# Mining Democracy: Using Data Mining to Analyze a Country's Politics Vincent Etter, Julien Herzen, Matthias Grossglauser and Patrick Thiran

ÉCOLE POLYTECHNIQUE Fédérale de Lausanne

### 1. Motivation

- Governments across the world are adopting open government directives
  - Data about multiple aspects of state affairs are continuously released
- In parallel, voting advice applications (VAA's) are set up:
  - VAA's usually come in the form of a website, where politicians can answer a set of questions to express their opinion. Citizens then fill the same questionnaire and are recommended the candidates with the closest opinion.
- Using such tools and datasets, we take an exploratory, data-driven approach to the study of the political landscape of Switzerland

## 2. Datasets

We use three different, publicly available datasets:
Municipality votes

## 4. Vote Analysis at the Municipality Level

Analysis of votes at the municipality level for 245 national votes
Dimensionality reduction (SVD) highlights linguistic/cultural contrasts:



- Outcome (percentage of "yes") of all Swiss national votes between 1981 and 2011
- ► This results in 585,305 outcomes (245 votes in 2,389 municipalities)
- Smartvote pre-electoral opinions
  - Responses given by candidates and voters on the smartvote [1] VAA before the parliamentary elections of 2011 (32 questions on various political topics)
  - ▶ 2,985 candidates (82.4% of all candidates) and 229,133 voters (9% of total turnout)
- Votes in the parliament
  - ► Votes (yes/no) of the members of the National Council, from 2011 to 2013
  - ► This amounts to 2,494 votes for each of the 181 legislators, or 451,414 votes in total

## **3. Ideological Space**

- What questions discriminate best the opinions of candidates?
- The projection (SVD) of smartvote responses yields axes that are coherent with the expected left/right and liberal/conservative directions:



Despite the existence of distinct clusters and strong geographical correlations, some municipalities have a great predictive power for the vote outcomes at the national level:



Voters are more uniformly spread in the ideological space than candidates: Candidate density



Political polarization can be measured by the variance captured by the first singular vectors. Politicians are more polarized than voters:



## 5. More on VAA's: Possible Abuses and Counteractions

- Some candidates can abuse voting advice applications such as smartvote to optimize their ranking in the voting recommendations
- For instance, a simple yet efficient strategy consists in targeting a region of the ideological space that is void of candidate. Such an "artificially crafted" candidate would appear in the top-50 recommendations of nearly half of the visitors of smartvote:





However, one can also use the opinions expressed on VAAs by candidates in order to check the consistency of their votes in the parliament once they are elected

Parties overlap can be measured by computing the proportion of candidates closer to the median answer of other parties. The (i, j)-th element of the right plot shows the proportion of candidates of party i closer to the median answer of party j:



- From the opinions expressed on smartvote, it is possible to predict 80% of the votes at the parliament with an accuracy higher than 90%
- We can thus compute the **opinion shift** of a legislator by comparing her actual and expected votes. Members of the parliament are **consistent**: at most 3.75% of their votes are in opposition to their smartvote profile.

### **6.** References

# smartvote. www.smartvote.ch V. Etter, J. Herzen, M. Grossglauser, P. Thiran. Mining Democracy. submitted

### Laboratory for Communications and Applications Swiss Federal Institute of Technology in Lausanne (EPFL)