A tool for the visualisation of public opinion

Konstantinos Soulis

Informatics Development Agency,
Ministry of Administrative Reform & Electronic Government,
2-4 Filoxenou & Spintharou,
GR11636 Athens, Greece
E-mail: ksoulister@gmail.com

Iraklis Varlamis*

Department of Informatics and Telematics,
Harokopio University of Athens,
70, El. Venizelou,
GR17671 Athens, Greece
E-mail: varlamis@hua.gr
Website: http://www.dit.hua.gr/~varlamis/
*Corresponding author

Andreas Giannakoulopoulos

Department of Audio and Visual Arts,
Ionian University,
GR49100 Corfu, Greece
E-mail: agiannak@ionio.gr
Website: http://www.ionio.gr/~agiannak

Filippos Charatsev

Department of Informatics and Telematics,
Harokopio University of Athens,
70, El. Venizelou, GR17671 Athens, Greece
E-mail: it20838@hua.gr

Abstract: The rise in popularity of the web and social media has significantly changed the way voters communicate and form their opinions. National governments are also affected by the hype of social media, so they launch new debate tools and open social platforms where citizens are able to communicate, collaborate and exchange opinions. When the amount of opinions increases, then it becomes difficult to process and interpret them manually. In this case, opinion-mining techniques and information visualisation tools can be employed to depict the public opinion and give comprehensive visual summaries. In this work, we present an information visualisation tool for surveys, which allows users to select from a variety of graphs, drill down to selected periods or roll-up to a larger scale and supports input from both closed-end and open-end questions. In the latter case, the tool employs opinion-mining techniques to quantify voter's opinion.

Keywords: open-source tools; opinion mining; opinion visualisation; public opinion mining; voting advice applications.

Reference to this paper should be made as follows: Soulis, K., Varlamis, I., Giannakoulopoulos, A. and Charatsev, F. (2013) 'A tool for the visualisation of public opinion', *Int. J. Electronic Governance*, Vol. 6, No. 3, pp.218–231.

Biographical notes: Konstantinos Soulis is an employee of the Informatics Department Agency in the Ministry of Administrative Reform and Electronic Governance. He is a graduate of the National School of Public Administration, Department of Information System Management. He is also a Postgraduate Student in Informatics of Greek Open University. He has a degree in Physics from the University of Ioannina. Also, he has attended summer school classes at the research centre 'Demokritos' (2000) and at the Faculty of Philology in St. Petersburg State University (2006).

Iraklis Varlamis is an Assistant Professor at the Department of Informatics and Telematics of Harokopio University of Athens. He received his PhD in Computer Science from Athens University of Economics and Business, Greece. His research interests vary from data mining and the use of semantics in web mining to virtual communities and their applications. He has published several articles in international journals and conferences, concerning web document clustering, the use of semantics in web link analysis and web usage mining, word sense disambiguation using thesauruses, etc.

Andreas Giannakoulopoulos is a Lecturer at the Department of Audio and Visual Arts of the Ionian University. He holds a BA (Ptychio) in Economics from the University of Athens (UoA), a BA (Ptychio) in Communication and Media Studies from UoA, a Master of Arts in Communication and Media Studies from UoA, and a Master of Science in Logic from the University of Amsterdam. His doctoral dissertation, approved by the University of Athens, was in the field of web accessibility technologies. He has professional working experience in various mass-media, including magazines, radio, television and the web. The main field of his academic activities is Computer Mediated Communication and especially the web development technologies. His research interests focus on information architecture, web-based media, content management platforms and e-learning systems as means of effective communication via the web.

Filippos Charatsev is a Postgraduate Student at the MSc in Informatics and Telematics of Harokopio University of Athens. He received his degree from the same department. His research interests comprise text mining, information graphics and visualisation, and he is currently working on projects concerning the visualisation of information extracted from the news.

1 Introduction

The aim of applications that summarise and visualise public opinion is bi-fold: first to assist voters to find a political stance that stands closer to their preferences and beliefs and second to help governmental bodies to have an overview of what citizens think and believe. Concerning Voting Advice, it is really quite hard to find a rule of thumb for deciding whether an application is indeed a voting advice application (VAA), since the title is quite self-defined. The related survey work of Cedroni and Garzia (2010) reveals that there are several VAAs that follow different procedures to analyse peoples' opinions and provide advice. Concerning Public Opinion Mining, 'listening' to public's opinion can be considered as a first step towards e-participation (Stylios et al., 2010) and is currently one of the hottest research areas on information and communication technologies (ICT) for Governance and Policy Modelling (Lampathaki et al., 2010).

Although the technological solutions may vary, there is a common ground for the two disciplines, which includes politics and the participating stakeholders, namely the voters and the politicians. Every tool is created according to the needs of the participants in a political debate. Typically, VAAs attempt to capture citizens' opinion and create their profile through a predefined set of closed-ended questions. Although questionnaires with closed-ended questions are easy to create, fill and process, they introduce a bias to the citizens' answer and thus affect the formulation of the public opinion. On the other side, public political surveys or open government discussion tools, which allow for comments and free text answers, avoid this bias but are difficult to process. Opinion-mining tools are useful in this direction, since they are able to extract sentiment and opinion from text and thus quantify citizens' answers.

In this work, we present an application, which combines in tandem opinion mining and information visualisation techniques, and offers an interactive environment for the visualisation of public opinion. The application uses time-related charts to allow users to understand how opinions and stances evolve during an ongoing public consultation or survey. We follow the typical steps of Knowledge Discovery in Data (Fayyad et al., 1996) (data selection, pre-processing, transformation, data mining and interpretation/ evaluation) to analyse and visualise public opinion.

The contributions of this work comprise:

- The integration of opinion-mining techniques and information visualisation tools in a single public opinion-mining platform.
- The application of this platform in two consultations. The consultations span different periods, use two different languages (Greek and English), tackle two different topics (racism and recycling) and employ both open- and closed-ended questions.
- A simple yet flexible structure for providing input data to the application, which allows easy integration with more public opinion data sets.

In the following section, we provide an overview of existing solutions in public opinion mining and visualisation. In Section 3, we briefly explain how the various data-mining tools and information visualisation tools are combined in a single application and in Section 4 we provide empirical results from the two consultations. In Section 5, we evaluate the accuracy of our opinion-mining methods and, finally, in Section 6 we provide the conclusions of this work.

2 Related work

Opinion mining and sentiment analysis are the key techniques employed by public opinion-mining applications to summarise large data sets comprising comments or free text opinions. Such applications are mainly operated by governmental bodies that wish to have comprehensive quantitative summaries that depict the number of positive and negative opinions, the common sentiment against politicians or political acts, etc.

Grimmer and Stewart (2013) offer a great survey of methods for analysing political texts and extracting sentiment and opinions. Research works in opinion mining use:

- unsupervised solutions, which combine large corpora and probabilistic models (Efron, 2004) for defining the cultural orientation (i.e., to distinguish between left and right political convictions)
- supervised solutions, which are based on a pre-categorised 'reference' set of training samples (Laver et al., 2003)
- semi-supervised solutions, which assume the polarity of training samples using related information, such as comments and votes (Goldberg et al., 2007).

They usually classify samples into positive or negative (Thomas et al., 2006; Bansal et al., 2008; Mullen and Malouf, 2006) but in some cases the classification problem contains more than two classes (Hopkins and King, 2010). Our approach can be either supervised or semi-supervised and can support more than two class values.

Advanced opinion-mining systems, such as UbiPOL (Irani et al., 2010), or the system presented by Kwon et al. (2006) employ ontological knowledge to first identify the topic or subtopic of each opinion and consequently to find its polarity. However, the cost of creating domain-specific ontologies makes such approaches difficult to scale to additional domains and languages.

On the other side, VAAs combine technological novelties with political procedures (Louwerse and Rosema, 2011) to assist voters to align with predefined opinions or with other voters who share similar beliefs. There are quite numerous examples concerning VAA implementations in the world (Cedroni and Garzia, 2010; Sasaki et al., 2010). All of them follow a basic voters' classification process, which comprises a learning step, where representatives from different political parties or opinions answer a set of questions and train an opinion classifier and a real-time step where each voter answers the same set of questions and is assigned to the most similar predefined profile(s). The tools differ from one another in the details, in the interface they use and the features they provide to the user. *Smartvote* in Switzerland (Ladner et al., 2008), *Wahl-O-Mat* in Germany (Cedroni and Garzia, 2010, p.65), *HelpMeVote* in Greece (Andreadis and Chadjipadelis, 2011), *Do the Vote Test* in Belgium (Walgrave et al., 2006), *wahlkabine.at* in Austria (Cedroni and Garzia, 2010, p.174), *KohoVolit*[1] in Slovakia and the Czech Republic or *EU Profiler* across EU member states (Trechsel and Mair, 2009) are a few only VAAs that aim in assisting voters to make a decision in the elections.

The first evidences from the use of Voting Advice and Public Opinion Mining applications in elections and public consultations show that such applications have enhanced political disputes and awareness and have increased people's participation (Ladner and Pianzola, 2010). Another interesting finding is that the use of such applications influences the final decision of people, especially those who have no predefined decisions in mind (Andreadis and Chadjipadelis, 2011).

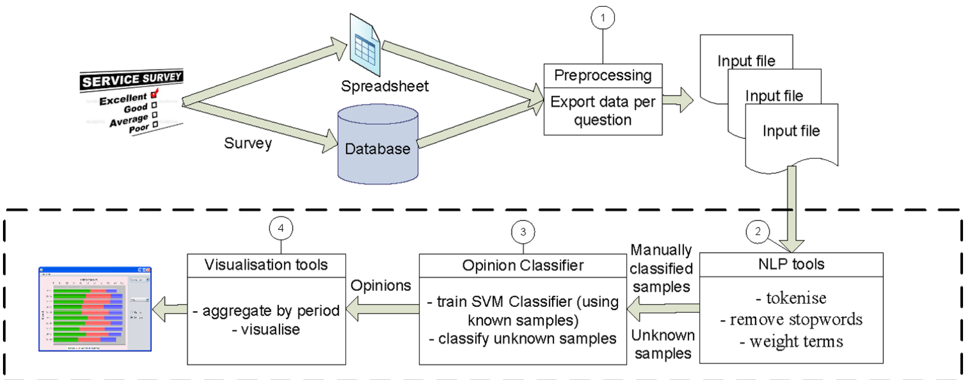
In both applications, voters are called to state their opinion on several political statements, using two or more predefined opinion alternatives (i.e., 'I strongly agree', 'I agree', 'neutral', etc.) or even free text argumentation. Voting advice tools estimate the distance between the voter and every party, showing explicitly the one closest to him/her and estimates the extent of agreement with the remaining parties. Opinion-mining tools process the free text argumentation to extract opinions, emotions or sentiment or simpler to categorise it to the respective opinion category. In many cases, the tools use different time-slots that span a period before and after the decision-making (e.g., elections) and comparatively present the results of public opinion.

The driving force behind these applications is data mining and the information visualisation techniques are their showcase. Although they choose from a variety of visualisations ranging from simple histograms, pies or plots (Trechsel and Mair, 2009) to graph-based argument maps (Kourmpanis and Peristeras, 2010), in the majority of cases each application uses a single visualisation. In contrast to them, our application supports multiple visualisations and time detail levels, thus constituting a more flexible solution. Most of the aforementioned tools use closed-ended questions with predefined or even scaled answers and rarely make use of the free text provided by the voters. Our system supports both scaled answers and free text and uses text-mining techniques to detect the polarity of an opinion.

3 Implementation

The proposed application incorporates several open-source data mining and visualisation tools in its process. The tools, as depicted in steps 2–4 of Figure 1, have been customised to better fit the needs of our application: content in multiple languages, classification of opinions in predefined categories and visualisation of evolution over time.

Figure 1 The flow of information, from the survey to the final visualisation (see online version for colours)



The first step of the process refers to feeding the survey data to the application. Surveys comprise many questions of different types and produce data that are afterwards processed to meet the format imposed by the statistical tool used in each research. Since there is no template or standard for survey data, the conversion cannot be avoided. To simplify the process, we adopt a simple schema for our input files, which assumes that every input file contains the answers to a single survey question. The input file contains the data part, which corresponds to the replies given to the survey question and the metadata that describe data. Each reply may contain a closed-ended question, with predefined answers and the respective open-ended question, with free text answers. Replies can optionally contain date stamps. Figure 2 depicts a sample input file from a survey on the use of Plastic Carrier Bags, where date information is not available in each reply, but answers have been grouped together by the order of appearance and the grouping information is available as metadata (see the @Period property). The metadata part contains information on the survey question title, the language and encoding used for the answers, the number of answers, which will be used to train the opinion-mining algorithm, the periods that data span (in different levels of detail) or the date format when a date stamp is used in each reply, the different user-defined opinion levels and their reference to positive and negative polarity values.

Figure 2 Sample input file fragment

```
@Title:Plastic Carrier Bags
@Language:english
@Encoding:UTF-8
@Training:1-500
@Period:Day:501-1000,1001-2000,2001-3000,3001-4000,4001-5000,5001-6000,6001-7000,7001-
@Period:Month:501-3000,3001-5000,5001-7500,7501-10000,10001-12000,12001-14000,14001-1
@Period:Season:501-5000,5001-10000, 10001-15550
@OpinionLevel:Strongly disagree, Disagree, No opinion, Agree, Strongly agree
@Polarity:Plus:Agree, Strongly agree
@Polarity:Minus:Strongly disagree, Disagree


Strongly agree  In spite of all well known figures: We owe it to ourselves and to the
Disagree  The European policy concerning the problem in focus should be adopted to the
Agree Plastic is not a stuff for the future, because it´s going to destroyed themself
No opinion  Important are not only the carrier plastic bags. At least as much damaging
Strongly disagree Regarding climate heating effect is a well recycled plastic bag much
Strongly agree  private opinion: Oil is too valuable to burn it
Strongly agree  1) Plastic bags are oil-products ---> oil becomes soon rare according
Strongly agree  Plastic bags are wasting the precious ressources of oil, so the ban o
Strongly agree  People have to go back to nature. Where are we going to, if we don`t
```

The second step refers to the pre-processing of data using linguistic analysis, which comprises to kenisation, stopword removal, frequency-based term weighting, etc.

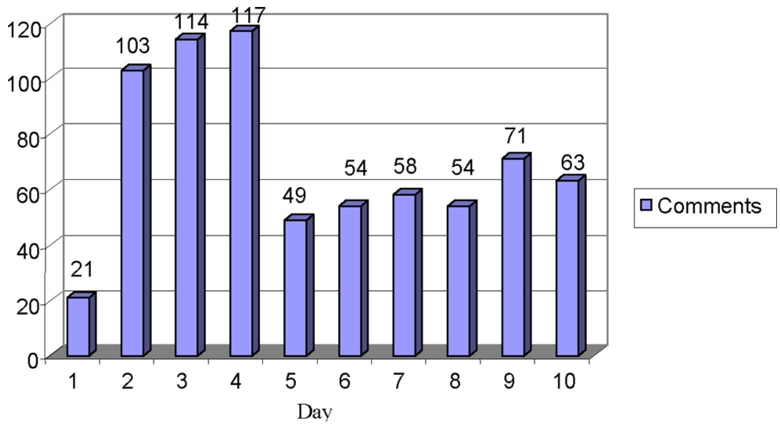
The third step is the application of the data-mining algorithm. Our application employs Weka data-mining suite,[2] which is a collection of machine-learning algorithms for data-mining tasks written in Java. More specifically, it uses the LibSVM classification algorithm, which is an implementation of Support Vector Machines algorithm (Chih-Chung and Chih-Jen, 2011). The classification algorithm uses the first entries of the input file (the number is defined by the @Training metadata option) for training and then applies the classifier to the remaining entries. As a result, all the free text answers are assigned an Opinion Level value (e.g., agree, disagree, etc.).

The last step of our process refers to the visualisation of the opinion-mining results. Among the various visualisation techniques and methods, presented in Soulis (2011), we selected and incorporated simple visualisations, which are easily interpreted by humans. The following section demonstrates some of the visualisations and the results in the two consultations that we processed.

4 Demonstration

The www.opengov.gr website in Greece is the first application of our tool. The site serves the online consultation organised by the government during the rule-making process, and offers to citizens the possibility to participate in the assessment of every article of a draft. Each citizen is able to choose the draft of his or her interest and comment on an article using free text. When the consultation is over, a report is created, which points out the most important issues stressed out by the participants. The report is written in plain text form by the consultation operator. We chose the consultation on "*Combating certain forms and expressions of racism and xenophobia by means of Penal Law*", organised by the Ministry of Justice, which contained 757 textual comments that span a 10-day period as shown in Figure 3. Each column in the bar chart of Figure 3 corresponds to a different day and the number of comments for a day is written above the respective bar.

Figure 3 Comment distribution (see online version for colours)



Our training sample consisted of 53 training samples (20 negative, 17 positive and 16 neutral) randomly chosen from the consultation data set. A snapshot of the data set is given in Figure 4. Both test and training samples have been pre-processed using stemming and stopwords' removal to better capture the morphological structure of the Greek language.

The first visualisation, depicted in Figure 5, is a stacked bar chart, which demonstrates the ratio of positive, negative and neutral opinions per day. JFree Chart API[3] is the open-source library that we incorporated in this step. The application programming interface (API) offers a wide range of charts, which can be employed for the visualisation of results.

The second example is a consultation held by the European Commission[4] for the citizens of the EU, who are called to express their opinion in free text and in a five-grade scale (strongly agree, agree, neutral, disagree and strongly disagree) in several questions. This scale is extremely helpful for classification purposes, since there is not any plain text that should be analysed and interpreted. The chosen consultation concerned environmental issues and the total set of answers were given in a spreadsheet.

The participation in this consultation is significantly bigger, with more than 15,000 replies. Although the number is not huge at European level, it is more than adequate for data-mining purposes. In the pre-processing phase, from the original spreadsheet we exported the user replies on the question "Reduce use of plastic carrier bags" to a tab-delimited format. Consequently, we added the metadata shown in Figure 2 and fed the answers to our application. In this example, we have chosen a spider web (or radar plot) visualisation, and a comparative tag cloud visualisation for positive and negative opinion. The radar plot, see Figure 6, allows the mapping of multivariate data in the form of a two-dimensional chart of three or more quantitative variables represented on axes starting from the same point. The results for the selected question, which are depicted in Figure 6, are based on the explicitly expressed user opinions (ratings) and not on free text answers and summarise answers in three custom seasons.

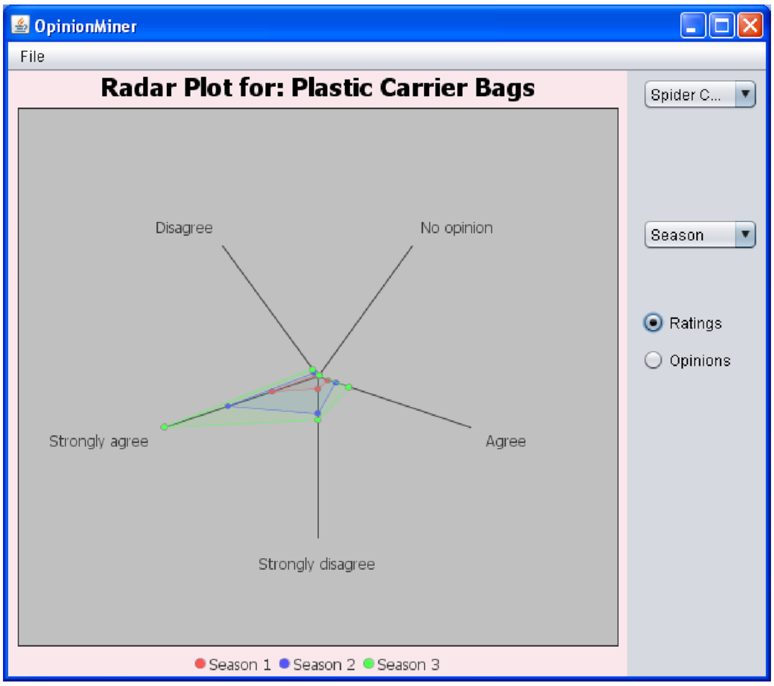
Figure 4 Input file fragment from the opengov.gr consultation (see online version for colours)

```
@Title:Penalise racism
@Language:greek
@Encoding:UTF-8
@Training:1-53
@DateFormat:DD/MM/YYYY
@OpinionLevel:Positive, Neutral, Negative
@Polarity:Plus:Positive
@Polarity:Minus:Negative

22/2/2011 Positive  "Ο όρος ""γενετήσιο προσανατολισμό"" πρέπει να αντικαταστα
αναχρονιστικός και ο έχει την έννοια και της γέννησης και έτσι εισάγει έμμεσα δ
και η Διεθνής Αμνηστία. http://www.amnesty.org.gr/right-to-sexual-identity Απο
εκφράστηκε στη Συνάντηση με τους Νέου τον περασμένο Νοέμβριο) είναι ξεκάθαρα εν
Νόμος αυτός να είναι σύγχρονος και να χρησιμοποιεί ορολογίες του 2011 που δεν α
22/2/2011 Positive  Επίσης στην 1η παραγραφό πρέπει να προστεθεί και η Ταυτότητ
...
...
...
22/2/2011 ? "Δεν διαθέτω νομικές γνώσεις και ενδεχομένως η φράση ""ο δράστης να
στα δικά μου μάτια, αλλά θα θέσω τις απορίες που μου δημιουργεί.  α)Αν η πράξη
παρών οποιαδήποτε χρονική στιγμή επιστρέψει στην Ελλάδα;  β)Αν ο δράστης είναι
θεωρείται ότι είναι σωματικώς παρών οποιαδήποτε χρονική στιγμή επισκεφθεί την Ε
```

Figure 5 Stacked bar chart visualisation (see online version for colours)

Figure 6 Radar plot for the 'plastic carrier bags' (see online version for colours)



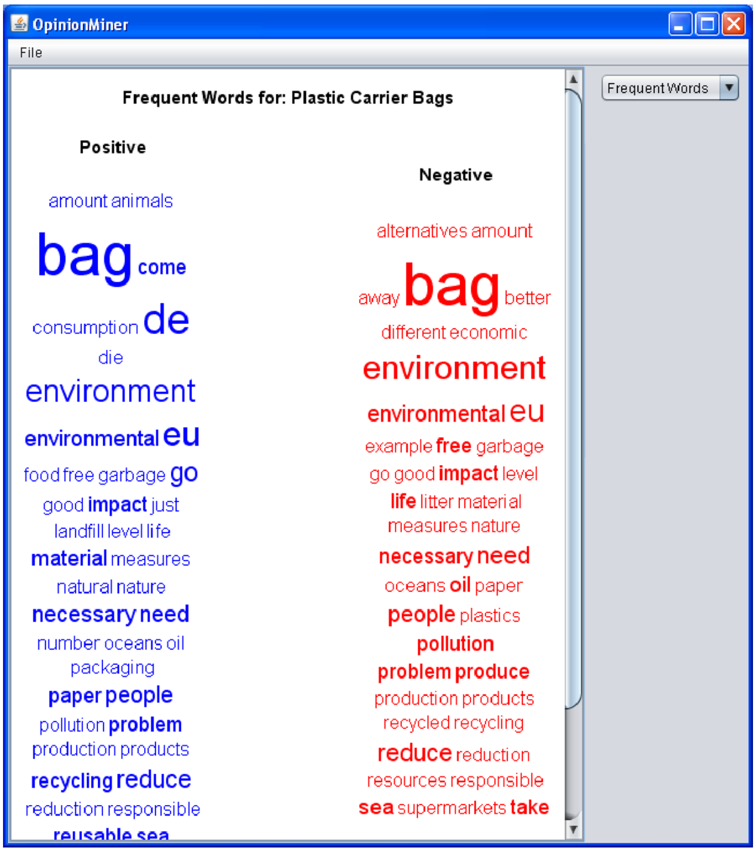
These kinds of plots offer a clear and immediate view of the information given by the participants' comments only in this kind of situations, where the different possible answers are strictly formulated (strongly agree, agree, neutral, disagree, strongly disagree) and not a plain text. Furthermore, their number must be limited (in this case 5) or else it will be very hard to read (like in a case of 10 different axes in the plot).

A far more interesting analysis is performed using the qualitative tools. In them, the arguments of the participants from different orientations are presented in a simple and clear way without having to study all the comments. OpenCloud[5] is a visualisation that complies with this specific prerequisite. It is open source written in Java. In Figure 7, we use two 'word cloud' visualisations to present in contrast to the frequent words in positive and negative comments. We can see that the two clouds have many words in common. However, words such as recycling and reusable are more common in positive comments, whereas pollution appears more frequently in negative comments.

The role of all the aforementioned visualisations is bi-fold. First, they provide useful summarisations of the public opinion, and help governmental bodies and other decision-makers and promoters to take critical decisions and be aware of the reaction of the public. Second, they help voters to understand what others believe by automatically analysing content and processing large amount of qualitative data and take position against or in favour of an issue using the output of text visualisation techniques.

Depending on the domain, the need for summarisation and the information available in the first level (i.e., before mining), different types of visualisations can be more or less useful. The aim of this work was to develop a prototype application with an initial set of interactive visualisations that combine text and opinion polarity, which will be further enhanced with more visualisations.

Figure 7 Frequent words chart, for positive and negative comments (see online version for colours)



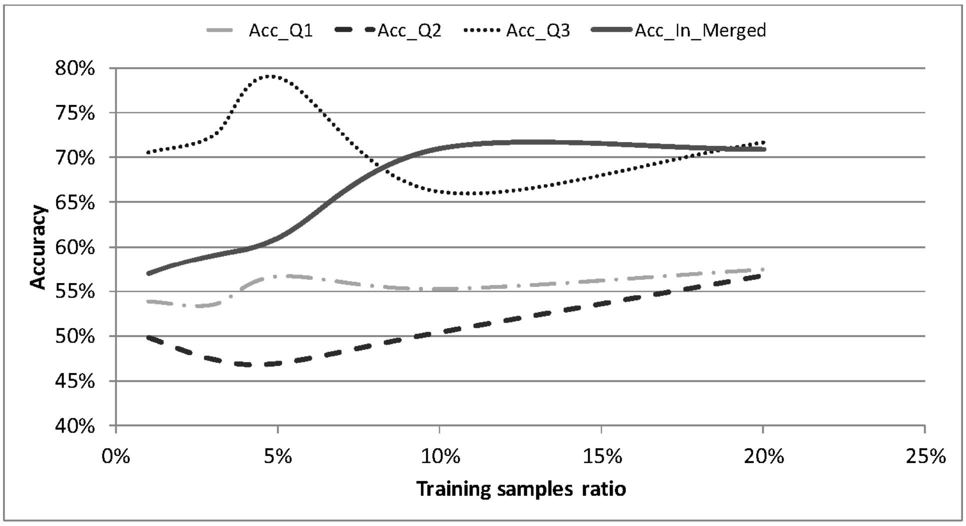
5 Evaluation

A research question that arises from the visualisation of public sentiment as expressed with text (Grimmer and Stewart, 2013) concerns the quality of the text-mining process and the correlation of text-based opinions with numeric polling data (Cummings et al., 2011; O'Connor et al., 2010). To evaluate the accuracy of our opinion-mining methods, we use the carrier bags data set as our benchmark, and more specifically three questions that contain both explicit opinions (Agree, Disagree, etc.) and textual justification of the opinion. For the evaluation, we ignore answers that omit the textual justification, since there is no need for opinion mining from text. This pre-processing results in three sets of <*rating,comment*> pairs, which comprise 3884, 2853 and 1326 pairs, respectively. All the sets use the same four rating values (i.e., Strongly disagree, Disagree, Agree, Strongly agree). The 'No opinion' rating does not appear in the sets, since it was never accompanied by a comment. The accuracy in the three sets, when a varying ratio of pairs is used for training, is depicted in Figure 8.

The dashed lines correspond to the three questions, whereas the thick black line depicts the accuracy in a merged corpus that comprises all answers in the three questions.

Accuracy scores at around 70% are very promising, especially when we have four different classes (ratings). Another interesting finding, which needs further evaluation, is that we need approximately 5% of the data set for training, which corresponds to 50–120 comments.

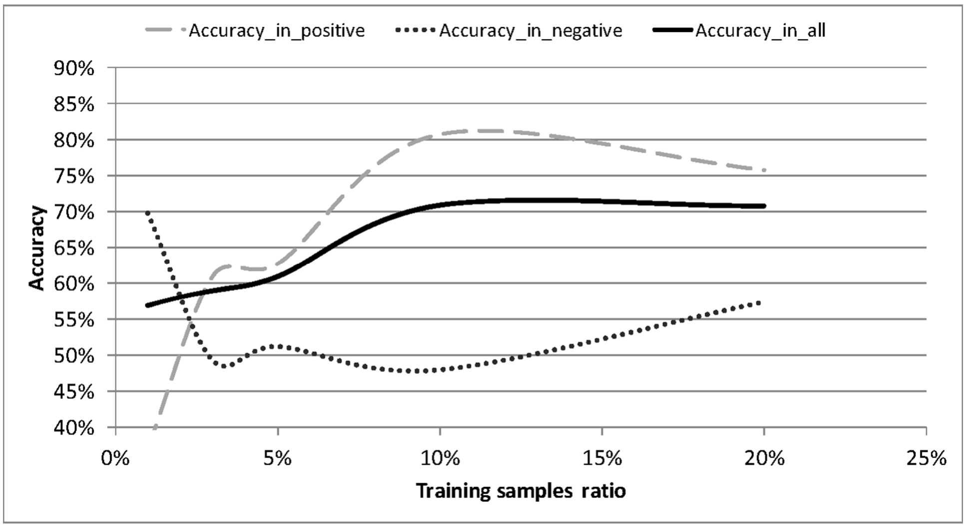
Figure 8 Accuracy of the opinion mining process in different datasets



To study the bias from the imbalanced polarity of the data set (positive opinions are double than negative ones), we depict accuracy separately on positive and negative cases by merging Strongly Agree with Agree and Strongly Disagree with Disagree.

The resulting scores (Figure 9) show a raise and higher accuracy values for positive comments when the training sample size reaches 5% of the data set. This results in a raise to the overall accuracy for all comments.

Figure 9 Accuracy of the opinion mining process in positive and negative ratings



6 Conclusions

As it is obvious throughout this paper, in the last 20 years a huge amount of effort has been put into combining technology with political procedures. This combination can be tackled at the level of legislation or regulatory efforts, through consultation in different policies, as well as at different levels of electoral process and this is where opinion mining and visualisation tools can be of assistance. However, if the creation and the assessment of various tools around the world remains in the plane of technological feasibility and ease of use, then the overall impact they might have in the political plane, in fields like e-democracy and e-participation, will most likely be decreased and consequently devalued.

When the tools are treated like indicators of political orientation, they can really offer an added value in these procedures and gain the trust of the citizens. In such a case, all the competing parties are 'exposed'. This means that they are not only obligated to answer the whole set of questions/statements but also to answer in a simple and clear way, being deprived of the right to ambiguous and blur answers. In this way, the voters can have a really explicit picture of the parties' views. In order, for this visual presentation, to be better served, different kinds of visualisations can be used, such as those presented and proposed in this paper.

References

Andreadis, I. and Chadjipadelis, Th. (2011) 'Voting advice applications and their impact on elections', *Proceedings of the 61st Political Studies Association Annual Conference 'Transforming Politics: New Synergies'*, London, UK, Available online at: http://www.psa.ac.uk/journals/pdf/5/2011/355_216.pdf

Bansal, M., Cardie, C. and Lee, L. (2008) 'The power of negative thinking: Exploiting label disagreement in the min-cut classification framework', *Proceedings of the International Conference on Computational Linguistics*, Manchester, UK.

Cedroni, L. and Garzia, D. (Eds.) (2010) *Voting Advice Applications in Europe: The State of the Art*, ScriptaWeb, Napoli, Italy.

Chih-Chung, C. and Chih-Jen, L. (2011) 'LIBSVM: a library for support vector machines', *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, No. 3, Article 27, May, pp.1–27.

Cummings, D., Haruki O. and Ningxuan, W. (2011) *Who Needs Polls? Gauging Public Opinion from Twitter Data. Stanford NLP Course, Final Project Report*, Available at: http://www-nlp.stanford.edu/courses/cs224n/2011/reports/ nwang6-davidjc-harukioh.pdf

Efron, M. (2004) 'Cultural orientation: classifying subjective documents by cociation [sic] analysis', *Proceedings of the AAAI Fall Symposium on Style and Meaning in Language, Art, Music, and Design*, Arlington, Virginia, pp.41–48.

Fayyad, U.M., Piatetsky-Shapiro, G. and Smyth, P. (1996) 'Advances in knowledge discovery and data mining', in Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R. (Eds.): *American Association for Artificial Intelligence*, Menlo Park, CA, USA, pp.1–34.

Goldberg, A.B., Zhu, J. and Wright, S (2007) 'Dissimilarity in graph-based semi-supervised classification', *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*, 21–24 March, 2007, San Juan, Puerto Rico, *Journal of Machine Learning Research – Proceedings series*, Vol. 2, pp.155–162.

Grimmer, J. and Stewart, B.M. (2013) 'Text as data: the promise and pitfalls of automatic content analysis methods for political texts', *Political Analysis*, Oxford University Press, Vol. 21, No. 3, Summer, pp.267–297.

Hopkins, D. and King, G. (2010) 'A method of automated nonparametric content analysis for social science', *American Journal of Political Science*, Vol. 54, No. 1, pp.229–247.

Irani, Z., Lee, H., Weerakkody, V., Kamal, M. and Topham, S. (2010) 'Ubiquitous Participation Platform for POLicy Makings (UbiPOL): a research note', *International Journal of Electronic Government Research*, IGI Publishing, Vol. 6, No. 1, January–March, pp.78–106.

Kourmpanis, V. and Peristeras, V. (2010) 'E-consultation: processing and summarizing contributions with argument maps', in Davies, J. and Janowski, T. (Eds.): *Proceedings of the 4th International Conference on Theory and Practice of Electronic Governance (ICEGOV '10)*, ACM, New York, NY, USA, pp.357–358.

Kwon, N., Shulman, S. and Hovy, E. (2006) 'Multidimensional text analysis for eRulemaking', *Proceedings of the 2006 International Conference on Digital Government Research (dg.o '06)*, Digital Government Society of North America, pp.157–166.

Ladner, A. and Pianzola, J. (2010) *Do Voting Advice Applications Have an Effect on Electoral Participation and Voter Turnout? Evidence from the 2007 Swiss Federal Elections*, Institutional Policies, IDHEAP Lausanne, France.

Ladner, A., Felder, G. and Schädel, L. (2008) 'From e-voting to smart-voting–e-Tools in and for elections and direct democracy in Switzerland', *Direct Democracy in and around Europe: Integration, Innovation, Illusion, and Ideology*, Center for Democracy Aarau (ZDA).

Lampathaki, F., Charalabidis, Y., Passas, S., Osimo, D., Bicking, M., Wimmer, M. and Askounis, D. (2010) 'Defining a taxonomy for research areas on ICT for governance and policy modelling', *EGOV*, pp.61–72.

Laver, M., Benoit, K. and Garry, J. (2003) 'Extracting policy positions from political texts using words as data', *American Political Science Review*, Vol. 97, No. 2, May, pp.311–331.

Louwerse, T. and Rosema, M. (2011) 'The design effects of voting advice applications: comparing methods of calculating results', *ECPR Conference*, Reykjavik, Iceland.

Mullen, T. and Malouf, R. (2006) 'A preliminary investigation into sentiment analysis of informal political discourse', *AAAI Symposium on Computational Approaches to Analysing Weblogs*, Stanford, CA, pp.159–162.

O'Connor, B., Balasubramanyan, R., Routledge, B.R. and Smith, N.A. (2010) 'From tweets to polls: Linking text sentiment to public opinion time series', *Proceedings of the International AAAI Conference on Weblogs and Social Media*, Washington DC, pp.122–129.

Sasaki, D. (Ed.), Avila, R., Chak, S., Górnicki, J., Heacock, R., Kaonga, V., Presley, S., Ribeiro, M.M., Singh, N. and Yang, C. (2010) *Technology for Transparency Network. The Role of Technology and Citizen Media in Promoting Transparency, Accountability and Civic Participation*, Available online at: http://www.right2info.org/resources/publications/technology-for-transparency

Soulis, K. (2011) *Visualizing the Public Opinion as it Takes Shape During a Public Consultation in Greece*, National School of Public Administration (in Greek).

Stylios, G., Christodoulakis, D., Besharat, J., Vonitsanou, M., Kotrotsos, I., Koumpouri, A. and Stamou, S. (2010) 'Public opinion mining for governmental decisions', *Electronic Journal of e-Government*, Vol. 8, No. 2, pp.203–214, Available online at www.ejeg.com

Thomas, M., Pang, B. and Lee, L. (2006) 'Get out the vote: determining support or opposition from congressional floor-debate transcripts', *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp.327–335.

Trechsel, A.H. and Mair, P. (2009) *When Parties (also) Position Themselves: An Introduction to the EU Profiler*, Robert Schuman Centre for Advanced Studies, European Union Democracy Observatory (EUDO).

Walgrave, S., Van Aelst, P. and Nuytemans, M. (2006) 'Do the vote test': electoral effects of a vote advice application at the 2004 Belgian elections', *ECPR Joint Workshop Sessions*, Nicosia, Cyprus.

Notes

[1]http://kohovolit.eu
[2]http://www.cs.waikato.ac.nz/ml/weka/
[3]http://www.jfree.org/jfreechart/
[4]http://ec.europa.eu/yourvoice/consultations/index_en.htm
[5]http://opencloud.mcavallo.org/