An Adaptable Legislative Advocacy and Accountability Framework

Matthew Weiden, Xifeng Yan
University of California, Santa Barbara
Goleta, CA USA
weiden@cs.ucsb.edu, xyan@cs.ucsb.edu
http://www.cs.ucsb.edu/~weiden

Abstract

Grassroots legislative advocacy groups face two fundamental challenges. The first is the challenge of quantifying the how supportive different legislators and legislative groups are of the cause in question. The second challenge is motivating legislators and their constituents to take some form of palpable action for the cause. The goal of the present work is to address these challenges by developing an automated framework for targeting advocacy efforts on determinative legislators or legislative groups and encouraging public participation by ensuring their trust in their participation's impact. Advocacy experts first provide quantitative reviews of bills on the framework's web application. Then, using this ground truth information and current voting records pulled from legislative data, a machine learning algorithm builds a model of each legislator's voting behavior. This model is then used to determine the extent to which the views of the legislator differ from the advocate and predict how the legislator will vote provided novel bill reviews. These rankings are published to a website in order to applaud supportive legislators and to hold the remaining legislators publicly accountable for being unsupportive of the cause.

1 Introduction

Political advocacy serves a special purpose in governance. At its best, advocacy serves to inform legislators and their constituencies about a good cause in hopes of inspiring positive legislative change. Constituents must be motivated to elect legislators supportive of the cause and to voice continued support so that the cause remains a priority once legislators have been elected. Complementarily, legislators must be motivated keep lines of communication open and to maintain a legislative program that reflects the concerns of their constituents.

A core challenge advocates face is to identify legislators that are unsupportive of a cause and anticipate how they will vote on future legislation relevant to it. This ensures that advocacy efforts can be targeted towards the exact groups that are most determinative in the legislative process. Further, this information can be published on the web to inform the broader public about the stance of each legislator on the cause.

Identifying determinative legislators and predicting their voting patterns is no easy task. Just the number of legislators that must be surveyed can pose a significant barrier. In the California State Legislature, for example, there are 120 legislators currently active within the Assembly and the Senate, a large number to analyze without adequate assistance. Fur-

[1]Special thanks are due to Kirsten Deshler and Monica Solorzano of the UCSB Office of Governmental Relations and to Everardo Diaz of the UCSB Political Science PhD program.

ther, canonical measures for predicting voting behavior such as the relative frequency that a legislator voted in concurrence with an advocates ideal voting record does not take into consideration the potential discriminative features of a piece of legislation. For example, at the federal level many conservative politicians might be against legislation that increases spending, and yet will often vote for legislation that increases funding for defense.

The advocacy framework presented here addresses the challenge of identifying determinative legislators and predicting their votes. First, advocates describe their position to the framework by submitting quantitative reviews of individual bills relevant to the cause and specifying their stance on each. Then, using the framework advocates can assess legislators stance on a cause relative to their own and calculate the probability that a legislator will vote for a novel bill based its quantitative review.

Importantly, the framework is adaptable to many situations. With small configuration changes it can be adapted to information from any state legislature or Congress. Even the dimensions upon which bills are evaluated can be customized. The only requirement is that the advocate must find sufficient legislation related to their cause.

The following sections of this report describe the framework and its performance in detail. The method section breaks down the organization of the software and describes the materials used. The experiment section describes the testing of the softwares practical utility using real data from California State Legislative sessions from 2010-2012.

2 Method

The framework is organized into two major components, a web interface and a backend set of recurrently scheduled jobs. The web interface provides a way for the user to input bill reviews and to view the predictions of the machine learning models once they have been computed. The recurrently scheduled jobs serve to keep the voting data on the server current, to compute models of each legislator, and to recommend new bills for review.

2.1 Web Interface

The web interface is comprised of the Bill Review, Dataset Completion, Crosscheck, Prediction, and Main pages. Each serves a specific purpose in the workflow of the site.

The Bill Review page allows the advocate to input quantitative reviews of recent legislation into the database from which the framework will determine the stance of each legislator and the advocate on the cause. On this page the advocate first selects a bill from a list of all legislation considered in the last two sessions of the specified legislative body. Upon selection, a bill quantitative review form is presented with selectors that allow the advocate to assess the bill on a number of dimensions. While the advocate can specify these dimensions, the set shown in Table 1 was used for prototyping the framework. Each of the prototype dimensions was chosen based on its ability to discriminate between political groups [5]. For each dimension the advocate specifies a number between 0 and 100 corresponding to their quantitative assessment of the bill on that dimension, for a detail of the scales used for each prototype dimension see Appendix A. Finally, the advocate identifies their position on the bill by selecting an ideal vote type from the option set Oppose, Support If Amended, Support. The resulting bill reviews and the legislatures voting records for each are saved to a database.

The Dataset Completion page serves to eliminate bias in the voting records that might negatively affect the performance of the machine learning models. Bias in legislators vote types is common–especially for bills on a specific topic–and poses a significant problem in modeling their behavior. For instance, if a given legislator favors few laws and has a high probability of voting Nay given a randomly selected bill, it will be statistically difficult to ascertain what kind of bills that they would favor. The Dataset Completion page lists in descending order the top 30 bills in the database that, if reviewed, will eliminate this bias in the dataset by maximizing vote type entropy across legislators. Similar to the Bill Review page, once one of these bills is selected from the list a review form is presented to the advocate to submit the review.

The purpose of the Crosscheck page is to make sure that the quantitative scores of each bill review are sensible relative those of other bills in each dimension. Since the quantitative scores for each bill are an abstract representation of their content rather than concrete measurements, it is the relationships between the scores for each bill that are important. On the Crosscheck page the advocate can first select one of the review dimensions from a drop-down menu. Once a dimension is selected, all of the reviewed bills will appear in a list sorted in descending order. Selecting a bill from the list will recall its review form, allowing the advocate to make adjustments in the review should they be necessary.

Once the reviews have been submitted in Bill Review, bias eliminated in Dataset Completion, and the collection of quantitative bill descriptions has been made consistent in Crosscheck, the advocate can use the Prediction page to predict how legislators will vote on novel legislation. In order to do this, the advocate submits a new, hypothetical bill review in a form identical to that on the Bill Review page and declares their position by selecting an ideal vote type. Once submitted, the framework uses a precomputed random forest model of each legislator to predict the probability that they will adopt the same position as the advocate [1]. The legislators are then listed sorted by this metric in ascending order in a pane on the same page. Unfortunately the framework is unable to accurately predict when legislators will abstain from voting at this time. Consequently, the models predict how the legislators will vote, given that they vote, and cannot accurately predict whether a bill will pass or fail. Importantly, the models still identify the relevant information for an advocate to direct his or her advocacy efforts.

Finally the Main page summarizes the random forest predictions of each legislators overall stance on the cause, commending those who are supportive and holding those unsupportive publicly accountable. Canonically, advocates have ranked legislators based on the relative frequency that their votes agreed with the advocates ideal votes. While the Main page rankings could use the same metric, the fit of each legislators average probability to agree with the advocates stance across all bills reviewed was used instead, as the model will likely better generalize to future voting behavior. Two lists of 20 legislators each list the legislators most supportive of the cause and those least supportive respectively. Each list entry displays the legislators name, party, district, and the average probability that their non-abstaining votes agreed with the ideal votes of the advocate. If specific legislators of interest are not listed, the visitor can input their address and a third list displaying the entries of their legislators will be displayed.

Cultural Locus	Economic Locus
Economic Intervention	Cultural Intervention
Government	Regulations
Taxes	Environment
Business	Labor
Religious Values	Cause Impact
Relevance to Cause	

Table 1: The dimensions listed above were used in the experiment. Alternatively the dimensions can be specified by the advocate.

In summary, the Bill Review, Dataset Completion, and Crosscheck pages aid the advocate in loading the database with consistent, unbiased bill reviews and an ideal voting record. The Prediction and Main pages allow advocates and public users to view predictions made by random forest models of the legislature. Accordingly, the framework serves two purposes: it helps advocates identify the stance of legislators relative to a cause and publishes these results to the web.

The web interface is built using the Django Web Framework [2] with HTML/CSS/JavaScript.

2.2 Backend

Since much of the remaining work related to the framework requires heavy computation that could not be implemented in real time, the backend set of machine learning and data management code is run periodically as a set of scheduled jobs. The main functions of the CRON jobs include updating local legislative data, computing legislator models, determining which bills will eliminate bias in dataset completion, and computing model accuracy.

The local database is updated weekly to insure that the frameworks metrics are current. Upon each update, JSON containing records from the last two legislative sessions is bulk downloaded from OpenStates project (Sunlight Foundation, Washington, DC). Legislator, voting, and bill information is then stripped out of the JSON and loaded into the local database for faster processing. Only the last vote in each legislative house is saved for each bill, as these votes reflect the most informed opinions of the legislators on the bill in its most refined form.With current legislative data loaded into the local database, models for each legislator can be computed.

The process starts by considering the voting record as a whole in order to smooth data for legislators with missing vote types. A k-Nearest Neighbor model relating all legislators based on this voting record is built with $k = 30$. For each legislator, if the bias in this voting record has not been eliminated and all of the votes non-abstaining votes are either yea or nay, a percentage of their abstaining votes are filled-in by the missing vote type from their 30 Nearest Voting Neighbors [3]. This smoothing method preserves the external validity of the model by replacing missing vote types with those of legislators with similar voting patterns. This strategy is better than filling in votes based on party because it takes into account that some legislators might be moderates, ideologically closer to colleagues in other parties than to the fringes of their own. Abstaining votes are then discarded from the resultant record and bill reviews corresponding to the remaining votes are selected from the database.

These bill reviews serve as samples and votes as class labels in training a random forest classifier. Given any feature vector the classifier can be leveraged in predicting the probability that the legislator will exhibit each vote type. In the framework this functionality is leveraged in two ways. First, feature vectors for all bill reviews are fed into the classifier and used to obtain the average probability that the legislator will vote in agreement with the advocate's ideal vote. While the legislator will have voted on portion of these bills, obtaining predictions for all bills will theoretically give a more comprehensive view of their stance on the cause that may better generalize to future voting behavior. This metric is used to rank legislators on their overall support of the cause. Second, legislators models can be queried with novel bill reviews on the prediction page in order to predict how they will vote under the simulated circumstances.

Classifier	Yeas	Nays	Mean
SVM: K=RBF	.90	.74	.82
SVM: K=Poly	.87	.74	.81
SVM: K=Linear	.86	.67	.76
Decision Tree	.80	.67	.73
Random Forest	.89	.73	.81
Prior	.69	.31	

Table 2: The results of testing other potential classifiers on the data. While SVMs provide better label classification, their probability estimates are unreliable.

The backend jobs were implemented using Python, Numpy/Scipy [4], and Scikit-Learn [6]. The jobs were scheduled using the standard CRON scheduler present in Unix-based operating systems.

3 Experiment

The framework was tested in a real advocacy scenario. California State Legislators were evaluated on their support of the University of California (UC) based on data from the California State Legislature and bill reviews adapted from the UC legislative program. For more background on the related issues see Appendix B.

The dependent measure in this experiment is the method used in predicting voting behavior among California State legislators. The independent measure chosen was the degree to which this voting behavior can be predicted, more specifically the Receiver Operating Characteristic Area Under the Curve (Az). Though only the most successful method will be detailed in this section, alternative methods including support vector machines (SVM), standard decision trees, and linear classifiers were also

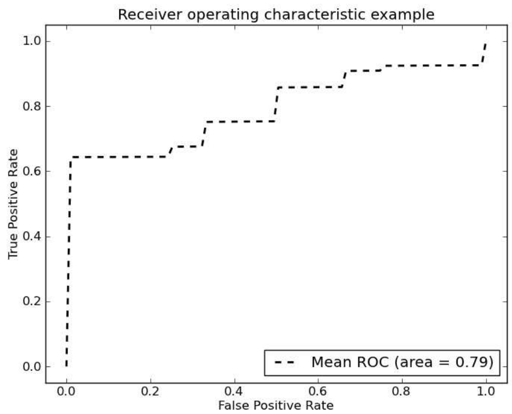

Figure 1: The Receiver Operating Characteristics of the Random Forest Classifier.

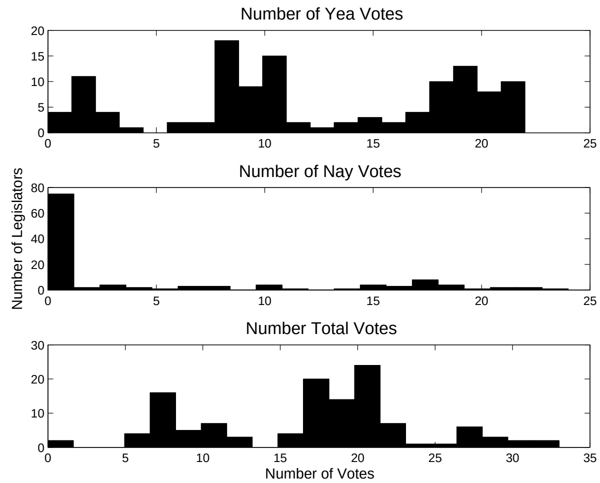

Figure 2: Histograms for the distribution of number of votes by vote type.

tested, the results of these tests are shown in Figure 2. While SVMs provided better overall classification, methods used to shoehorn probabilistic measures used in the framework from SVM resulted in poor Az scores [7].

3.1 Procedure

The California State Legislature database was first downloaded from OpenStates, deserialized and loaded into the local database. Then, the forty most recent bills were selected from the legislative program published on the University of California Office of the Presidents (UCOP) State Government Relations (SGR) website. The quantitative reviews entered in the web interface were based on detailed letters written by SGRs legislative analysts to legislators expressing the position of the UC on each bill. Bias in the resulting bill reviews was reduced using the Dataset Completion page and checked for inconsistent scoring between bills using the Crosscheck page.

3.2 Stimuli

In Experiment B, the same type and number of stimuli were presented as in Experiment A but the number of target and non-target stimuli per block was altered according to the target load condition. In addition to the single-target training sequences of the first experiment, three, five and seven target sequences were displayed to each participant in 25 sequence sets. These were used as testing data.

3.3 Results

The framework can be evaluated on its accuracy in cross-validation, its Main page rankings, and its predictive power for the Predict page.

The accuracy of the random forest classifiers is evaluated using a leave-five-out cross-validation. This method was chosen because of the limited training data available, running cross-validation with larger test set sizes would have significantly decreased the size of the training set, giving an inaccurate picture of the models effectiveness.

After 1000 Monte Carlo iterations, the framework correctly classified each legislators non-abstaining votes with an $Az = 0.79$ as shown in Figure 1. This means that each sample has a 79% chance of being correctly classified. Broken down into the component vote types, 88.93% of the yea votes and 73.26% nay votes were correctly classified. The discrepancy in accuracy between yea and nay vote types may be

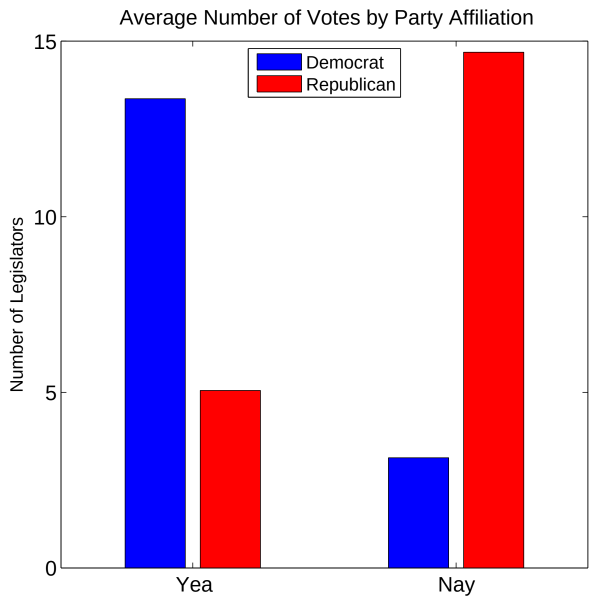

Figure 3: Vote types by party.

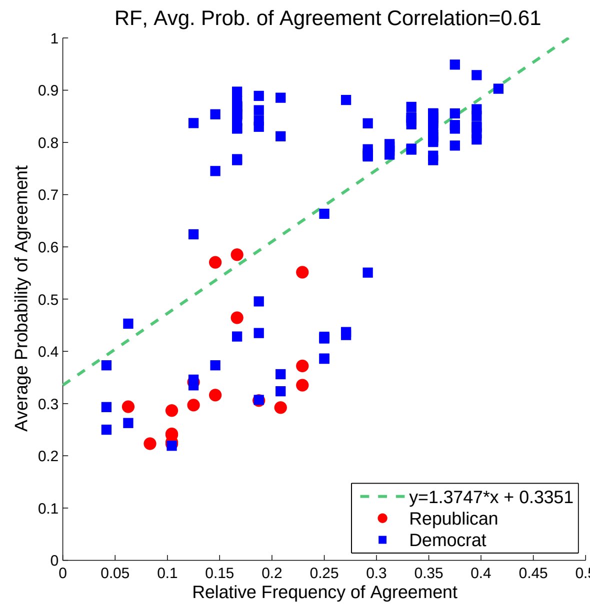

Figure 4: Correlation between Relative Frequency of Agreement and Average Probability of Agreement.

accounted for by the remaining vote type bias in the voting data. Even with significant input from dataset completion the final voting data had a mean of 12.26 yea votes and 5.03 nay votes per legislator (which was not evenly distributed, see Figure 2). More samples for the yea class likely eliminated some sampling error and led to increased classification accuracy of novel yea samples over novel nay samples.

The vote type bias is likely caused by a combination of the state legislative process and political differences between the Democratic and Republican parties. Considering the legislative process, most bills that would have a low probability of passing through the legislature–those that would draw nay votes–die or are revised in committees before they make it to the assembly or senate floor. The bills that do make it through these committees therefore have a much higher chance of garnering yea votes. Political differences between the Republican and Democratic parties are another potential cause of vote type bias. As shown in Figure 3, on average Republicans voted nay more than they voted yea, the converse is true of Democrats. In combination with the proportion of Republicans in the legislature–19/120–this meant far fewer nay votes were present in the voting dataset overall. This interaction of party type and average number of votes by type could be caused by two factors. First, the current lack of support amongst Republicans for increases in regulation, spending, and government. Second, since Democrats outnumber Republicans in the California State Legislature by six to one, most of the bills that make it to a floor vote will be authored by Democrats and unpalatable to Republicans.

The Main page rankings are evaluated on their power to predict a legislator's latent stance on the cause relative to the rest of the legislature. In the framework the legislators are ranked on the average probability that their votes would agree with the ideal votes of the advocate. As the relative frequency of each legislators votes agreeing with the advocates ideal votes has canonically been used to perform this task, it is contrasted with the average probability of agreement here, the comparison is depicted in Figure 4. Taking a Pearsons correlation reveals a very strong

Legislator	Party	APoA	Vote
Grove, Shannon L.	Rep.	0.1	Nay
Calderon, Ron	Dem.	0.167	Nay
Mansoor, Allan	Rep.	0.167	Nay
Lowenthal, Alan	Dem.	0.178	Nay
Jones, Brian W.	Rep.	0.2	Nay
Wright, Roderick D.	Dem.	0.211	Nay
Olsen, Kristin	Rep.	0.233	Nay
Wagner, Donald P.	Rep.	0.233	Nay
Yee, Leland Y.	Dem.	0.233	Nay
Achadjian, Katcho	Rep.	0.267	Nay
Corbett, Ellen M.	Dem.	0.267	Nay
Valadao, David G.	Rep.	0.267	Nay
McLeod, Gloria N.	Dem.	0.3	Nay
Norby, Chris	Dem.	0.3	Nay
Padilla, Alex	Dem.	0.3	Nay
Morrell, Mike	Rep.	0.317	Nay
DeSaulnier, Mark	Dem.	0.333	Nay
Halderman, Linda	Rep.	0.333	Nay
Jeffries, Kevin	Rep.	0.333	Nay
Leno, Mark	Dem.	0.333	Nay

Table 3: Prediction results for a novel bill that Democrats might dislike. Republicans remain on this list probably due to heavy bias towards voting no. "APoA" is the Average Probability of Agreement calculated by the model.

Legislator	Party	APoA	Vote
Grove, Shannon L.	Rep.	0.0	Nay
Blakeslee, Sam	Rep.	0.027	Nay
La Malfa, Doug	Rep.	0.033	Nay
Achadjian, Katcho	Rep.	0.067	Nay
Gaines, Beth	Rep.	0.067	Nay
Wagner, Donald P.	Rep.	0.067	Nay
Donnelly, Tim	Rep.	0.1	Nay
Halderman, Linda	Rep.	0.133	Nay
Jones, Brian W.	Rep.	0.133	Nay
Mansoor, Allan	Rep.	0.133	Nay
Olsen, Kristin	Rep.	0.133	Nay
Anderson, Joel	Rep.	0.167	Nay
Conway, Connie	Rep.	0.167	Nay
Morrell, Mike	Rep.	0.167	Nay
Berryhill, Tom	Rep.	0.2	Nay
Cook, Paul	Rep.	0.2	Nay
Dutton, Bob	Rep.	0.2	Nay
Hagman, Curt	Rep.	0.2	Nay
Harman, Tom	Rep.	0.2	Nay
Nielsen, Jim	Rep.	0.2	Nay

Table 4: Prediction results for a novel bill that Republicans might dislike.

positive relationship $r = 0.61, N = 118, p << 0.01$, but there are many ordinal differences in the lists of legislators when sorted by each metric. To determine which had more predictive power, the mean difference in place number between a ground truth ranking and the rankings of each of the metrics is taken across legislators. The ground truth ranking was based on their mean relative frequency of agreement of legislator votes in the test set agreeing with the advocates ideal votes. After 1,000 trials the each element in the average probability of agreement ordering was off by 15.33 places and the relative frequency of agreement were off by 22.85 places on average relative to the ideal ranking based on the testing data; the difference given the sample size is highly significant $F(1, 1998) = 11830.61, p << 0.01$. Though the framework uses of predictions for bills on which the legislator has not voted, the average probability of agreement is proven a better metric on which to rank legislators than the canonically used relative frequency of agreement.

Lastly, the framework can be evaluated on the speculative power of the Predict page. Given that novel bill reviews input in the Predict page are consistent with patterns in the dataset, the accuracy of the predictions will be much the same as the those in the random forest classifier accuracy evaluation. Further, a demonstration of differences between political parties is useful. Table 3 presents the results of suggesting a bill that Democrats would dislike–one with communal cultural locus, individual economic locus, stringently pro-business, anti-government, tax reductions, deregulation, and is anti-labor–is shown, and in Table 4 the results of evaluating a contrasting bill–this one with an individual cultural locus, communal economic locus, anti-business, pro-government, tax increases, regulations, and is pro-labor–designed to upset Republicans is shown. As expected, the framework predicts members of the Republican party will object to the liberal bill. For the conservative bill, however, the framework predicts that the strongest objectors will be both Republicans and Democrats, a counter-intuitive result probably due to the heavy nay vote type bias in the Republican voting record.

4 Discussion

In summary, political advocacy groups face challenges in quantifying the support of legislators for a cause, identifying the determinative groups for outreach, and publicizing the results. The present work addresses these problems with a framework that demonstrably improves on prior methods in its accuracy and level of automation.

The framework is more accurate than canonical methods. As shown in the Main page ranking analysis, the random forest classification methods used provide better assessments of legislators support for a cause than relative frequency of agreement. Consequently, the identification of the determinative legislators and their constituencies is more accurate as well.

Additionally, the framework automates this process. It is difficult to obtain an objective political profile of legislators–whose political autobiographies are often woefully inadequate–especially those in the local levels of government where press coverage is sparse. While there is no replacement for pouring over news on each legislator, it would take a significant amount of effort and a sizable staff to accurately and frequently analyze all of the members of a whole legislature. For example, suppose assessing each the California State Legislatures 120 active members took two hours once a month. The resulting workload would require a full time staff member to keep the analysis current. Comparatively, the framework can be updated just as frequently as new data becomes available, leaving the advocates free to do more detailed analysis and outreach.

One major challenge of using the framework is the sparsity of the data. The relatively small number of bill reviews and proportion of non-abstaining votes both restrict the predictive power of the classifier. Cause topics must be sufficiently broad in order to relate to a sufficient number of bills. Advocates must be diligent and review as many bills from this subset as they can to maximize the training set. Not only does the added data improve the performance of the classifier, but the proportion of legislators that vote on each bill is surprisingly low–for the dataset described in the experiment the mean number of non-abstaining votes per legislator was less than half the number of roll calls for which they were eligible. It is likely that performance of the classifier suffered greatly for lack of data in the presented experiment. The framework would likely be far more effective with additional bill reviews.

A further limitation of the framework is that it only evaluates legislators on what political scientists have termed "position-issues" and largely ignores "valance-issues". Position-issues have multiple options or a range of preferences from which legislators can choose. Comparatively, valence-issues are positive or negative issues that are merely associated with a legislator or political party. For example, taxation is a position-issue, legislators determine the level of taxation and populations to tax on continuous scales. In contrast, fighting unemployment is a valence-issue, no legislator is *against* fighting unemployment, yet some legislators make it more central to their rhetoric and legislative programs than others. As there are far too many valence-issues to take into account, they are excluded from the framework. This poses a problem for the model's accuracy as many legislative decisions are made on the basis of valence-issues [8].

Going forward I hope to expand the framework to allow anyone to select a state legislature and build their own advocacy database. This would better empower individuals to effect meaningful political change for whichever cause they should choose and increase governmental transparency by digesting incomprehensibly large amounts of data to objective analysis of our legislators.

References

[1] Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001. 10.1023/A:1010933404324. 3

[2] Django Software Foundation. Django web framework, 2005–. 3

[3] Yoshikazu Fujikawa and TuBao Ho. Cluster-based algorithms for dealing with missing values. 2336:549–554, 2002. 4

[4] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. 4

[5] David Nolan. Classifying and analyzing politico-economic systems. *The Individualist*, 1971. 2

[6] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python . *Journal of Machine Learning Research*, 12:2825–2830, 2011. 4

[7] John C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *ADVANCES IN LARGE MARGIN CLASSIFIERS*, pages 61–74. MIT Press, 1999. 5

[8] Donald E. Stokes. Spatial models of party competition. *The American Political Science Review*, 57(2):pp. 368–377, 1963. 8

A Description of Prototype Dimension Scales

Dimension	0 Rating	100 Rating
Cultural Locus	Individual	Community
Economic Locus	Individual	Community
Economic Intervention	None	Extreme Changes
Cultural Intervention	None	Extreme Changes
Government	Reductions in Programs	Increases in Programs
Regulations	Decreases in Regulation	Increases in Regulation
Taxes	Tax Decrease	Tax Increase
Environment	Environmental Degradation	Environmental Protection/Amelioration
Business	Anti-Business	Pro-Business
Labor	Anti-Labor	Pro-Labor
Religious Values	Offends Religious Groups	In Accordance with Religious Groups
Cause Impact	Negative	Positive
Relevance to Cause	None	Wholly Relevant

Table 5: Scale descriptions for the dimensions used to prototype the experiment. A rating of 50 in each dimension translates to a neutral or indeterminate rating.

B California's Disinvestment in the University of California 2008-Present

Since the global financial collapse of 2008, California's funding of its public research university system, The University of California (UC), has plummeted. Between 2008 and the present date alone the university lost $1.1 billion dollars in state funding and more cuts in the future are likely. Consequently, tuition costs have jumped and UC student tuition now comprises a larger portion of the UC Budget than state contributions. This disinvestment could result in the dismantlement of one of the worlds premier research universities and higher education becoming inaccessible to many California citizens. For more information please visit www.ucforcalifornia.org.