entries is [7] that performs the analysis of feelings and mining opinions in these entries, showing the relevance of learning systems automatic as a resource for the detection of information of opinion.

2.1. Opinion classification in cross domain

The cross domain opinion mining problem [1, 25] focuses on the challenge of training a binary classifier from one or more domains to classify the product based sentiment data.

Current research in cross domain sentiment classification methods have focused broadly on two approaches: one is to use out-of-domain data to build supervised learning models capable of performing well on other domains. The second class of methods essentially amounts to an unsupervised approach where documents are evaluated with the assistance of pre-existing knowledge, requiring no training data. The unsupervised sentiment classifier takes into account a document’s linguistic clues, and often relies on a sentiment lexicon: a language resource that associates terms with sentiment information.

The supervised classification methods cannot perform well when training data and test data are from different domains. The main reason for this variation in performance is that training data do not have the same distribution with test data. As each domain has its own vocabulary to express sentiment, the sentiment information shared between two unrelated domains will be less. The poor performance found in the above-mentioned scenario can hence be accounted to the following two basic reasons. First, each domain has its own domain specific words and it will be different from one domain to another domain.

For instance, the word “warmhearted” frequently appears in hotel reviews, but it hardly appears in electronics reviews. Secondly, words that have high correlations with certain class labels in the training domain may not have the same degree of correlations with the same class labels in the target domain.

For instance, the word “portable” may be positive in electronics reviews, but it means nothing in hotel reviews.

3. Methodology proposed

First, we will carry out the classification of opinions in the same domain, that is when the training and test sets correspond to the same subject. This will give us a reference against which we can make comparisons while performing the classification in cross domains. This is to say that the training and test domains are different.

Table 1
Examples of corpus opinions

Class	Positive opinion	Negative opinion
DVD	This film can be interpreted any number of ways: as a farce; a tragedy; or just plain Shakespearean comedy. Brenda Blethyn has, in the past, acted in each of these genres and she brings that experience into Saving Grace with exquisite skill	Horror movies are a dime a dozen now, and nothing really speical about them anymore. Sad to say, considering its probably my faviorte genere. This movie was a disappointment from the start. Whoever did editing on this movie needs to be fired. I mean they are standing there talking, and it just cuts away when you think something is going to happen. I didnt watch the whole movie, it sucked that bad. I bought this movie used, and I'd gladly give it back to the owner free of charge.
Books	This is one of the best biographies I have ever read. You can tell the authors put a lot of time and effort into this work - it's a true labor of love. Filled with beautiful photos and extensive bibliographical notes, this one is a keeper. Who knew Miss Francis was such a "wild child"? Whether you're a film scholar or a movie buff, Lynn Kear's book deserves a special spot in your bookcase	I tried reading this book but found it so turgid and poorly written that I put it down in frustration. It reads like a translation from another language by an aca-demic bureacrat. The theme is interesting, the execution poor. Cannot recommend
Kitchens	I would certainly highly recommend this juicer to anyone! It is stylish, sleek,easy to clean, and has plenty of power to handle anything. I feel sorry for michigan, mix 3 apples and a handful of baby carrots for a great drink! Awesome! Awesome The first negative credit for this must be a lazy ol' prune who wouldn't be happy with any-thing! Get it now!	The design for using the DC adapter with LED Booklight is extremely poor. It is possible that I missed the warning that tells you to remove the batteries before using the DC adapter. But, it must not have been very big. I combined the two and ended up with batteries that leaked all over inside the battery case. These were brand new batteries. There does not seem to be anything protecting the batteries from the DC adapter and so the power goes right into them. This is bad. Modern devices are expected to work with either battery or adapter without having to remove the batteries.
Electronics	It worked much better than expected. The page clips also were very good, but one fell off about 5 minutes into owning/using the item (I put it back on, but fell off a couple more times). I also get the feeling that the other parts are fragile as well.	I was hoping for a great deal on these batteries, but like other reviewers I've found that they run out of juice incredibly fast. I'm amazed that other reviewers have had such good luck because things like my wireless keyboard which uses two AA's ran out of juice in a couple days of normal use while my automatic sprinkler time that uses two AA's lasted only a week.

The classification of documents will be performed using four machine learning techniques: Naïve Bayes (NB), Nearest Neighbors (KNN), and Support Vector Machines (SVM).

Training and test sets were constructed so that the test instances are not seen by the training instances and in this way, they tend to have a greater similarity with a real life scheme and therefore a greater degree of reliability in the system proposed.

3.1. Corpus description

The corpus used in the present work was created within the Department of computing and computer science of the University of Pennsylvania. It contains 8,000 documents in electronic format corresponding to subjective opinions of different products on the Amazon website: books, films in general (DVDs), electrical appliances and kitchen applications. Each

of these Amazon products is treated as a different domain. For each of these four categories, there are one thousand positive opinions and one thousand negative ones. This corpus is freely available and is the one mostly used in automatic classification experiments when the emotional valency of the opinion is meant to be counted.

As an example of the opinions that exist in the corpus, in the Table 1 we show an example of opinion of each class (Books, DVDs, Electronics and Kitchen).

Each review consists of a rating (0 – 5 stars), a reviewer name and location, a product name, a review title and date, and the review text. Reviews with *rating* > 3 were labeled *positive*, those with *rating* < 3 were labeled *negative*, and the rest discarded because their polarity was ambiguous.

Each labeled dataset was split into a training set of 1,600 instances and a test set of 400 instances. All the experiments use a classifier trained on the training set

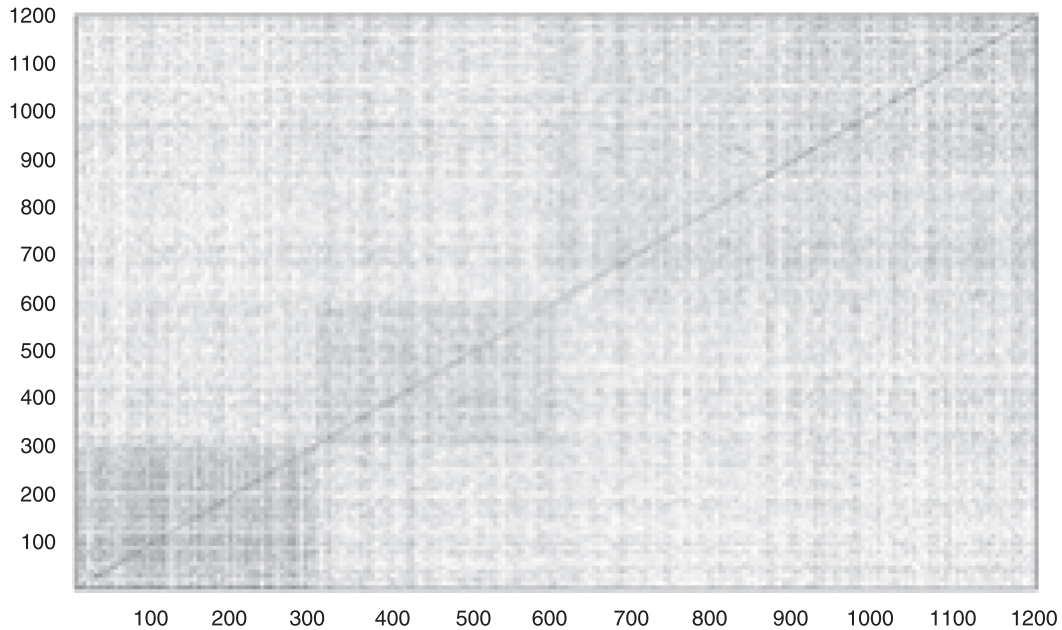


Fig. 1. Similarity chart of the four classes of the corpus.

of one domain and tested on the test set of a possibly different domain. The baseline is a linear classifier trained without adaptation, while the gold standard is an in-domain classifier trained on the same domain as it is tested.

A pre-processing of the corpus was carried out, which consisted of removing HTML tags, eliminating punctuation marks as well as empty words (also known as stopwords) which are meaningless words. The purpose of the pre-processing is to leave the opinions entering the classifier in plain text so as to avoid ascribing any possible tendency for the labels in the text. After pre-processing, the vocabulary of the training set existing in the corpus contained a total of 1,729,933 words and 5,218 different words.

In order to have an idea of the complexity of the task, Fig. 1 shows a similarity chart for the 4 different domains of the corpus under study. For the development of this graph, only 300 random files of each class were taken. It can be seen that in the following graph the axes do not have labels, instead, they have numbers ranging from 0 to 1,200 in each of them. Each of these numbers represents an opinion from the corpus used for the analysis.

In this similarity graph, we notice a high degree of overlap in the vocabulary used, but we also see the use of more frequent words in each domain (more obscure parts).

3.2. Machine learning methods

Machine learning is a field of artificial intelligence that uses statistical techniques to give computer systems the ability to learn (for example: progressively improve performance on a specific task) from data, without being explicitly programmed [30]. In order to have a perspective of the type of classifier that can best deal with the problem of analysis of opinions on four specific products (books, DVDs, kitchens, and electronics), we have selected four different types of classifiers: Naïve Bayes, Supporting Vector Machines, and K-Nearest Neighbor.

3.2.1. Classifiers employed

Naïve Bayes. Bayesian Classification represents a supervised learning method as well as a statistical method for classification. It assumes an underlying probabilistic model allowing capturing uncertainty about the model in a principled way by determining probabilities of the outcomes (see Fig. 2).

It is appropriate to solve diagnostic and predictive problems since it allows combining prior knowledge and observed data in a useful perspective for understanding and evaluating many learning algorithms. It calculates explicit probabilities for hypothesis and it is robust to noise in input data.

A Naïve-Bayes classifier is constructed by using the training data to estimate the probability of each

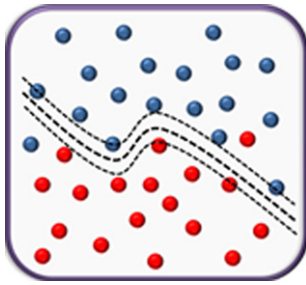


Fig. 2. Schematic of the classification principle of the Naïve-Bayes Classifier.

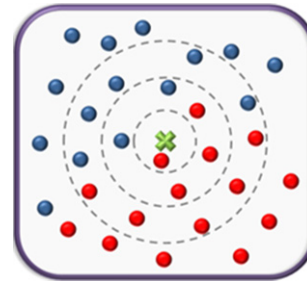


Fig. 4. Schematic of the classification principle of the K-Nearest Neighbor classifier.

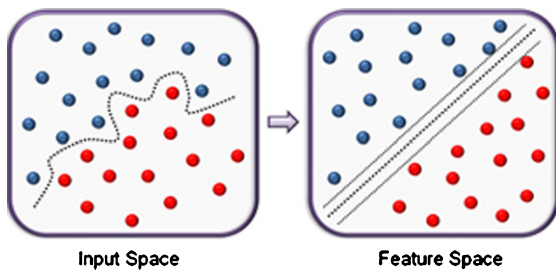


Fig. 3. Schematic of the classification principle of the Linear Support Vector Machine.

category, given the feature values of a new instance. We use Bayes theorem to estimate the probabilities [8].

Despite the fact that the assumption of conditional independence is generally not true for word appearance in documents, the Naïve-Bayes classifier is surprisingly effective.

Supporting Vector Machines. The support vector method (SVM) was introduced by Boser, Guyon and Vapnik [2] but they have only recently been gaining popularity in the learning community. In its simplest linear form, an SVM is a hyperplane that separates a set of positive examples from a set of negative examples with maximum margin, as shown in Fig. 3.

SVMs have been shown to yield good generalization performance on a wide variety of classification problems, including: handwritten character recognition [15], face detection [24], and most recently text categorization [13]. We used the simplest linear version of the SVM because it provided good classification accuracy, is fast to learn and fast for classifying new instances.

Training an SVM requires the solution of a QP (quadratic programming) problem. Any QP optimization method can be used to learn the weights, on

the basis of training examples. However, many QP methods can be very slow for large problems such as text categorization. We used a new and very fast method developed by Platt which breaks the large QP problem down into a series of small QP problems that can be solved analytically [11]. Additional improvements can be made because the training sets used for text classification are sparse and binary. Once the weights are learned, new items are classified by computing the vector of learned weights. After training the SVM, we fit a sigmoid to the output of the SVM using regularized maximum likelihood fitting.

There is no explicit error minimization involved in computing the Find Similar weights. Thus, there is no learning time so to speak, except for taking the sum of weights from positive examples of each category. Test instances are classified by comparing them to the category centroids using the Jaccard similarity measure. If the score exceeds a threshold, the item is classified as belonging to the category.

K-Nearest Neighbor. A decision tree was constructed for each category using the approach described by Chickering [4]. The decision trees were grown by recursive greedy splitting, and splits were chosen using the Bayesian posterior probability of model structure. We used a structure prior that penalized each additional parameter with probability 0.1, and derived parameter priors from a prior network as described in [4] with an equivalent sample size of 10. A class probability rather than a binary decision is retained at each node (see Fig. 4).

3.3. Evaluation

In order to evaluate the results obtained of the proposed approach on the experiments performed, we have employed precision, recall and *F*-measure,

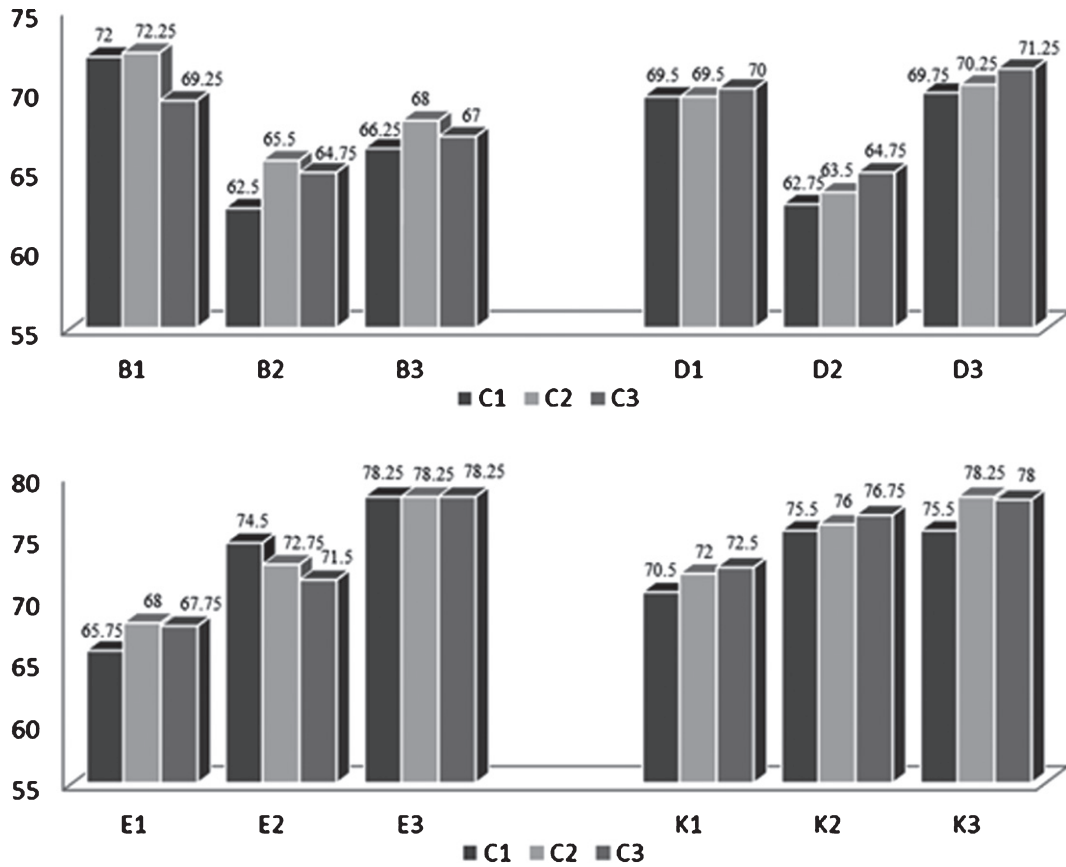


Fig. 5. Results using Naïve-Bayes.

which are standard performance measures for machine learning and are used in the classification process [34].

The metrics employed in these experiments can be formally defined in terms of True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN) values as follows:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

There exist a metric that combines precision and recall as the harmonic mean of both and it is defined as follows:

$$F - measure = 2 * \frac{Precision * Recall}{Precision + Recall}$$

4. Obtained results

Each type of product is considered as a separate domain and for each of these domains, the same amount of training data –2,000 reviews, split in a balanced way into positive and negative examples – is available.

Three classification experiments were performed for each case: Without preprocessing (classifier 1, represented by C1); Preprocessing the corpus and eliminating the stop words with a short list (classifier 2, represented by C2) and Preprocessing the corpus and eliminating the stop words with a long list (classifier 3, represented by C3).

4.1. Naïve Bayes

Figure 5 shows the results obtained from the categorization in cross domains using Naïve-Bayes.

Table 2
Gold-standard registration using Naïve-Bayes.

Group	Gold-Standard	Domain	Sub-group 1	Sub-group 2	Sub-group 3
B	73.5	Books	DVD-Books	Electronics-Books	Kitchen-Books
D	79.5	DVD	Books-DVD	Electronics-DVD	Kitchen-DVD
E	78.25	Electronics	Books-Electronics	DVD-Electronics	Kitchen-Electronics
K	82	Kitchen	Books-Kitchen	DVD-Kitchen	Electronics-Kitchen

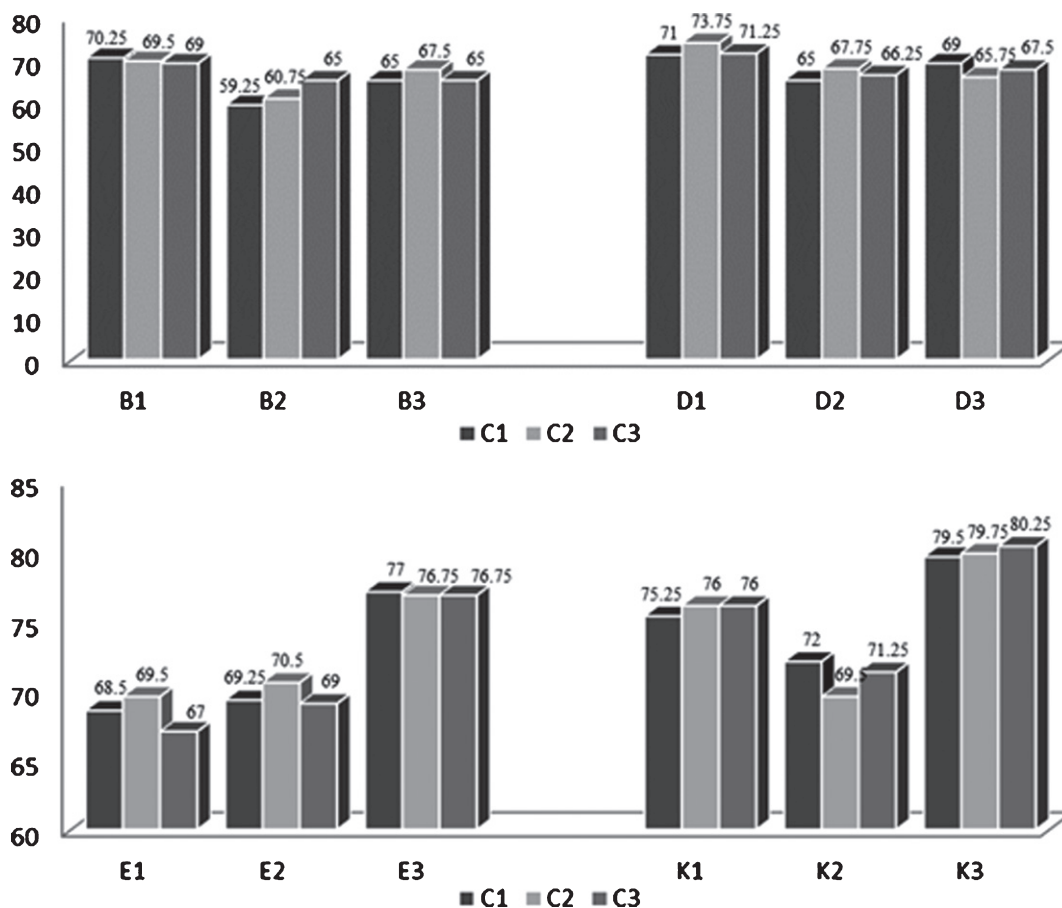


Fig. 6. Results using Support Vector Machine.

Table 2 shows the record of the values of the gold-standard of each category. In all cases, the first word corresponds to the training domain and the second word corresponds to the test domain.

4.2. Support Vector Machine

Figure 6 shows the results obtained from the categorization in crossed domains using SVM (also see table 3).

4.3. K-Nearest Neighbor

Figure 7 shows the results obtained from the categorization of all pairs of domains with the neighbor's.

Table 4 shows the record of the values of gold-standard in each categorized domain.

In Table 5 it is possible to appreciate the measurements of evaluation for the best results obtained in each categorization technique.

Figure 1 shows the gray similarity graph for the 4 different domains (B, D, E and K) of the Blitzer

Table 3
Gold-standard registration using Support Vector Machine.

Group	Gold-Standard	Domain	Sub-group 1	Sub-group 2	Sub-group 3
B	73.5	Books	DVD-Books	Electronics-Books	Kitchen-Books
D	77.25	DVD	Books-DVD	Electronics-DVD	Kitchen-DVD
E	82.5	Electronics	Books-Electronics	DVD-Electronics	Kitchen-Electronics
K	84.25	Kitchen	Books-Kitchen	DVD-Kitchen	Electronics-Kitchen

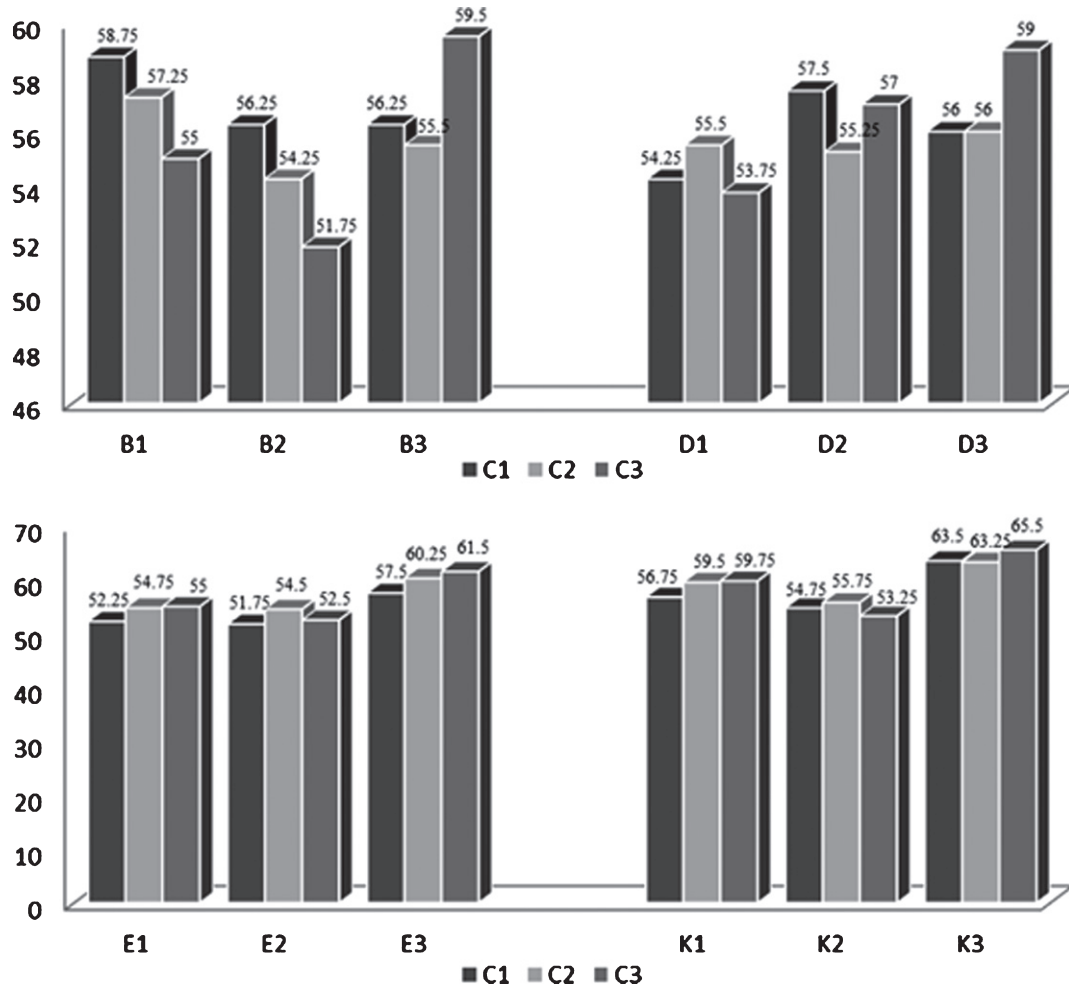


Fig. 7. Results using K-Nearest Neighbor.

Table 4
Gold-standard registration using K-Nearest Neighbor.

Group	Gold-Standard	Domain	Sub-group 1	Sub-group 2	Sub-group 3
B	61.25	Books	DVD-Books	Electronics-Books	Kitchen-Books
D	58.25	DVD	Books-DVD	Electronics-DVD	Kitchen-DVD
E	63.25	Electronics	Books-Electronics	DVD-Electronics	Kitchen-Electronics
K	66.25	Kitchen	Books-Kitchen	DVD-Kitchen	Electronics-Kitchen

Table 5
Finally results according to the evaluation measures.

Group	Sub-Group	Algorithm	Precision	Recall	Fallout	Category
B	B1	NB	0.719	0.73	0.285	Negative
B	B1	NB	0.726	0.715	0.27	Positive
E	E3	NB	0.723	0.915	0.35	Negative
E	E3	NB	0.884	0.65	0.085	Positive
B	B1	SVM	0.677	0.745	0.355	Negative
B	B1	SVM	0.717	0.645	0.255	Positive
K	K3	SVM	0.795	0.815	0.21	Negative
K	K3	SVM	0.81	0.79	0.185	Positive
D	D1	J48	0.656	0.705	0.37	Negative
D	D1	J48	0.681	0.63	0.295	Positive
K	K3	J48	0.796	0.74	0.19	Negative
K	K3	J48	0.757	0.81	0.26	Positive

corpus. For the realization of this graph, only 300 files were taken from the training set of each domain.

5. Conclusions and future work

In this article, we have presented a study about different possible strategies for building a statistical classifier for a target domain with only annotated data for one or several other domains. We presented a domain adaptation method for document-level sentiment classification.

This is a preliminary work where the objective was to consider the performance of different methods of classification combined with different preprocessing performed for the training and test sets used in the work. It can be seen from the results obtained that the best preprocessing appears in the first test and the classification method that shows better performance are SVM, followed by Naïve Bayes.

Three experiments were carried out and in each of them three learning methods were used. From figures 5, 6 and 7 we can see that the best gold standard is obtained with SVM. While, the highest precision is obtained with Naïve Bayes, followed by SVM.

A conclusion of the present work is that, with the experiments carried out, we can measure the impact of stopwords in the classification of feelings in crossed domains, by having a comparison of the classification results obtained in the different scenarios.

The most direct extension of this work is to apply the same method to other kinds of statistical classifiers.

References

- [1] J. Blitzer, M. Dredze and F. Pereira, Biographies, Bolly-wood, boom-boxes and blenders: Domain adaptation for sentiment classification, In *Proceedings of the Association for Computational Linguistics (ACL)*, 2007.
- [2] B.E. Boser, G. I. and V. Vapnik, A training algorithm for optimal margin classifiers, *Fifth ACM Workshop on Computational Learning Theory (COLT)*, 1992, pp. 144–152.
- [3] J. Carbonell, *Subjective Understanding: Computer Models of Belief Systems*, Yale, PhD thesis, 1979.
- [4] D.M. Chickering, D. Heckerman and C. Meek, A bayesian approach to learning bayesian networks with local structure, In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, UAI'97*, San Francisco, CA, USA, 1997, pp. 80–89. Morgan Kaufmann Publishers Inc.
- [5] G. Colmenares, Máquinas de vectores de soporte, *Instituto de Investigaciones Económicas y Sociales* (2009).
- [6] M. Díaz Galiano, M.R.A. and T. Martín Valdivia, Mejora de los sistemas multimodales mediante el uso de ganancia de información, *Sociedad Española para el Procesamiento de Lenguaje Natural*, 2007.
- [7] J. Fernández, E. Boldrini, J.M.G. Soriano and P. Martínez-Barco, Análisis de sentimientos y minería de opiniones: El corpus emotiblog, *Procesamiento del Lenguaje Natural* **47** (2011), 179187.
- [8] I.J. Good, *The Estimation of Probabilities: An Essay on Modern Bayesian Methods*, MIT Press, Cambridge, MA, USA, 1965.
- [9] R. Guzmán-Cabrera, M. Montes-y Gómez, P. Rosso and L. Villaseñor-Pineda, A web-based self-training approach for authorship attribution, In B. Nordström and A. Ranta, editors, *Advances in Natural Language Processing*, Berlin, Heidelberg. Springer Berlin Heidelberg, 2008, pp. 160–168.
- [10] R. Guzmán-Cabrera, RPM-y-GM Villaseñor Pineda and L., Clasificación automática de textos, *Research Computing Science*, 2012.
- [11] M. Hearst, S. Dumais, E. Osman, J. Platt and B. Schölkopf, Support vector machines, *IEEE Intelligent Systems and Their Applications* **13**(4) (1998), 18–28.
- [12] M. Hu and B. Liu, Mining and summarizing customer reviews, In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04*, New York, NY, USA, 2004, pp. 168–177. ACM.
- [13] T. Joachims, Text categorization with support vector machines: Learning with many relevant features, In C. Nédellec and C. Rouveirol, editors, *Machine Learning: ECML-98*, Berlin, Heidelberg, Springer Berlin Heidelberg, 1998, pp. 137–142.

- [14] H. Kanayama and T. Nasukawa, Fully automatic lexicon expansion for domain-oriented sentiment analysis, In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, EMNLP '06*, Stroudsburg, PA, USA, 2006, pp. 355–363. Association for Computational Linguistics.
- [15] Y. Lecun, L. Jackel, L. Bottou, A. Brunot, C. Cortes, J. Denker, H. Drucker, I. Guyon, U. Muller, E. Sackinger, P. Simard and V. Vapnik, Comparison of learning algorithms for handwritten digit recognition. In F. Fogelman and P. Gallinari, editors, *International Conference on Artificial Neural Networks Paris*, 1995, pp. 53–60. EC2 & Cie.
- [16] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu and H.W. Lauw, Detecting product review spammers using rating behaviors, In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM '10*, New York, NY, USA, 2010, pages 939–948. ACM.
- [17] B. Liu, M. Hu and J. Cheng, Opinion observer: Analyzing and comparing opinions on the web, In *Proceedings of the 14th International Conference on World Wide Web, WWW '05*, New York, NY, USA, 2005, pp. 342–351. ACM.
- [18] F. Martín Mateos and J. Ruiz Reina, *Inteligencia Artificial II, Procesamiento del lenguaje natural*, 2011.
- [19] G.A. Miller, R. Beckwith, C. Fellbaum, D. Gross and K. Miller, Wordnet: An on-line lexical database, *International Journal of Lexicography* **3** (1990), 235244.
- [20] T. Mitchell, *Machine Learning*, McGraw Hill, 1996.
- [21] S. Morinaga, K. Yamanishi, K. Tateishi and T. Fukushima, Mining product reputations on the web, In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '02*, New York, NY, USA, 2002, pp. 341–349. ACM.
- [22] T. Nasukawa and J. Yi, Sentiment analysis: Capturing favorability using natural language processing, 2003.
- [23] K. Nigam and M. Hurst, Towards a robust metric of polarity, In *Computing Attitude and Affect in Text: Theories and Applications*, 2006.
- [24] E. Osuna, R. Freund and F. Girosi, Training support vector machines: An application to face detection, 1997, pp. 130–136.
- [25] S.J. Pan, X. Ni, J.-T. Sun, Q. Yang and Z. Chen, Cross-domain sentiment classification via spectral feature alignment, In *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, New York, NY, USA, 2010, pp. 751–760. ACM.
- [26] B. Pang, L. Lee and S. Vaithyanathan, Thumbs up?: Sentiment classification using machine learning techniques, In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP '02*, Stroudsburg, PA, USA, 2002, pp. 79–86. Association for Computational Linguistics.
- [27] C. Pérez and D. Santín, *Data Mining, Soluciones con Enterprise Miner*, Alfaomega grupo editor, 2006.
- [28] P. Procházková, *Fundamentos de la lingüística de corpus, Concepción de los corpus y métodos de investigación con corpus*, Alfaomega grupo editor, 2006.
- [29] E. Riloff and J. Wiebe, Learning extraction patterns for subjective expressions, In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, EMNLP '03*, Stroudsburg, PA, USA, 2003, pp. 105–112. Association for Computational Linguistics.
- [30] A. Samuel, Some studies in machine learning using the game of checkers, *Journal for Literary and Linguistic Computing* **3** (1983), 210229.
- [31] B.P. Sánchez and D. Pinto, Identification of verbal phraseological units in mexican news stories, *Computación y Sistemas* **19**(4) (2015).
- [32] B.P. Sánchez and D. Pinto, Idiom polarity identification using contextual information, *Computación y Sistemas* **22**(1) (2018).
- [33] B.P. Sánchez, D. Pinto and S. Mejri, Evaluating polarity for verbal phraseological units, In *Human-Inspired Computing and Its Applications - 13th Mexican International Conference on Artificial Intelligence, MICAI 2014*, Tuxtla Gutiérrez, Mexico, Proceedings, Part I, 2014, pp. 191–200.
- [34] B.P. Sánchez, D. Pinto and V.K. Singh, Compositionality versus non-compositionality verification based on lexical domain for verbal phraseological units, *Journal of Intelligent and Fuzzy Systems* **34**(5) (2018), 3059–3067.
- [35] F. Su and K. Markert, Subjectivity recognition on word senses via semi-supervised mincuts, In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, NAACL '09*, Stroudsburg, PA, USA, 2009, pp. 1–9. Association for Computational Linguistics.
- [36] P.D. Turney, 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*, Stroudsburg, PA, USA, 2002, pp. 417–424. Association for Computational Linguistics.
- [37] P.D. Turney and M.L. Littman, Measuring praise and criticism: Inference of semantic orientation from association, *ACM Trans Inf Syst* **21**(4) (2003), 315–346.
- [38] F. Vallejos, *Minería de datos, Alfaomega grupo editor* (2006).
- [39] Y. Wilks and J. Bien, Beliefs, points of view and multiple environments, In *Proceedings of the International NATO Symposium on Artificial and Human Intelligence*, New York, NY, USA, 1984, pp. 147–171. Elsevier North-Holland, Inc.
- [40] T. Wilson, J. Wiebe and R. Hwa, Just how mad are you? finding strong and weak opinion clauses, In *Proceedings of AAAI*, 2004, pp. 761–769.
- [41] H. Yu and V. Hatzivassiloglou, Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences, In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, EMNLP '03*, Stroudsburg, PA, USA, 2003, pp. 129–136. Association for Computational Linguistics.