Social Voting Advice Applications - Definitions, Challenges, Datasets and Evaluation

Ioannis Katakis, Nicolas Tsapatsoulis, Fernando Mendez, Vasiliki Triga, Constantinos Djouvas

***Abstract*—Voting Advice Applications (VAAs) are online tools that have become increasingly popular and purportedly aid users in deciding which party/candidate to vote for during an election. In this paper we present an innovation to current VAA design which is based on the introduction of a social network element. We refer to this new type of online tool as a *Social Voting Advice Application* (SVAA). SVAAs extend VAAs by providing (a) *community-based* recommendations, (b) comparison of users' political opinions and (c) a channel of user communication. In addition, SVAAs, enriched with data mining modules, can operate as *citizen sensors* recording the sentiment of the electorate on issues and candidates. Drawing on VAA datasets generated by the Preference Matcher research consortium, we evaluate the results of the first VAA –Choose4Greece– which incorporated social voting features and was launched during the landmark Greek national elections of 2012. We demonstrate how a Social VAA can provide community based features and, at the same time, serve as a citizen sensor. Evaluation of the proposed techniques is realized on a series of datasets collected from various VAAs, including Choose4Greece. The collection is made available on-line in order to promote research in the field.**

I. Introduction

One of the most significant advancements of the world wide web is the establishment of social media, where users form online communities, share content and interact with each other. With 23% of web traffic generated by social network usage [1], social media are radically changing the way content is developed and distributed. This is especially the case in the political domain where the impact of social media is being attributed with an increasingly significant role, for instance in facilitating the recent popular mobilizations in various Arab spring countries [2].

In this paper we introduce a relatively new online phenomenon that involves the development of so-called Voting Advice Applications (VAAs), which are deployed during an election campaign. As the name implies, VAAs are essentially vote recommendation systems. For many advocates of such online tools, if appropriately designed VAAs could be beneficial for the electoral process since they promote more rational reasoning on the part of voters, fill important information gaps and can ultimately have positive impact on voter turnout [3]. A typical VAA involves users expressing their preferences on a set of policy statements that cover a range of issues deemed relevant to the wider election campaign. Most VAAs have around 30 policy statements, but some such as Smartvote in Switzerland can include over 70 policy items. The user is invited to submit their opinion, whether they strongly disagree, disagree, agree, strongly agree, on a particular policy statement. The policy positions of parties or candidates are already encoded in the system so that once a user has filled in the online policy questionnaire the tool then matches the user with the closest party or candidate. In short, VAAs currently provide a voting recommendation based on the overall congruence of a user's stated policy preferences with those of parties or candidates in the system –with the closest match ranked highest.

In this paper we build on current VAA research and practice by exploring an innovation which we refer to as a Social Voting Advice Application (SVAA). As a complementary feature of a VAA, it can be distinguished from current issue based recommendations by providing community based recommendations. The latter enables users to compare political opinions with their friends, other users, as well as providing a channel of inter-network communication through a blog like feature. The principal analytical focus of this paper will be on community recommendations based on like-minded users. Building on this social element, we also discuss the utility of SVAAs as citizen sensors and, in particular, how SVAAs enriched with data mining modules, can sense the emotion of the electorate on specific issues and parties/candidates.

To make our argument we will draw on datasets gathered by a VAA has been developed by a research consortium, Preference Matcher (www.preferencematcher.org), that brings together researchers from three universities (Cyprus University of Technology, University of Zurich and the University of Twente). For the purposes of illustrating our social voting model we will be principally relying on a pioneering case, the Choose4Greece VAA, which was deployed during the landmark Greek elections of May 2012 by the Preference Matcher team. However, for the purposes of comparison we will also include a number of Preference Matcher VAAs deployed in five other country cases between 2010 and 2013.

The contribution of this work can be summarized into the following points:

- The proposal of a number of different approaches for generating vote advice in SVAAs. Two of these approaches are specifically designed for the problem of providing community-based advice in a Voting Advice Application.
- An extensive experimental evaluation and discussion of the proposed vote advice systems on real datasets.
- A collection of new datasets, taken from real VAAs, that

Ioannis Katakis, Nicolas Tsapatsoulis, Vasiliki Triga and Costantinos Djouvas are with the Department of Communications and Internet Studies, Cyprus University of Technology (email: ioannis.katakis@cut.ac.cy, nicolas.tsapatsoulis@cut.ac.cy, vasiliki.triga@cut.ac.cy, costas.tziouvas@cut.ac.cy)

Fernando Mendez is with the e-Democracy Center, University of Zurich (email: fernando.mendez@zda.uzh.ch)

were launched in various countries. As far as we know, this includes the first application of community based recommendation in a VAA.

This work extends and improves the preliminary study presented in [4][1]. In addressing this topic, which lies at the intersection between Computer and Political Science, our aim is to make this contribution as comprehensible to both communities as possible. The structure of the paper is as follows. In the following section we review work on voting advice applications and recommendation systems. In Section III we define the concept of a Social VAA, present its advantages, and introduce a SVAA developed by our research team (Choose4Greece). Next, (Section VI) a number of issue-based and community-based vote suggestion approaches are presented. Section V presents the collection of datasets compiled by our research team. Evaluation setup is described in section VII while results are discussed in Section VIII. Section IX includes a discussion on how SVAAs can be utilized as sensors. Finally, the findings and potential applications of our research are summarized in Section X.

II. Related Work

Voting Advice Applications are currently mostly studied from a political science perspective [5]. In this work we extend VAAs by incorporating social network information in order to provide vote advice. This requires the re-definition of the problem as a special kind of collaborative recommendation [6] or data classification task (see Section IV). In addition we investigate how SVAAs can be utilized as citizen sensors.

Our research is related to a number of different fields, namely a) Voting Advice Applications (currently addressed by Political Science), b) Recommendation Systems (from Computer Science), and c) Social Networks as Citizen Sensors. In the last subsection, we discuss our work in relation to other dataset collections.

A. Voting Advice Applications

Largely as a result of their recent proliferation, much of this aided by the rise of social media, as well as their potential to have an impact on electoral outcomes, VAA research during this formative stage has been largely taken up by political scientists [7]. A number of research agendas are beginning to emerge at this, admittedly, early stage of diffusion (see [8] for the most recent review). Some authors, not surprisingly, adopt a rather positive assessment of VAA potential ([9], [10]) whilst others are more skeptical [11]. An obvious starting point has been to focus on the characteristics of VAA users ([12],[13], [14]). Some have tried to measure possible effects of VAAs on political mobilisation ([15], [16]). In fact according to some latest research, scholars have estimated a positive effect on turnout in the region of 4 per cent in the Dutch case of 2006 ([17]). Another perspective is to take up the issue of how VAAs are designed. A number of methodological concerns have emerged. The area most relevant to this paper is research on the choice of matching algorithms used by VAAs. In [18], Mendez compares four models for calculating the user-party congruence. The first two are based on how close the answers of the party and the user are (proximity models) and are implemented either by Euclidean or City Block metrics. Other models are based on so-called directional theory, which take into account the polarity of the opinions (i.e. if the answer of the voter and the candidate lie on the same side (disagree - agree) of the Likert scale). Two metrics are used, the Scalar Product and a hybrid metric. The basic claim of the paper is that directional inspired models perform better. In [19] the authors share their concern that the output of voting assistance tools might be strategically manipulated by political actors and that VAAs might be most advantageous to non-programmatic political parties. Finally, Walgrave *et. al.* [20] study the effect of the selection of statements and its impact on the recommendations that are produced. The paper suggests that certain configurations might favour certain parties.

All in all, VAA research has been currently dominated by political science perspectives. One limitation of current VAA design, which has been largely dominated by the political science approach, is that thus far little has been done in terms of taking advantage of the voter community in order to enable advanced features like collaborative vote suggestions or interactions between voters.

B. Recommendation Systems & Data Mining

A recommendation engine is an information system that suggests items (e.g. books, movies, etc) to users. The methods that have been proposed for recommendations can be organized into the following categories (for an extended review of the field see [21]):

- Content-based: Users are recommended items similar to the ones they preferred in the past [22], [23].
- Collaborative-filtering: Users are recommended items that people with similar preferences liked in the past [24], [25], [26].

Concerning the first approach, the preferences of the user for political parties other than the one they intend to vote for are typically unknown. Furthermore, the voting history of a user is generally not collected or not existing (new voters). Moreover, as we discuss in Section IV, the voting recommendation problem has an additional dimension in comparison with conventional content-based recommendation problems. In this case, there are the users (voters), the items (political parties) and the questions (policy statements). In order to produce recommendations we need to exploit all three elements.

As far as we are aware, the only recent work that deals with the VAA problem from a Computer Science perspective is that of Terán et al [27], [28]. However, this is still an issue-based recommendation, since the vote intention of the users is not taken into consideration and voter modeling is not constructed (see Section IV). The authors use a fuzzy clustering approach to provide a visualization-based recommendation. In order

[1]in the following ways: a) five approaches of voting advice are discussed with many additional variations, b) five datasets are utilized in the evaluation section that are made available online, c) an extensive experimental evaluation is included, using more datasets, methods, and evaluation metrics, d) the discussion includes additional topics like user evaluation, stream classification and the cold start problem, correlation of recommendations and social VAAs as citizen sensors.

to achieve this, they apply dimensionality reduction to the profiles of the candidates and the VAA user seeking for advice. Another recent work towards dimensionality reduction is presented in [29].

C. Mining Social Networks

The rise of social networks during the last years draw the attention of the research community. Some representative approaches are discussed in this section. ArnetMiner is one of the most popular publication/researcher search engines. In [30] many challenging tasks are confronted like the modeling of the research community using generative probabilistic models. In [31] the authors present LikeMiner which is a software that can capture the represenativeness and influence of objects. The system is based on mining a 'like'-graph in a social network and its effectiveness is demonstrated on Facebook data. Another interesting work on a similar topic is presented in [32], where the authors utilize what they define as "social endorsements networks" in order to assign tags to entities existing in social systems like Twitter or Flickr.

Contrary to the aforementioned social networks, in contemporary VAAs, like Choose4Greece, only a few types of user connections are recorded. Hence, since many approaches in this category are graph-based methods [33], they are not suitable for our case. However, the latest Preference Matcher platform will allow such connections by tracking relationships between a user recommending a VAA and the user accepting the invitation and filling the questionnaire (for more information on this feature please refer to Section X).

D. Social Networks as Citizen Sensors

Social Networks are constantly gaining popularity since they provide users with valuable community based features. Moreover, the vast information exchange as well as user interactions in an online community can be exploited for a variety of cases. In [34], insights into the dynamics of rating content by users of social tagging systems are described while [35] overviews approaches that analyze group information in Flickr in order to improve image retrieval.

Social media, like Twitter, are utilized for political analysis and for sensing the emotion of the electorate. [36] investigates the role of Twitter in political deliberation and participation by analyzing the ways in which South Korean politicians use Twitter. In [37] the authors study the process of opinion convergence by analyzing Twitter data of Singapore General Elections 2011. Canover et al. [38] describe several methods for predicting the political alignment of micro-blogging users based on the structure and content of their communication. Finally, a study on user comments on the Facebook walls of politicians is described in [39].

E. Dataset Contribution

The area of Machine Learning / Data Mining is one of the most applied research fields in Artificial Intelligence with rapid emergence of new interesting application domains. For the newly introduced problems, experts of the field have to invent new approaches or adapt the already existing ones to the problems' specific requirements. The main obstacle in many cases is the unavailability of real-world benchmark datasets that would enable the evaluation of novel solutions. Inevitably, artificial data are created but the results obtained are considered equivocal.

In order to overcome this obstacle researchers of the field, with the aid of domain experts, introduce such benchmarks in various data mining / machine learning venues [40], [41], [42], [43]. This paper introduces and makes available a number of real-world datasets for the domain of Voting Advice Applications. The authors' motivation is to promote research in the field.

III. Choose4Greece & Social VAAs

As mentioned above our focus is on one particular deployment of a VAA developed by the Preference Matcher team, the Choose4Greece (http://www.choose4greece.org) VAA. Where relevant we also draw insights through comparison with VAAs deployed by Preference Matcher team in other countries. The key feature of the Choose4Greece experiment explored in this paper is the introduction by the research team of a social voting advice module into the VAA design. More specifically, Choose4Greece implemented the following social features:

- *Collaborative voting advice*. In addition to voter-party similarity, Choose4Greece is able to provide with community based recommendations (see next section).
- *Friend functions*. Users can compare political views with each other. Each user is provided with a unique private PIN number used to save results and a public PIN (Figure 1a) that can be sent to other users for comparison (Figure 1b).
- *Blog*. Users could leave comments and feedback at the Choose4Greece blog where they also interact with the research team.

There were additional, innovative features of the VAA, which include the way in which the political parties were coded (using a Delphi system) and the presentation of results based on two different matching algorithms. The description of these additional features can be found in a report [44] in the FAQs sections of the Choose4Greece VAA. However, the methodological issues addressed in that report are not relevant to the analysis conducted in this paper.

As with any self-administered online survey there is a self selection problem with VAAs. Respondents, especially in the case of VAAs, tend to be more interested in politics and more educated than the average voter, and therefore more likely to complete a 30 policy statement questionnaire. Nonetheless with many of the successfully deployed VAAs, as the number of users increases it may be possible to make inferences about the wider population of voters. Indeed, there are a number of papers that do precisely this by using weighting techniques to make inferences about, say, the dimensionality of the political space [45] or the positions of political parties in the policy space [46]. However, representativeness is less of an issue for our paper since we are not making inferences about the entire voter population but rather those that participated by completing the VAA questionnaire. In the case of chooe4greece, this amounted to more than 75,000 users after data cleaning (more details for the dataset are presented in Section V).

A typical VAA includes two basic elements: a) a questionnaire that voters complete and b) the output of the VAA which is a similarity match between the user and all the political parties included in a VAA. However, with an appropriately designed infrastructure it is possible that a VAA can operate as a social network (*Social Voting Advice Application - SVAA*) where the users can take advantage of the community and interact with each other. It is important to note that adding a Social Networking component to a VAA presents some challenging issues: First of all, users are - and should be - completely anonymous. Anonymity prohibits user registration, web browser cookies or IP tracking. Hence, the VAA community consists of unique visitor ids. Second, information stored in a VAA is politically sensitive and should be protected by robust data protection safeguards. Lastly, vote advice procedures should be open and transparent to public.

The Choose4Greece consortium made every effort to address the above issues. In particular, users were never requested to register to the system. In order to provide some advanced features, user identification was necessary. In that cases users were provided with a private and a public key. Using those two keys voters could return to their results page without re-filling the questionnaire. The friend function (comparison of agreement with friends that had filled in the questionnaire) was also based on the public and private keys. In addition, no data were recorded that could aid tracking the real identity of the users. As far as the transparency issues are concerned, the Choose4Greece consortium is publishing scientific publications describing the approaches and, at the same time, additional articles were written in a non-technical language targeted to the broader audience. Many of these articles are referenced at the Preference Matcher website.

IV. Problem Definition & Notation

In the problem of voting suggestion there is a set of N users $U = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_N}\}$, a set of M questions (or issues or statements) $Q = \{q_1, q_2, \ldots, q_M\}$, and a set of T political parties (or candidates) $P = \{\vec{p_1}, \vec{p_2}, \ldots, \vec{p_T}\}$. Each user $\vec{u_i} \in U$ and each political party $\vec{p_j} \in P$, has answered each question $q_k \in Q$. The answers of users are recorded through on-line questionnaires like the one in Choose4Greece. The answers of political parties are either coded by experts or answered by representatives of political parties.

Based on their answers, every political party and user can be represented in a vector space model:

$$\vec{u_i} = \{u_{(i,1)}, u_{(i,2)}, \ldots, u_{(i,k)}, \ldots, u_{(i,M)}\} \tag{1}$$

$$\vec{p_j} = \{p_{(j,1)}, p_{(j,2)}, \ldots, p_{(j,k)}, \ldots, p_{(j,M)}\} \tag{2}$$

where $u_{(i,k)}, p_{(j,k)} \in L$ are the answers of the i-th user and j-th party, respectively, to the k-th question. In most cases, vectors $\vec{u_i}$ and $\vec{p_j}$ are called *profiles*.

A typical set of answers is a 6-point Likert scale: L ={1 (Strongly disagree), 2 (Disagree), 3 (Neither agree nor disagree), 4 (Agree), 5 (Strongly agree), 6 (No opinion)} but in practice the sixth point is not taken into consideration since it does not correspond to a particular stance. As a result the set L, in the context of this work, becomes: L ={1,2,3,4,5}.

The *task*: *Given the answers of a specific user* $\vec{u_a}$ *suggest a ranking of political parties based on user-party relevance.*

Essentially, the goal is to approximate the hidden function $h : \mathbb{R}^M \times \mathbb{R}^M \rightarrow \mathbb{R}$, where $h(\vec{u}, \vec{p})$ is the estimation of the relevance of user $\vec{u}$ with political party $\vec{p}$. Typically $h(\vec{u}, \vec{p}) \in [0, 1]$. In each case, the top suggestion s for user u should be:

$$s_{top} = \underset{p}{\operatorname{argmax}}[h(\vec{u}, \vec{p})] \tag{3}$$

where $s_{top} \in \{p_1, \ldots, p_T\}$. Similarly, we could consider a function $r(\vec{u}, \vec{p}) \in [1, T]$ that returns the *rank* of the political party p for the user u, if all political parties are ranked according to relevance (similarity) with this specific user. Having learned function $h(\vec{u}, \vec{p})$ it is straightforward to calculate $r(\vec{u}, \vec{p})$.

In order to produce vote recommendations, the most simple approach is to define $h(\vec{u}, \vec{p}) = d(\vec{u}, \vec{p})$ where d is a distance function between $\vec{u}$ and $\vec{p}$. This is, in general, the approach used by most VAAs. A number of such distance measures are discussed in [18].

In the Preference Matcher VAAs, the information of *vote intention* $\nu_i \in P$ of many users $\vec{u_i}$ is available as it is included as a supplementary question in the on-line surveys. This kind of information, as we discuss later on, can be utilized to provide voting recommendations.

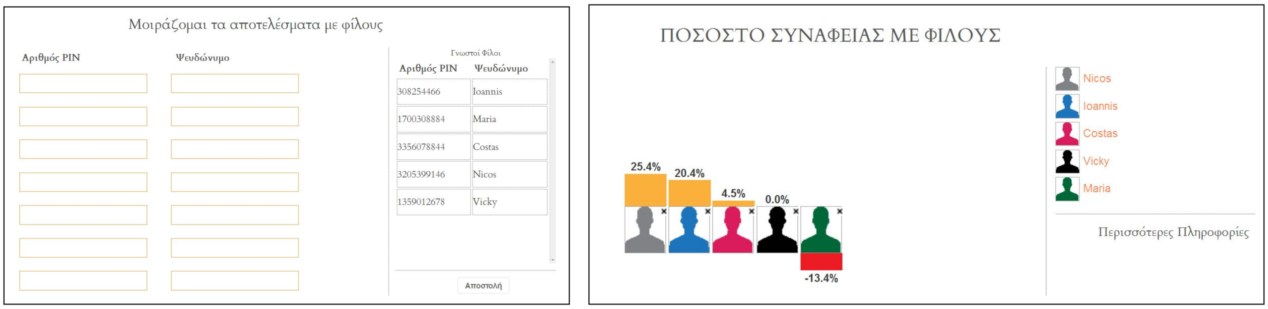
If we ignore the information of political party profiles, then the problem can be defined as a single-class data classification problem, with the class obviously being the vote intention of the user. The data matrix ($D_{N\times M}$) consists of all the users profiles and the class (label) vector

$$V_{N\times 1} = \{\nu_1, \ldots, \nu_N\}^T \tag{4}$$

consists of the vote intentions of all N users. Hence, $D_{N\times M}$ and $V_{N\times 1}$ constitute the training *examples* of the learning problem. In this case many classifiers can be represented as score functions $f(\vec{x}, y)$ that output the probability that instance $\vec{x}$ belongs to class y (i.e. $P(y|\vec{x})$), and $\sum_{y=1}^{T} f(\vec{x}, y) = 1$. Hence, in order to solve the VAA problem, it is straightforward to set $h(\vec{u}, p) = f(\vec{x}, y)$ and use f as one of the variety of classifiers (Decision Trees Classifiers, Bayesian Classifiers, Support Vector Machines, Neural Networks, etc). In Section VIII we evaluate and comment on a number of such classifiers. In essence what is achieved in this case is the modeling [47] of the users' political preferences based on the policy statements and the corresponding vote intention. It is important to note that in some cases, classifiers can model the extracted knowledge in an easily interpreted way by humans. Decision Tree algorithms for example [48] can model voters behaviour in a hierarchy of tests on the various VAA questions. This characteristic is of great value for VAAs and is used to model and generate a vote recommendation.

At this point we have to make a distinction between the two different kinds of voting advice.

- *Issue-based advice*: The VAA suggests to the user the candidate or political party that has the highest *degree of congruence* in the issues recorded by the VAA.



(a) Submitting Pin Numbers (b) Comparison with Friends

Fig. 1: Screenshots of Choose4Greece - Friends Function

TABLE I: Symbols used in the paper and descriptions

Symbol	Description
u_i	User i
$\vec{u_i}$	The profile of user u_i
N	The total number of users
U	The set of users
q_k	k-th question
M	Total number of questions
Q	The collection of questions in a VAA
p_j	Political Party (candidate) j
$\vec{p_j}$	The profile of political party p_j
T	Total number of political parties (candidates)
P	The set of parties participating in elections
L	The set of possible answers for all questions
$h(\vec{u}, \vec{p})$	Relevance of $\vec{u}$ with $\vec{p}$
s_i	The suggestion (advice) of the VAA to user u_i
$d(\vec{u}, \vec{p})$	The (Euclidean) distance between $\vec{u}$ and $\vec{p}$
$a(\vec{p_j})$	Profile of the average voter of p_j
ν_i	The vote intention of user i
D,D_{train},D_{test}	Dataset (user profiles), training data, testing data
V	Collection of vote intentions of all users

- *Community-based advice*: Used in Social VAAs like Choose4Greece. The SVAA suggests to the user the candidate or political party that *similar users* (i.e. users that answered similarly to the questions) will vote for.

Since the users of a VAA could be interested in both types of advices, we discuss both of them in this paper and provide both of them in Choose4Greece.

For the convenience of the reader, we summarize all notation and symbols used in the paper in Table I.

V. The Datasets

In this section we provide information about the newly generated VAA datasets that are available at the Preference Matcher website. Detailed information is presented for the Choose4Greece dataset since it was the pioneer case study in which the community based recommendations was first applied. Information about the other datasets appears in the Preference Matcher website.

A. The Choose4Greece Dataset

The dataset consists of information collected from the usage of the Choose4Greece system during the period April - May 2012 for the 2012 National Elections in Greece. There were two rounds of elections in Greece 2012 (May 6 & June 17). The dataset under study includes data collected for the elections on May 6th. The dataset contains 75294 voters out of which 26335 provided with their vote intention and 15 parties.

Users of Choose4Greece had to submit their opinion for 30 issue statements plus some supplementary questions asking for demographic information, voting intention and self-placement on the main political dimensions (left/right, traditional/liberal). The issue statements are presented at Table II. For each issue statement, the user had to choose one of the following answers: 1) Completely agree, 2) Agree, 3) Neither agree nor disagree 4) Disagree 5) Completely disagree 6) No opinion.

B. The VAA Dataset Collection

In addition to the Choose4Greece dataset we included 4 datasets from elections in other countries (Cyprus, Brazil, Peru, Scotland). Summary information about the collection can be found in Table III. Information in this table includes: the number of questions included in the questionnaire, the number of political parties participated in the elections, the number of users after the cleaning process, the number of users that declared their vote intention, the time period of the election and the election type. All datasets and accompanying information files are available at http://www.choose4greece.org/datasets/.

C. "Cleaning" the Datasets

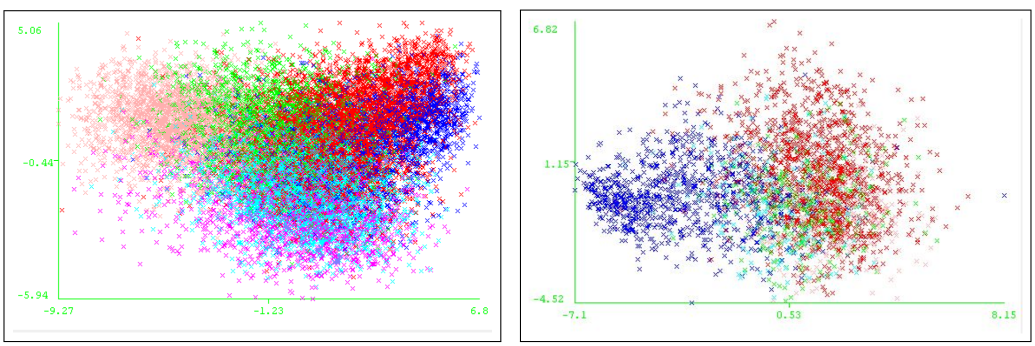
The datasets had to be pre-processed in order to remove invalid records. The first step was to filter all user-entries that did not exceed the time threshold of 120 seconds during the whole session. We considered that if a user spent less that 120 seconds to complete the full questionnaire (that is approximately 4 seconds per question - not considering the supplementary questions which are not mandatory) then probably she would answer the questions randomly. Another important step was to remove user entries that did not complete the full 30-questions.

D. Dimensionality Reduction

Dimensionality reduction is a common step in data analysis. In the case of VAA data, it is important to be able to reduce the original data into a two dimensional space in order to provide with human comprehensible visualizations [28], [27].

TABLE II: The Choose4Greece 30 Questions

1) Priority should be given to economic growth even if this leads to a wider gap between rich and poor
2) Taxes on large corporations should be increased
3) The exploitation of the mineral wealth of the country will lead to environmental degradation rather than contribute to economic growth
4) The installation of solar panels on productive and fertile agricultural land should be prohibited
5) Privatization will help reduce the deficit
6) Strong policing in town centers should be implemented to tackle crime.
7) The installation and use of closed circuit cameras in public places (e.g. shopping centers) should be prohibited
8) In order to maintain order in the cities, the state should take more restrictive measures on demonstrations
9) Granting Greek citizenship on favorable terms to second generation immigrants will encourage further immigration to Greece
10) Multiculturalism in Greece is a positive phenomenon
11) To combat unemployment, workers must accept the new forms of flexible working conditions (e.g. part time jobs)
12) Co-funding the universities by private investors will have negative effects in higher education
13) The introduction of university fees will ensure a better functioning of the public university
14) Merging smaller hospital and creating large hospitals will deteriorate the provision of health services
15) It is feasible to reorganize the public sector without dismissing civil servants
16) The pay cuts are necessary for overcoming the crisis
17) The proper implementation of the IMF/EU/ECB memoranda will lead Greece to overcome the financial crisis
18) The loans provided to Greece have only benefited the banks and lenders
19) Signing the memoranda with the IMF/EU/ECB means selling out Greece to foreigners
20) The second memorandum with the IMF/EU/ECB was necessary to prevent Greece from bankruptcy
21) The Inter-governmental Fiscal Treaty will not help Europe to overcome the crisis
22) Greeces exit from the euro zone would help address the economic crisis
23) Possession of soft drugs (eg, cannabis) for personal use should be decriminalized
24) Reducing defence spending (e.g. closing military camps) will provide resources for the welfare state
25) Economic growth can be achieved by liberalizing all of the closed professions
26) Church property should be exempted from taxation so that the church can engage in charitable causes
27) The participation of (non-MP) technocrats in the cabinet will lead to a better management of the economy
28) Greece should strengthen its ties with Israel
29) The name issue of the Former Yugoslav Republic of Macedonia (FYROM) should be resolved by Greece's acceptance of a composite name with a geographical qualifier
30) Greece should leave the European Union



(a) Greece (b) Cyprus

Fig. 2: VAA users after PCA leading to a two dimensional space. Color represents vote intention (political party)

Given a data matrix $D_{N\times M}$ a dimensionality reduction technique R is able to be applied in D and produce a new table $D'_{N\times M'}$ where $M' < M$ or preferably $M' << M$. In Principal Component Analysis, attributes in M' are linear combination of the attributes in M. Moreover, with PCA, the information contained in the new attributes (Principal Components) is evaluated and hence only the most informative attributes can be maintained. In Figure 2(a) the users of Choose4Greece are presented in the first two Principal Components. Colour represents the vote intention (political party). For presentation reasons we kept only the users of the 6 most popular political parties. Since only two Principal Components are presented there is information loss. However, we still observe that there are space regions with voters of the same political party. In other words, voters of the same political party tend to reside close to each other. This type of dimensionality reduction, like PCA is named unsupervised dimensionality reduction, because class-label (vote intention) is not taken into consideration. Figure 2(b) depicts the same representation for the Cyprus dataset where we observe similar behaviour. The rest of the datasets present similar behaviour, with the exception of the Scotland dataset where, due to the information loss, the boundaries between users of different political parties are not so clear. This is consistent with the findings of [46] in which it is shown that a three dimensional space best captures the Scottish political landscape.

A second approach for dimensionality reduction is to evaluate each attribute based on the information they provide in the task of identifying the class of an object (supervised dimensionality reduction). Such measures include information

TABLE III: Dataset Information

	Questions	Parties	Users	Users with VI	Time Period	Type
Greece	30	15	75294	26335	April-May 2012	Parliamientary
Cyprus	30	11	4900	3596	April-May 2011	Parliamentary
Brazil	30	10	16513	16254	September-November 2010	Presidential
Peru	30	5	40627	32043	March-April 2011	Presidential
Scotland	30	11	20730	10615	April - May 2011	Parliamentary

gain the x^2 measure, etc [49]. In table IV we observe the top-5 questions of the Choose4Greece dataset ranked according to their information gain. This procedure could provide with a smaller questionnaire that will demand less time on behalf of the users. Note that 4 out of 5 highest-rank questions are related to Greece's debt. In Table V the same information is presented for the Cyprus dataset. We omit similar tables for the rest of the datasets due to space limitations.

TABLE IV: The five most informative attributes (questions) according to information gain with respect to class (vote intention)

Rank	Info Gain	QID	Question
1	0.3535	20	The second memorandum with the IMF/EU/ECB was necessary to prevent Greece from bankruptcy
2	0.3502	19	Signing the memoranda with the IMF/EU/ECB means selling out Greece to foreigners
3	0.3312	17	The proper implementation of the IMF/EU/ECB memoranda will lead Greece to overcome the financial crisis
4	0.2991	18	The loans provided to Greece have only benefited the banks and lenders
5	0.295	8	In order to maintain order in the cities, the state should take more restrictive measures on demonstrations

TABLE V: Most Informative Questions based on Information Gain (Cyprus)

Rank	InfoGain	QID	Question
1	0.4726	1	In the negotiations for the Cyprus' problem, the Government has made unacceptable concessions
2	0.4057	9	Cyprus must apply for membership in the program "Partnership for Peace"
3	0.3709	8	The position of Cyprus is in NATO.
4	0.3284	26	The views of the Church of Cyprus should be seriously taken into account regarding the formulation of the country's policy-making.
5	0.2679	2	A bi-zonal, bi-communal federation with one sovereignty and citizenship is an acceptable solution.

VI. Candidate Recommendation Systems

In this section we present a number of approaches for providing vote-advices: *issue based* and *community based.* All methods, with the exception of Party-Coding similarity (see Section VI-A), are first applied and make calculations on the existing database of the VAA (D_{train}), assuming that there is a number of user profiles with the corresponding vote intention. Then, the recommendations (advices) are provided by the VAA to new users (D_{test}) given only their profile. The Party-Coding similarity method is directly applied to D_{test} since no D_{train} is required.

A. (Weighted) Party-Coding Similarity

This is the issue-based approach most widely used in Voting Assistance Applications. In this case $h(\vec{u},\vec{p}) = d(\vec{u},\vec{p})$, where d is the Euclidean distance:

$$d_{\text{Euc}}(\vec{u}_i,\vec{p}_j) = \sqrt{\sum_{k=1}^{M} w_k (u_{(i,k)} - p_{(j,k)})^2} \tag{5}$$

Where w_k is the weight of question k. Weight could be: a) defined by experts for all users (based on what the expert considers critical issues for a community), b) defined by the voters (based on personal preference), c) or, as in our case, automatically defined by evaluating each attribute in terms of information gain (see Section V-D).

Naturally normalization is necessary if h is required to be in $[0,1]$, with 0 meaning identical profiles. However, since the recommendation is $s = \underset{p}{\operatorname{argmax}}[h(\vec{u_a},\vec{p})]$ normalization is not required, even if a ranking of political parties is requested.

The advantage of this approach is that it provides the degree of agreement / disagreement with each political party. This information normally demands significant effort on behalf of the user. Another positive aspect of this approach is computational simplicity. The main disadvantage is that the profiles of political parties / candidates are not easy to collect. Another concern with this method is that usually users do not vote based on agreement with political parties (non-issue voters). Many citizens tend to vote based on other criteria like personal relations with party, personality of the party leader, effectiveness in solving the problems, etc (see [18] for more information). An algorithmic description of this approach can bee seen in Algorithm 1.

B. (Weighted) Average Voter

This is a simple community-based approach that calculates the distance between the user and the average voter of each party. The party with the nearest average voter comprises the recommendation in this approach. The average voter of party

Algorithm 1: The party coding algorithm

foreach *user* u_i *in* D_{test} **do**
  Create list l_i to contain <party,distance> tuples;
  foreach *party* p_j **do**
    l_i.insert($< p_j$,$d(\vec{u}_i, \vec{p}_j) >$);
  Sort l_i by distance;
  $s_i \leftarrow l_i$.get(first).getParty();

p_j is defined as:

$$a(\vec{p_j}) = \frac{1}{N_j}\{\sum_{i=1}^{N_j} u_{(i,1)}, \ldots, \sum_{i=1}^{N_j} u_{(i,k)}, \ldots, \sum_{i=1}^{N_j} u_{(i,M)}\} \quad (6)$$

where N_j is the total number of voters of political party p_j. In this approach $h(\vec{u_i}, \vec{p_j}) = d(\vec{u_i}, a(\vec{p_j}))$ where d is the distance between the user under study and the average voter of each party. Similarly to the party coding approach we are evaluating a simple version and a weighted one where we utilize the weights obtained from the information gain. As discussed previously depending on the application requirements h should be normalized.

In this particular research, we have included two measures of distance: a) the (weighted) Euclidean distance, which can be defined similarly to Equation 5, and b) the Mahalanobis distance which is defined as follows:

$$d_{\text{Mah}}(\vec{u_i}, a(\vec{p_j})) = a_j \sqrt{(\vec{u_i} - \vec{p_j})^T {S_j}^{-1} (\vec{u_i} - \vec{p_j})} \quad (7)$$

where S_j is the covariance matrix of the political party j and a_j is a stabilizing parameter tuned for political party j. We consider the Mahalanobis distance more suitable for the particular problem since many issues in the VAA's questionnaire are correlated. Using the covariance matrix, this distance takes into consideration the correlation between the various issue-questions. In the experimental section, we compare the Euclidean version of the Average Voter approach against the Mahalanobis version in order to investigate if such correlations are exploitable.

In general, the advantage of this approach is that it does not require the profile of each political party and that it is computationally undemanding. However, a sufficient number of users is necessary in order to calculate the average voters. In recommendation system literature this issue is known as "cold-start" problem [21]. An algorithmic description of this approach is shown in Algorithm 2.

Algorithm 2: The average voter algorithm

foreach *user* u_i *in* D_{train} **do**
  $\nu_i \leftarrow$ Get vote intention of u_i;
  Update average voter of party ν_i based on $\vec{u_i}$;
foreach *user* u_i *in* D_{test} **do**
  Create list l_i to contain <party,distance> tuples;
  foreach *party* p_j **do**
    l_i.insert($< p_j$,$d(\vec{u}_i, a(\vec{p_j})) >$);
  Sort l_i by distance;
  $s_i \leftarrow l_i$.get(first).getParty();

C. Clustering

This community-based approach is based on data clustering [50]. Given a set of data points in a multi-dimensional space, a clustering algorithm is able to organize data points into similar groups (*clusters*). Partitioning algorithms, like the widely known k-means, organize data based on feature space distance. Essentially, points that are close are assigned to the same group. In this work, we exploit clustering in order to organize voters into clusters: Voters will be similar in terms of their feature vector which expresses their answers in a VAA's issue-questions. Therefore, clustering will produce groups of like-minded users. This research direction of exploiting clustering in order to organize users of Social Media is suggested by other researchers as well [51].

After creating clusters, the system will be able to produce vote recommendations for new users. This is achieved by calculating the closest cluster to the new user. Then, the system suggests the political party that has the greatest number of voters (considering the vote intention) in that cluster. Note that the proposed approach is essentially a framework in which any clustering algorithm can be utilized. We investigate two such algorithms in the evaluation section.

A clustering algorithm (Φ) consists of two basic functions. A `clusterData()` function that organizes a set of data items into clusters given the number of clusters (k) and a `clusterItem()` function that identifies the cluster that a new item belongs to. This is usually the closest cluster to this item. An algorithmic description of this approach can be seen in Algorithm 3. We assume that the number of clusters is defined by the system engineer. Such is the case of widely used clustering algorithms like k-means or Expectation-Maximization [52]. Nevertheless, an optimal value for k can be identified through experimentation.

Algorithm 3: The clustering algorithm

$k \leftarrow$ desired number of clusters;
initialize VoteCounts= **0** (A $k \times T$ array that stores the number of votes of each political party for each cluster);
Clusters $C = \{C_1, \ldots, C_k\} \leftarrow \Phi.clusterData(D_{train}, k)$;
foreach *Cluster* C_l *in* C **do**
  foreach *user* u_i *in* C_l **do**
    $\nu_i \leftarrow$ Get vote intention of u_i;
    Update VoteCounts[l][ν_i];
foreach *user* u_i *in* D_{test} **do**
  $C_x \leftarrow \Phi$.clusterItem(u_i);
  $s_i \leftarrow$ party with most votes in C_x (calculated from VoteCounts[C_x][]) ;

A particular example of voting recommendation based on clustering in Choose4Greece is shown in Figure 3a. Once the user completed the set of 30 questions the closest cluster according to her profile is identified. Percentages of voting intention of the members of this cluster are used as recommendation and are illustrated as a bar chart (see Figure 3a). Note, however, that the results refer to cluster members that answered the supplementary question on voting intention. In summary: 38.2% of the cluster members choose to vote for 'Demokratiki Aristera', 33.3% choose to vote for 'Syriza', etc. Thus the voting recommendation was 'Demokratiki Aristera'. There are only four parties in the bar chart because only voters of those parties appeared in the particular cluster.

The corresponding recommendation based on the party-coding similarity scheme, for the same user, is shown in Figure 3b. We observe some differences both in the similarity scores and ranking of parties.

An obvious advantage of the user clustering approach is that it is not necessary to obtain the profiles of each political party / candidate. However, the most important characteristic is that it enables the organization of users into clusters. This feature will provide with three more advantages.

Firstly, it will enable the production of more accurate recommendations than the average voter and the user-candidate similarity since it will enable to create finer groups of users that will vote for the same candidate. Experimental evaluation (Section VIII-A) confirms this statement.

Secondly, it provides with valuable insight of the electorate. See for example the clusters produced after applying k-means on the Choose4Greece data (Table VI). One could note some interesting observations on this outcome. For example, we observe that Cluster 5 consists mostly of Siriza voters, Cluster 8 is a group of left party voters (KKE, Siriza, Dimokratiki Aristera), Cluster 4 consists of voters with right-conservative orientation (Nea Dimokratia, Anexartitoi Ellines, Xrisi Augi) and finally, Cluster 7 and 9 seems to have voters from various political parties.

Finally, each cluster can be represented by a centroid (average vector of each cluster). This representation is of great importance since it enables the interpretation of the opinions that dominate each cluster and can be exploited as a data compression technique.

D. Classifiers

As discussed in Section IV the problem of social voting advice can be formulated as a data classification task, where the class represents the political party. Note that this is a general framework, and any classifier (Θ) could be used. Each classifier can be *trained* on a dataset D_{train} and for each new item (voter) can output a vector that contains the membership probability (θ_i) for all classes (political parties) (see Algorithm 4).

Algorithm 4: The classification algorithm

$\Theta.train(D_{train})$;
foreach *user* u_i *in* D_{test} **do**
  $\{\theta_1, \ldots, \theta_T\} \leftarrow \Theta$.classify($u_i$);
  $s_i \leftarrow$ party with greatest θ_i

E. Collaborative Filtering

It is quite natural to perceive the voting advice task as a recommendation problem. Common approaches utilized in recommender systems are collaborative filtering techniques like the user-based and item based method. In this section we explain why the off-the-self collaborative filtering algorithms are not suitable for the particular task and propose a variation that is more suited to addressing the problem. This variation could be applied to other domains as well in case they share some similar characteristics with the vote suggestion task. Therefore, we propose a general method for collaborative filtering rather than a VAA-specific technique.

Typically, in collaborative filtering approaches, there exist two entities: *users* and *items*. Moreover, there is a data table

$$U_{(m,n)} = \begin{pmatrix} r_{(1,1)} & r_{(1,2)} & \cdots & r_{(1,n)} \\ r_{(2,1)} & r_{(2,2)} & \cdots & r_{(2,n)} \\ \vdots & \vdots & \ddots & \vdots \\ r_{(m,1)} & r_{(m,2)} & \cdots & r_{(m,n)} \end{pmatrix} \tag{8}$$

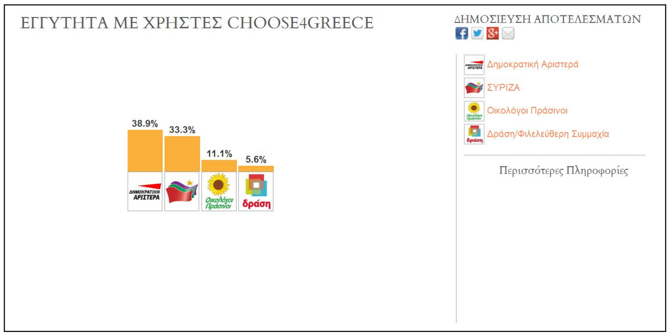
where $r_{(i,j)}$ is the rating that user u_i assigned to item o_j. Based on the above table, a collaborative filtering system provides recommendations to user u_x based on it's profile (ratings) by: a) identifying similar users (neighbors) and recommending to u_x items that were highly rated by the neighborhood (user-based approach), b) identifying and recommending similar items with the ones that where highly rated by user u_x (item based approach).

Unfortunately, in the case of VAAs as in other domains, there is no direct relationship of users and items. In the case of VAAs, users are not assigning "ratings" to political parties. Usually there is a supplementary question that asks the user what they have voted in previous elections and what do they intend to vote in the next elections (vote intention). Hence rating typically exists for just one party and it is actually a binary value ($r_{i,j} \in \{\texttt{vote, no-vote}\}$). Therefore the collaborative filtering model is not directly applicable to VAA data. However, we could utilize the profiles of users and parties (candidates) in order to provide collaborative based techniques. In particular, we have implemented a collaborative filtering technique by calculating the similarity of voters (users) utilizing their profile distance. In other words,

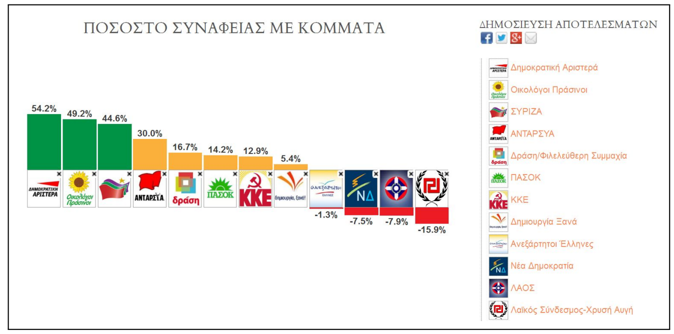
$$sim(u_1, u_2) = distance(\vec{u_1}, \vec{u_1}) \tag{9}$$

where $\vec{u_1}$ and $\vec{u_2}$ are the profiles of users u_1 and u_2. Remember that the profiles contain information about the opinion of the users regarding the issue-statements and not directly their preference to the political parties. In collaborative filtering, the purpose of calculating similarities is to identify the top-k similar users to the user under study (u_x). This list of most-similar users (usually called "neighbors") can assist in predicting the vote intention of u_x. Typically, in our case the prediction for u_x would be the most popular party in the neighborhood of u_x. In principle, this approach is similar to the kNN classifier [53].

In the context of collaborative filtering, a rather intuitive and reasonable approach would be to equate (opinion) similarity between party and voter profile as political preference (ranking



(a) Clustering-based Recommendation



(b) Party-Coding based Recommendation

Fig. 3: Screenshoots of Choose4Greece

TABLE VI: The ten clusters created using k-means ($k = 10$) in Choose4Greece dataset

#	PASOK	ND	KKE	LAOS	SIRI	DIAR	DISI	ANEL	OP	KISI	ARPO	DRASI	ANTA	XA	DIKS
1	14%	2%	2%	0%	12%	30%	2%	2%	10%	1%	0%	16%	2%	1%	6%
2	14%	30%	1%	3%	2%	6%	8%	8%	2%	1%	0%	14%	0%	8%	4%
3	21%	9%	0%	1%	0%	5%	9%	1%	2%	0%	0%	47%	0%	1%	4%
4	1%	11%	2%	5%	8%	3%	1%	32%	2%	1%	0%	2%	1%	29%	2%
5	1%	1%	8%	1%	29%	10%	1%	26%	5%	1%	1%	4%	2%	9%	2%
6	1%	1%	14%	1%	23%	3%	0%	31%	2%	0%	0%	1%	3%	19%	1%
7	4%	9%	3%	2%	12%	14%	3%	21%	4%	1%	0%	13%	1%	9%	4%
8	1%	0%	11%	0%	54%	10%	0%	4%	7%	1%	0%	1%	8%	1%	1%
9	5%	6%	6%	1%	26%	17%	2%	18%	6%	2%	0%	3%	2%	4%	3%
10	0%	0%	32%	0%	40%	1%	0%	1%	1%	0%	0%	0%	24%	0%	0%

of political parties - see Equation 8). Having done so a user-based collaborative approach could be applied as well. We do not include such an alternative for the following reason: We have observed that in most VAA data the preference of a voter to a political party is weakly correlated with the concordance of views (see Section VIII-A). In essence, people do not vote for (or prefer) political parties that agree with (non-issue voters). Many times, citizen's vote is based on other criteria (e.g. trust to the party leader, personal relationship with the party, etc). Hence we have rejected the assumption that political preference is synonym to coincidence of opinion.

VII. Evaluation Setup

This section presents in detail the evaluation procedure used to estimate the predictive performance of the methods discussed in the previous sections.

We use the 10-fold cross validation procedure to calculate the evaluation measures. Cross validation splits randomly the data into 10 equally-sized parts and uses 9 of them for training and the remaining one for calculating the evaluation metric. This procedure is repeated 10 times by choosing different parts for training and testing. Eventually, the evaluation metric is averaged over these repetitions.

In Average Voter, the training set is used to calculate the average vectors for each political party. In the Clustering approach, the training set is used to organize the voters into clusters and calculate the vote distributions for each cluster. In the Classifier approach the training set is used to train the classifier. The evaluation of all approaches (calculation of evaluation measures) was carried out in the test set. For Clustering and Classification, the Weka [54] implementations of Data Mining algorithms were exploited. We have evaluated one classifier for each category. More specifically, we included the Naive Bayes (NB) classifier, a Decision Tree classifier (J48), a Support Vector Machine (SMO), a Neural Network (NN - Multi-Layer Perceptron), and a Rule Learning algorithm (JRip). All classifiers were set with Weka default parameter settings.

A. Evaluation Measures

In order to evaluate and compare the aforementioned approaches in terms of quality of prediction, we exploit the following evaluation measures:

1) Accuracy: This is a widely used evaluation measure for classification problems. It calculates the percentage of correct predictions. If a prediction of an approach h for user i is $p_i = \underset{p}{\operatorname{argmax}}[h(\vec{u_i}, \vec{p})]$ and the vote intention is ν_i then, accuracy for method h in dataset D is calculated as:

$$acc(h, D) = \frac{1}{|D|} \sum_{i=1}^{|D|} e(p_i, \nu_i) \tag{10}$$

where $|D|$ denotes the cardinality (i.e., number of user entries) of set D, and

$$e(p_i, v_i) = \begin{cases} 1 & \text{if } p_i = v_i \\ 0 & \text{if } p_i \neq v_i \end{cases}$$

Accuracy is a strict measure that considers only the cases where the recommendation system has placed first the correct political party / candidate.

2) Weighted Mean Rank: This is a measure that evaluates how high did the recommendation system placed the correct political party. Weighted Mean rank differs from Accuracy in the following way: Consider two recommendation systems h_1 and h_2 with ranking functions r_1 and r_2 and a user u_a with vote intention ν_a. If $r_1(u_a, \nu_a) = 2$ and $r_2(u_a, \nu_a) = 4$ then these cases will be treated equally in accuracy since none of these methods ranked first the correct political party (ν_a). On the other hand, weighted mean rank is defined as follows:

$$wmr(h, D) = \frac{1}{T} \sum_{j}^{T} w_j \frac{1}{N_j} \sum_{i}^{N_j} r(u_i^j, p_j) \tag{11}$$

where T is the number of political parties, w_j is the percentage of voters that party j collected in the training set, N_j is the number of voters of party j in the evaluation (test) set, r is the ranking function corresponding to recommender h and u_i^j is the i user (voter) of political party j. wmr takes into consideration the ranking of the correct political party (vote intention) and the number of the voters of each political party. In short, the weighted mean rank takes into account the vote share of the political party in the respective election. Naturally, the closer wmr is to 1 the better.

3) Mean Rank: Mean Rank (mr) is essentially very similar to Weighted Mean Rank. The difference is that in Mean Rank, the percentage of votes of each political party (weight - w_j) is ignored.

4) Precision, Recall, F-measure: $prec$, rec, f measure are well known metrics from the field of information retrieval. Note that in each case we calculate the average $prec$, rec and f. First, the precision, recall and f-measure for each party is calculated. Then the average over all parties is calculated. We intentionally left this average un-weighted, meaning that in these metrics we do-not take into consideration the number of voters of each party. This way, these measures significantly penalize the approaches that make mistakes in small parties. However, since VAA data are particularly imbalanced in terms of voters, we mostly comment on the accuracy of the approaches.

VIII. Results and Discussion

This section presents and discusses the experimental outcome of our work.

A. Comparative Results

Table VII displays the results of the first five approaches (Party Coding, Weighted Party Coding, Average Voter using Euclidean Distance, Weighted Average Voter, Average Voter using Mahalanobis distance and Clustering) in six evaluation measures: Accuracy (acc), Mean Rank (mr), Weighted Mean Rank (wmr), Precision (pre), Recall (rec), F-measure (f). For clustering, we use k-means algorithm with $k = 200$, we elaborate on the selection of k later on.

We observe that the Average Voter using the Mahalanobis distance performs better in all metrics. Clustering presents the second best performance. This result confirms our initial intuition that clustering will organize users into like-minded voters who tend to vote for the same political party. The Average-Voter algorithm presents better predictive performance than the Party-Coding similarity. The bad performance of Party-Coding suggests that voters are not fully consistent in terms of the overlap of their policy preferences and the parties they vote for. However, it is interesting that the Weighted Party Coding approach outperformed the basic Party Coding approach. This outcome underlines the utility of weighting the questions based on the information gain. In addition, Table VII presents the training and testing time for each method (to train/test the whole dataset). Naturally, the party coding requires no time for training and the clustering approach is the most time consuming approach mainly because of the execution of the repetitious clustering algorithm. Most importantly, all methods presented good respond times in testing (i.e. providing recommendations).

B. The effect of number of clusters

In Figures 4a and 4b we observe the variation of performance for clustering with respect to k in the Choose4Greece dataset. In both metrics the performance seems to be stable with respect to k and better than the other two approaches. However, if we observe Figures 4c and 4d we note that the performance is reduced when the number of clusters increases drastically. This can be explained by the fact that with such large values of k, small clusters (clusters with only few members) will be created. Obviously, small clusters do not contain enough number of voters to comprise a sufficient block of like-minded voters. This is a rather important conclusion, since it suggests that the clustering approach is independent of the number of clusters (k) as long as it enables a sufficient number of voters at each cluster. In the case of Greece this number is close to 500.

C. The effect of clustering algorithm

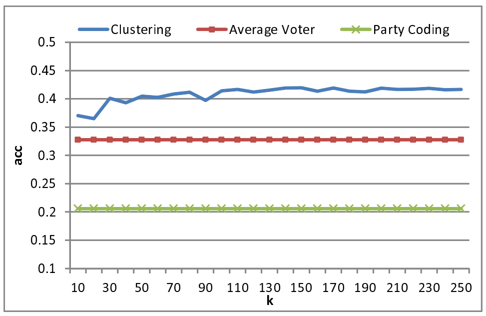
In this section we compare the two most widely used clustering algorithms in the context of the clustering algorithm presented in Section VI-C. In Figure 5a and Figure 5b we observe a comparison between kMeans and EM algorithm [52] when used for recommendation in the proposed framework. We observe that kMeans presents better performance in both metrics. This result confirms previous studies on data clustering [55].

D. Voting advice as Data Classification

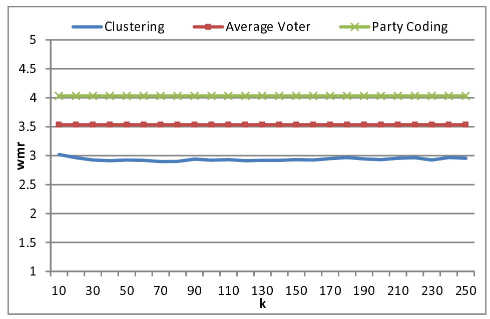
In Table VIII the predictive performance of the classifiers is displayed including training and testing time. We observe that Support Vector Machine's performance is superior in comparison with the other classifiers as well as with the clustering approach (see Table VII). Support Vector Machines have presented good classification performance in various domains [56]. Interestingly enough, the Collaborative Filtering approach presented a decent classification performance. We should underline the significant difference in predictive performance between the Party Coding method (acc 0.19) which is the current approach used in VAAs and the best

TABLE VII: Comparison of voting recommendation schemes

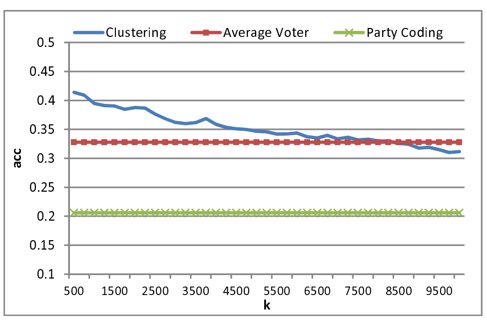
	acc	*mr*	*wmr*	*prec*	*rec*	*f*	**train (millisec)**	**test (millisec)**
Party Coding	0.19	4.18	4.36	0.29	0.22	0.25	0	143
Var	.0000	.0092	.0043	.0002	.0003	.0002		
Weighted Party Coding	0.22	4.12	4.20	0.27	0.23	0.25	0	154
Var	.0001	.0153	.0072	.0002	.0003	.0002		
Average Voter (Euclidean)	0.31	3.47	3.52	0.19	0.21	0.20	190	106
Var	.0001	.0096	.0023	.0002	.0003	.0002		
Weighted Average Voter (Euclidean)	0.31	3.54	3.56	0.19	0.20	0.20	195	114
Var	.0001	.0096	.0023	.0002	.0003	.0002		
Average Voter (Mahalanobis)	0.43	3.63	3.90	0.35	0.38	0.37	570	376
Var	.0001	.0096	.0023	.0002	.0003	.0002		
Clustering (k=200)	0.40	5.00	2.99	0.19	0.17	0.18	966835	770
Var	.0001	.0130	.0010	.0001	.0000	.0000		



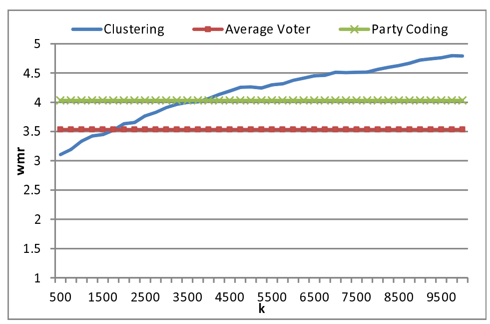
(a) Clustering: *acc* vs k (10 to 250, step: 10)



(b) Clustering: *wmr* vs k (10 to 250, step: 10)

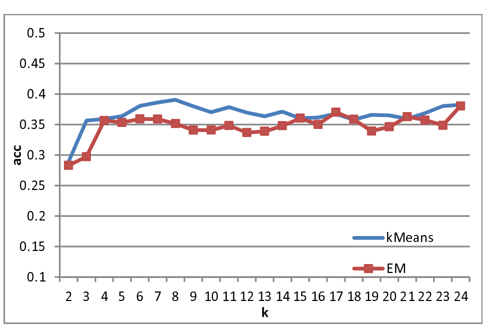


(c) Clustering: *acc* vs k (500-10000, step: 250)

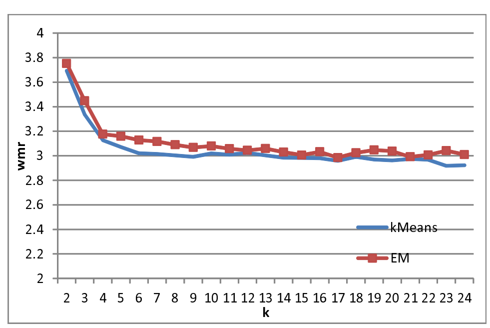


(d) Clustering: *wmr* vs k (500 to 10000, step: 250)

Fig. 4: The effect of the number of clusters



(a) kMeans vs EM - acc



(b) kMeans vs EM - wmr

Fig. 5: The effect of the clustering algorithm

Classifier (acc 0.45). In Table IX we present a comparison of the issue-based approach (Party-Coding) and a community-based approach (SMO) in all measures for all datasets. Again this table underlines the fact that community-based advice are more accurate than the issue based matching.

TABLE IX: Accuracy and Weighted Mean Rank for the issue-based approach and the community-based approach (SMO)(All datasets)

	acc	*mr*	*wmr*	*prec*	*rec*	*f*
Greece						
Party-Coding	0.19	3.80	4.36	0.29	0.22	0.25
SMO	0.45	2.47	2.70	0.29	0.27	0.28
Cyprus						
Party-Coding	0.30	3.24	2.12	0.19	0.17	0.18
SMO	0.64	1.85	1.86	0.32	0.28	0.30
Brazil						
Party-Coding	0.42	3.50	1.90	0.19	0.21	0.20
SMO	0.57	1.91	1.63	0.25	0.24	0.25
Peru						
Party-Coding	0.17	2.48	3.52	0.30	0.37	0.33
SMO	0.67	2.47	1.55	0.36	0.31	0.34
Scotland						
Party-Coding	0.40	4.14	2.06	0.22	0.21	0.22
SMO	0.61	2.10	1.75	0.34	0.27	0.30

As one can easily observe from the above tables, predictive performance of the studied approaches is far from being optimal. However, this can be explained by the following facts: a) in most cases the political parties are large in number (ranging from 5 to 15 in our datasets). This fact makes the classification problem more difficult (many classes), b) trying to predict what citizens will vote based on a questionnaire is a difficult problem since there are voters that answer identically but choose to vote for different political parties. In such cases any recommendation model is going to fail. In any case, voting advice applications are not built to predict the vote of the user but aim at advising the user. Moreover, the users of Choose4Cyprus seemed to find the social recommendation component quite useful (see next section).

E. User Evaluation

In addition to the above experimental evaluation we have launched a new feature in the latest VAA operated by our consortium that intends to measure user satisfaction for the community and issue based recommendations. This VAA was operated during the latest presidential elections in Cyprus (http://www.choose4cyprus.com). More specifically, a like-button was introduced so as the users could "like" or "dislike" the recommendation of each algorithm. There was a neutral button as well. This button appeared under the results of each algorithm.

We have calculated user satisfaction as follows:

$$\text{satisfaction} = \frac{\text{likes}}{\text{likes} + \text{neutral} + \text{dislikes}} \tag{12}$$

The satisfaction of the community-based approach was 65% while the issue-based approach was 38%. These figures indicate that the users of the VAA were more satisfied with the community-based recommendation (see Table X). Since the Social Recommendation algorithm results were presented last, the corresponding algorithm collected a smaller number of evaluations.

	Likes (Pct)	Total
Issue-based	844 (38.4%)	2196
Social Rec	121 (65.0%)	186

TABLE X: User Evaluation

F. Correlation of Recommendations

In this section we study the degree of correlation among the decisions of each of the approaches. More specifically, we investigate if the proposed approaches tend to correctly advise the same users or if there is diversity in their recommendations. If such complementarity exists, then the VAA designers should probably exploit a combination of the above approaches either to create voting advices or to model user behaviour.

In order to study this issue, we exploited a pair-wise classifier correlation metric, Yule's Q-Statistic [57]. The Q statistic between two classifiers a and b in a evaluation dataset D_{test} is equal to:

$$Q_{(a,b)}^{D_{test}} = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}} \tag{13}$$

where N^{11} is the number of times both a and b are correct, N^{00} is the number of times both classifiers are wrong, N^{01} is the number of times a is wrong and b is correct and N^{10} is the inverse case. The closer the Q-Statistic is to 1 the more correlated the two approaches are. Results on the Choose4Greece dataset can be seen in Table XI.

We observe that the most remote approach is the Party Coding method. This can be explained by the fact that the Party Coding approach relies on a different source of information. More specifically, Party Coding depends on the profiles of political parties instead of taking into consideration the vote intention of the VAA's user base.

G. Stream Classification and the Cold Start Problem

This section discusses the cold start problem in Social Voting Advice Applications. In particular we answer the following question: *How many users are required in order for the SVAA to reach the peak of its predictive performance?*

For this purpose, we consider the problem of vote recommendation as a stream classification problem. This means that the previous evaluation model of training and testing dataset does not hold. The learning method should be able to create the classification model incrementally (or even update it regularly) as users register to the system.

For this reason we experiment with two additional stream learning classifiers [58]. The first one is the Incremental Naive Bayes classifier (*INB*) which is able to update the model with each new voter. This is an example of an *online* or *incremental* type of stream learning algorithm. The second one is the Batch Support Vector Machine (*BSVM*) which is re-trained every $b = 250$ voters. This is an example of a *batch* type of stream learning algorithm. For every b users, the vote intention of the

TABLE VIII: Performance of Various Classification Methods

	acc	***mr***	***wmr***	***prec***	***rec***	***f***	**train (millisec)**	**test (millisec)**
Naive Bayes	0.40	2.83	2.66	0.36	0.35	0.36	519	882
Var	.0001	.0013	.0014	.0001	.0001	.0001		
Decision Tree	0.31	6.39	5.16	0.21	0.20	0.20	21450	67
Var	.0000	.0021	.0026	.0001	.0000	.0000		
Support Vector	0.45	4.60	2.70	0.29	0.27	0.28	268443	2211
Var	.0001	.0061	.0014	.0003	.000	.0001		
Neural Network	0.43	4.56	2.97	0.28	0.26	0.27	638837	130
Var	.0001	.0133	.0031	.0003	.0001	.0001		
Rule Learning	0.34	6.66	3.96	0.34	0.18	0.23	183388	105
Var	.0001	.0061	.0025	.0011	.0000	.0001		
Collab. Filt.	0.42	4.93	2.95	0.31	0.24	0.27	28	88839
Var	.0001	.0036	.0012	.0011	.0000	.0002		

TABLE XI: Pairwise Q-Statistic and Average Q-Statistic

	Party Coding	**Average Voter**	**Clustering**	**Naive Bayes**	**J48**	**SMO**	**NN**	**Jrip**	**Collab. Filt.**
Party Coding	1.00	**0.50**	**0.60**	**0.16**	0.63	**0.65**	**0.69**	**0.62**	**0.62**
Average Voter	0.50	1.00	0.96	0.79	**0.62**	0.86	0.83	0.67	0.79
Clustering	0.60	0.79	1.00	0.87	0.80	0.94	0.91	0.89	0.95
Naive Bayes	**0.16**	0.96	0.87	1.00	0.68	0.92	0.87	0.72	0.90
Decision Tree - J48	0.63	0.62	0.80	0.68	1.00	0.85	0.84	0.83	0.84
Support Vector Machine - SMO	0.65	0.86	0.94	0.92	0.85	1.00	0.96	0.93	0.97
Neural Network	0.69	0.83	0.91	0.87	0.84	0.96	1.00	0.88	0.93
Rule Learning - Jrip	0.62	0.67	0.89	0.72	0.83	0.93	0.88	1.00	0.94
Collaborative Filtering	0.62	0.79	0.95	0.90	0.84	0.97	0.93	0.94	1.00
AVERAGE	0.56	0.75	0.84	0.76	0.76	0.88	0.86	0.81	0.87

last b voters is considered known and hence, the classifier is re-trained with this additional information.

For the evaluation of these approaches the acc metric is not sufficient. It is now required to evaluate the variation of predictive performance with respect to users registering to the system. For this purpose we use the following metrics.

- *Rolling Accuracy - racc*: Given a window size w, at each instance $t > w$, $racc_t = acc(h, D_{(t-w,t)})$ (see section VII-A), where $D_{(t-w,t)}$ is the dataset containing instances $t-w$ to t. In other words $racc$ is the accuracy of the recommender for the last w instances and is tracked at each instance. Therefore, for a dataset D we shall obtain a *series* of Rolling Accuracies that can be represented in a graph (see graph below). Naturally in our case, instances are users registering to the system and declaring their voting intention.
- *Batch Accuracy - bacc*: Given a batch size b, $bacc$ is recorded every b instances (i.e. when $t \bmod b = 0$) and is equal to $bacc_t = acc(h, D_{(t-b,t)})$.

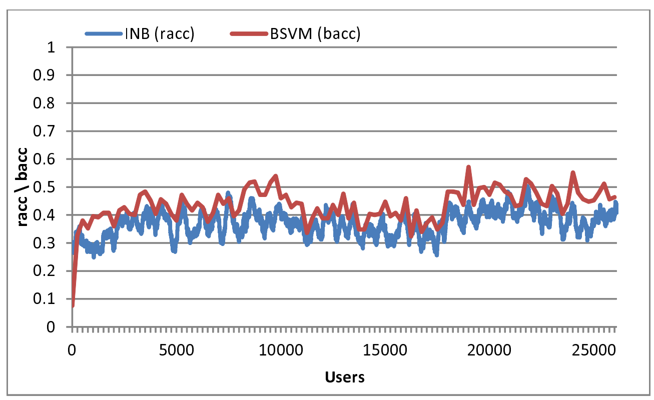


Fig. 6: Rolling Accuracy (*racc*) of Incremental Naive Bayes (*INB*) and Batch Accuracy (*bacc*) of Batch Learning Support Vector Machine (*BSVM*)

Figure 6 displays the rolling accuracy (*racc*, $w = 250$) of the Incremental Naive Bayes classifier and the batch accuracy (*bacc*, $b = 250$) for the Batch Learning Support Vector Machine (*BSVM*). The most important observation is that both classifiers reach a sufficient performance (near their average and their peak) in the first few hundreds of users. From this result we can conclude that the designers of a SVAA could rely on the recommendations of a machine-learning classifier from the first set of users. Moreover, this finding suggests that: a) the behaviour of the voters can be modeled from a small number of users and b) there is no significant variation in the behaviour of the users with the passing of time.

These findings can be supported by Figure 7 where the accuracy of a decision tree classifier (J48) is presented using various percentages of training instances (users). It is observed that in most cases even with a very small percentage of training instances (e.g. 0.1%) the classifier presents a performance similar to its peak accuracy.

The above experimental observations can be explained by the following theoretical reasoning. The sequence in which users access a VAA solely depends on the dissemination activities of the research team behind the platform. If the dissemination is not biased towards media of certain ideology or political preference, then, at time t_0 users will start entering the system from specific distributions $\Pr_{t_0}(\vec{u})$, $\Pr_{t_0}(v)$ and $\Pr_{t_0}(v|\vec{u})$. Where $\Pr(\vec{u})$ expresses the ideologies of the users,

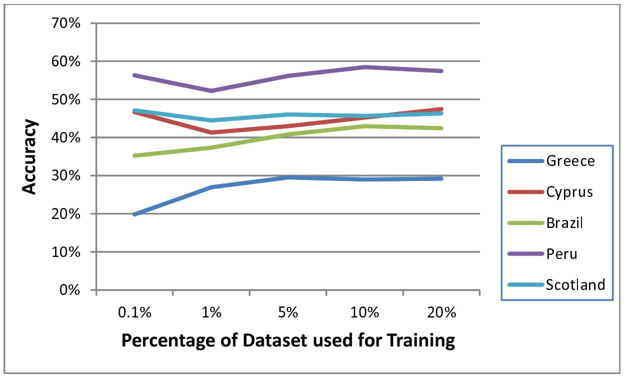


Fig. 7: Accuracy using various sizes of training datasets for all datasets using a J48 classifier

$\Pr(v)$ their vote intention distribution and $\Pr(v|\vec{u})$ the vote intention given the ideology of the user. The latter is actually the distribution that is required to be modeled in order to predict the vote intention of new users. After a certain number n_c of users entered the system the statistics will be sufficient to calculate the aforementioned probabilities. This parameter n_c depends on the number of political parties T and the number of questions M, which can vary in different VAAs. After this time point t_{n_c} and under the assumption that users will follow the same distributions (i.e. $\Pr_{(t<t_{n_c})}(v|\vec{u}) \sim \Pr_{(t>t_{n_c})}(v|\vec{u})$) then the classifier will remain accurate. However, since the above assumption might not apply, the classifier can be re-trained (e.g. through a Batch Learning Support Vector Machine) or incrementally updated (e.g. through an Incremental Naive Bayes classifier) as presented in this section.

IX. Social VAAs as Citizen Sensors

Based on our previous results, in this section we summarize the various cases where Social VAAs can operate as citizen sensors. Information recorded in a VAA can be utilized by political analysts, social scientists, political parties or any other independent organization in order to discover knowledge about the electorate's perceptions and feelings on certain issues, voting behaviour, relationships between voters and candidates as well as about many other issues.

Clustering - As we discussed in section VI-C, clustering can organize users into groups of like-minded voters. Clusters could consist of voters of one party or many. A political party studying the distribution of their voters into the different clusters could identify the different ideological groups of their voters. Moreover, a cluster where the answer "None" dominates in vote intention could imply that this cluster is not identified with any political party. Such groups are important to further be studied in the context of political instability and transformation. In addition, opinions in such clusters are known and expressed by the cluster centroids.

Dimensionality Reduction - Despite information loss, a two-dimensional representation (Section V-D, see Figures 2a and 2b) can aid in understanding groups of voters by visualizing them. In fact, political parties having a profile can be represented in the same axes as well. This could visually demonstrate the differentiation between the voters of a candidate and the candidate.

Attribute Selection - Using measures such as information gain, an analyst could identify which are the most important issues in terms of vote intention (see Section V-D). In addition, using this information, the design team of the VAA could reduce the number of questions by selecting only the ones that are strongly correlated with vote intention.

User Modeling - Many classifiers can provide with human comprehensible representation of the voter modeling. Using J48 for example, the voter's behaviour can be represented as a tree while using JRip as a set of rules.

Party Coding - This approach (see Section VI-B) presented the worst predictive performance. This can be partly explained by the fact that voters tend to base the reasons for choosing a particular candidate / party on criteria other than programmatic policy positions. Other factors such as perceived competence, charisma of the party leader or processes of socialisation affect how a prospective voter identifies with particular parties. Indeed, models of party identification based on socialization and emotional factors have provided a prominent theory of voting behaviour [59]. It is not surprising, therefore, that issue matching based on policy congruence between users and parties performs worse than community based recommendations. The latter appear to better incorporate the social factors that have informed older theories of voting as opposed to the purely programmatic factors.

Stream Classification - The sentiment of the electorate can be tracked in real-time. A statement made by a leader could change the behaviour of the voters. In this case the classifier trained from the initial data will drop in accuracy. This problem is known as concept drift [60] and supports the use of the stream learning approaches presented in Section VIII-G.

X. Conclusions

In this work we proposed Social Voting Advice Applications (SVAAs) where the user can benefit from community-based advices and features. We first formulated the problem of providing voting advice to users of VAAs and SVAAs. Next, we proposed a number of approaches that could be used for community-based vote recommendation. An analysis of how SVAAs can operate as citizen sensors was then presented. The approaches were evaluated in terms of predictive accuracy on five real VAA datasets. We made this collection available online in order to promote research in the field.

The conclusions of the evaluation study can be summarized into the following points:

- There are two main categories of voting advice. The first one is based on user-candidate distance (issue based) and the second one on voter modeling (community-based).
- Voter modeling provides in general more accurate predictions, which is largely explained by the fact that voters frequently do not agree with the policy position of the party/candidate they vote for.
- Voter modeling based on data mining classifiers, and particularly Support Vector Machines, achieved the best performance.

- The cold-start problem in VAAs is not an insurmountable obstacle since data mining classifiers can rapidly learn the limited variability of the electorate's behaviour.
- Social VAAs can operate as citizen sensors by tracking voter perceptions and feelings on issues and candidates.

By incorporating a community based voting recommendation, for the first time in real world setting, the Choose4Greece VAA (and all subsequent VAAs deployed by the Preference Matcher research team) has provided users with alternative recommendations that go beyond those solely based on the matching of policy preferences. In doing so, the approach has highlighted the importance of the social/community dimension that influences an individual's vote choice –a formulation of the problem which is closer to alternative theories of voting behaviour which current VAA design has neglected by focusing exclusively on matching users' policy preferences with party positions.

XI. Future Work

Evidently there are many interesting research directions that remain unexplored in this paper. Important issues such as data privacy in VAAs and transparency in the VAA's recommendation engine demand the proper attention from the scientific community. Another interesting research challenge is the automated collection of the party profiles by crawling the world wide web and using Natural Language Processing techniques in order to extract such information.

It is worth pointing out some of the issues that arise when well-known data mining and recommendation algorithms are applied to Voting Advice Applications. In fact, VAAs have certain specific requirements that, as far as we are aware, no single mining/recommendation scheme can confront.

- Data are in general sparse. Many missing values arise because users avoid answering all questions.
- Vote intention in most VAAs is an imbalanced attribute since usually there are some political parties that have much more voters than the smaller parties.
- There are strong correlations and patterns among the issue-questions in a VAA. A recommendation engine should take into consideration such patterns.
- Recommendation schemes should be able to adapt to changes in data (see Section VIII-G ("Stream Classification and Cold Start Problem")) either by batch or incremental learning.
- Another interesting issue is the vulnerability of the community based module to attacks of political parties that will probably try to manipulate the recommendations.

The multi-agent system (MAS)[61] paradigm can be utilized in order to model the interactions of voters within the SVAA community. Users of the system can invite others through social networks (Facebook, Twitter). In the current version of the Preference Matcher engine, interactions between users are recorder anonymously. Such information includes "who-invited-whom" tuples, which in future research could be used to identify influential users and interactions among voters. In this context, the modeling of such interactions can be implemented through the exploitation of multi-agent-systems. Finally, user evaluation of all algorithms will be continued as the Preference Matcher consortium will be establishing new Voting Advice Applications in future elections.

Acknowledgments

The authors would like to thank the anonymous reviewers and the associate editor for carefully reading the manuscript and for suggesting valuable improvements.

References

[1] The Nielsen Company, "State of the media: The Social Media Report, Q3 2011," Report, 2011.

[2] P. Norris, "Political mobilization and social networks. the example of the arab spring," *E-democracy. Electronic Democracy*, 2012.

[3] A. Ladner and J. Pianzola, "Do voting advice applications have an effect on electoral participation and voter turnout? evidence from the 2007 swiss federal elections," in *Electronic Participation*. Springer Berlin / Heidelberg, 2010, vol. 6229, pp. 211–224.

[4] I. Katakis, N. Tsapatsoulis, V. Triga, C. Tziouvas, and F. Mendez, "Clustering online poll data: Towards a voting assistance system," in *Proceedings of the 7th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP 12)*, 2012.

[5] T. Chadjipantelis, U. Serdült, and V. Triga, "Special issue on "voting advice applications and state of the art: Theory, practice, and comparative insights"," *International Journal of Electronic Governance*, 2012, (in print).

[6] Q. Liu, E. Chen, H. Xiong, C. Ding, and J. Chen, "Enhancing collaborative filtering by user interest expansion via personalized ranking," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, no. 1, pp. 218 –233, feb. 2012.

[7] L. Cedroni and D. Garzia, *Voting Advice Applications in Europe: The state of the art*. Napoli: Scriptaweb, 2010.

[8] V. Triga, U. Serdlt, and T. Chadjipadelis, "Special issue on voting advice applications and state of the art: Theory, practice, and comparative insights," *International Journal of Electronic Governance*, vol. 5, no. 3/4, 2012.

[9] J. de Graaf, "The irresistible rise of stemwijzer," in *Voting Advice Applications in Europe: The State of the Art*, L. Cedroni and D. Garzia, Eds. Napoli: Scriptaweb, 2010.

[10] A. Ladner, G. Felder, and J. Fivaz, "More than toys? a first assessment of voting advice applications in switzerland," in *Voting Advice Applications in Europe: The State of the Art*, L. Cedroni and D. Garzia, Eds. Napoli: Scriptaweb, 2010.

[11] S. Walgrave, M. Nuytemans, and K. Pepermans, "Voting aid applications and the effect of statement selection," *West European Politics*, vol. 32, no. 6, pp. 1161–1180, 2009.

[12] M. Hooghe and W. Teepe, "Party profiles on the web: an analysis of the logfiles of non-partisan interactive political internet sites in the 2003 and 2004 election campaigns in belgium," *New Media & Society*, vol. 9, no. 6, pp. 965–985, 2007.

[13] M. Wall, M. L. Sudulich, R. Costello, and E. Leon, "Picking your party online an investigation of ireland's first online voting advice application," *Information Polity*, vol. 14, no. 3, pp. 203–218, 2009.

[14] A. Perera, U. Serdlt, and Y. Welp, "Questo pblica: First voting advice application in latin america,," in *Electronic Participation: Third IFIP WG 8.5 International Conference, ePart 2011, Delft, The Netherlands, August 29 - September 1, 2011 Proceedings*, E. Tambouris, A. Macintosh, and H. de Bruijn, Eds. Heildelberg: Springer, 2011.

[15] S. Marschall, "The online making of citizens: Wahl-o-mat," in *The Making of Citizens in Europe: New Perspectives on Citizenship Education*, V. B. Georgi, Ed. Bonn: Bundeszentrale fr politische Bildung, 2008, pp. 137–41.

[16] S. Marschall and C. K. Schmidt, "The impact of voting indicators: The case of the german wahl-o-mat," in *Voting Advice Applications in Europe: The State of the Art*, L. Cedroni and D. Garzia, Eds. Napoli: Scriptaweb, 2010.

[17] K. Gemenis, O. Ruusuvirta, and M. Rosema, "The effect of voting advice applications on electoral turnout," *Electoral Studies*.

[18] F. Mendez, "Matching voters with political parties and candidates: An empirical test of four algorithms," *International Journal of Electronic Governance*, 2012, (in print).

[19] A. Ramonaite, "Voting advice applications in lithuania: Promoting programmatic competition or breeding populism?" *Policy & Internet*, vol. 2, 2010.

[20] S. Walgrave, M. Nuytemans, and K. Pepermans, "Voting aid applications and the effect of statement selection," *West European Politics*, vol. 32, no. 6, pp. 1161–1180, 2009.

[21] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. on Knowl. and Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.

[22] M. Balabanović and Y. Shoham, "Fab: content-based, collaborative recommendation," *Commun. ACM*, vol. 40, no. 3, pp. 66–72, Mar. 1997.

[23] Z.-S. Chen, J.-S. R. Jang, and C.-H. Lee, "A kernel framework for content-based artist recommendation system in music," *IEEE Transactions on Multimedia*, vol. 13, no. 6, pp. 1371–1380, 2011.

[24] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. Riedl, "Grouplens: applying collaborative filtering to usenet news," *Commun. ACM*, vol. 40, no. 3, pp. 77–87, Mar. 1997.

[25] D. Li, Q. Lv, X. Xie, L. Shang, H. Xia, T. Lu, and N. Gu, "Interest-based real-time content recommendation in online social communities," *Knowl.-Based Syst.*, vol. 28, pp. 1–12, 2012.

[26] J.-K. Min and S.-B. Cho, "Mobile human network management and recommendation by probabilistic social mining," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 41, no. 3, pp. 761 –771, june 2011.

[27] L. Terán, A. Ladner, J. Fivaz, and G. Stefani, *Using a Fuzzy-Based Cluster Algorithm for Recommending Candidates in eElections*, ser. 2. IGI Global, 01/2012 2012, ch. 6, pp. 115–138.

[28] L. Terán and A. Meier, "Smartparticipation – a fuzzy-based platform for stimulating citizens' participation," vol. 4, pp. 501 – 512, 12/2011 2011.

[29] D. Garcia, F. Mendez, U. Serdült, and F. Schweitzer, "Political polarization and popularity in online participatory media: an integrated approach," in *Proceedings of the first edition workshop on Politics, elections and data*, ser. PLEAD '12. New York, NY, USA: ACM, 2012, pp. 3–10. [Online]. Available: http://doi.acm.org/10.1145/2389661.2389665

[30] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su, "Arnetminer: extraction and mining of academic social networks," in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, ser. KDD '08. New York, NY, USA: ACM, 2008, pp. 990–998. [Online]. Available: http://doi.acm.org/10.1145/1401890.1402008

[31] X. Jin, C. Wang, J. Luo, X. Yu, and J. Han, "Likeminer: a system for mining the power of 'like' in social media networks," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, ser. KDD '11. New York, NY, USA: ACM, 2011, pp. 753–756. [Online]. Available: http://doi.acm.org/10.1145/2020408.2020528

[32] T. Lappas, K. Punera, and T. Sarlos, "Mining tags using social endorsement networks," in *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, ser. SIGIR '11. New York, NY, USA: ACM, 2011, pp. 195–204. [Online]. Available: http://doi.acm.org/10.1145/2009916.2009946

[33] C. Giatsidis, F. D. Malliaros, and M. Vazirgiannis, "Advanced graph mining for community evaluation in social networks and the web," in *Proceedings of the sixth ACM international conference on Web search and data mining*, ser. WSDM '13. New York, NY, USA: ACM, 2013, pp. 771–772. [Online]. Available: http://doi.acm.org/10.1145/2433396.2433495

[34] S. Papadopoulos, A. Vakali, and Y. Kompatsiaris, "The dynamics of content popularity in social media," *In International Journal of Data Warehousing and Mining*, vol. 6, no. 1, 2010.

[35] R.-A. Negoescu and D. Gatica-Perez, "Flickr groups: Multimedia communities for multimedia analysis," in *Internet Multimedia Search and Mining*, X.-S. Hua, M. C. Worring, and Tat-Seng, Eds. Bentham Science Publishers, 2012.

[36] M. Kim and H. Park, "Measuring twitter-based political participation and deliberation in the south korean context by using social network and triple helix indicators," *Scientometrics*, vol. 90, pp. 121–140, 2012.

[37] Y. Wu, J. Wong, Y. Deng, and K. Chang, "An exploration of social media in public opinion convergence: Elaboration likelihood and semantic networks on political events," *Dependable, Autonomic and Secure Computing, IEEE International Symposium on*, vol. 0, pp. 903–910, 2011.

[38] M. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer, "Predicting the political alignment of twitter users," in *Proceedings of 3rd IEEE Conference on Social Computing (SocialCom)*, 2011.

[39] S. P. Robertson, S. Douglas, M. Maruyama, and L.-W. Chen, "Political dialog evolution in a social network," in *Proceedings of the 13th Annual International Conference on Digital Government Research*, ser. dg.o '12. New York, NY, USA: ACM, 2012, pp. 40–48.

[40] S. Bukhari, F. Shafait, and T. Breuel, "The iupr dataset of camera-captured document images," in *Camera-Based Document Analysis and Recognition*, ser. Lecture Notes in Computer Science, M. Iwamura and F. Shafait, Eds. Springer Berlin / Heidelberg, 2012, vol. 7139, pp. 164–171.

[41] H. Yannakoudakis, T. Briscoe, and B. Medlock, "A new dataset and method for automatically grading esol texts," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, ser. HLT '11. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011, pp. 180–189.

[42] J. De Belder and M.-F. Moens, "A dataset for the evaluation of lexical simplification," in *Computational Linguistics and Intelligent Text Processing*, ser. Lecture Notes in Computer Science, A. Gelbukh, Ed. Springer Berlin / Heidelberg, 2012, vol. 7182, pp. 426–437.

[43] M. Lu, S. Bangalore, G. Cormode, M. Hadjieleftheriou, and D. Srivastava, "A dataset search engine for the research document corpus," *Data Engineering, International Conference on*.

[44] K. Gemenis, F. Mendez, and V. Triga, "Analytical outline of the calculation methodology of choose4greece," Technical Report, 2012.

[45] J. Wheatley, "Using VAAs to explore the dimensionality of the policy space: experiments from brazil, peru, scotland and cyprus," *International Journal of Electronic Governance*, vol. 5, no. 3/4, pp. 318–348, 2012.

[46] J. Wheatley, C. Carman, F. Mendez, and J. Mitchell, "The dimensionality of the scottish political space: Results from an experiment on the 2011 holyrood elections," *Party Politics*.

[47] S. C. H. Hoi, J. Luo, S. Boll, D. Xu, and R. Jin, Eds., *Social Media Modeling and Computing*. Springer, 2011.

[48] J. R. Quinlan, *C4.5: programs for machine learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.

[49] G. Forman, "An extensive empirical study of feature selection metrics for text classification," *Journal of Machine Learning Research*, vol. 3, pp. 1289–1305, 2003.

[50] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.

[51] V. Koutsonikola, A. Vakali, E. Giannakidou, and I. Kompatsiaris, "Clustering of social tagging system users: A topic and time based approach," in *Web Information Systems Engineering - WISE 2009*, ser. Lecture Notes in Computer Science, G. Vossen, D. Long, and J. Yu, Eds. Springer Berlin / Heidelberg, 2009, vol. 5802, pp. 75–86.

[52] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, vol. 39, no. 1, pp. 1–38, 1977.

[53] D. W. Aha, D. F. Kibler, and M. K. Albert, "Instance-based learning algorithms," *Machine Learning*, vol. 6, pp. 37–66, 1991.

[54] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009.

[55] B. Zhang, "Comparison of the performance of center-based clustering algorithms," in *Advances in Knowledge Discovery and Data Mining*, ser. Lecture Notes in Computer Science, K.-Y. Whang, J. Jeon, K. Shim, and J. Srivastava, Eds. Springer Berlin / Heidelberg, 2003, vol. 2637, pp. 569–569.

[56] D. Meyer, F. Leisch, and K. Hornik, "The support vector machine under test," *Neurocomputing*, vol. 55, no. 1a2, pp. 169 – 186, 2003, ¡ce:title¿Support Vector Machines¡/ce:title¿.

[57] L. I. Kuncheva and C. J. Whitaker, "Measures of diversity in classifier ensembles," *Machine Learning*, vol. 51, no. 2, 2000.

[58] I. Zliobaite and B. Gabrys, "Adaptive preprocessing for streaming data," *IEEE Transactions on Knowledge and Data Engineering, in press.*

[59] G. A. Almond and S. Verba, *The civic culture: Political attitudes and democracy in five nations*. Boston: Little, Brown & Company, 1963.

[60] I. Zliobaite, A. Bifet, G. Holmes, and B. Pfahringer, "Moa concept drift active learning strategies for streaming data," *Journal of Machine Learning Research - Proceedings Track*, vol. 17, pp. 48–55, 2011.

[61] S. Parsons and M. Wooldridge, "Game theory and decision theory in multi-agent systems," *Autonomous Agents and Multi-Agent Systems*, vol. 5, no. 3, pp. 243–254, Sep. 2002. [Online]. Available: http://dx.doi.org/10.1023/A:1015575522401