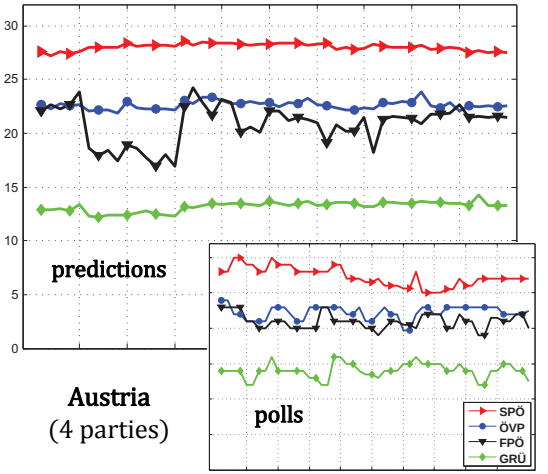
A user-centric model of voting intention from Social Media

Vasileios Lampos, Daniel Preoţiuc-Pietro & Trevor Cohn
Computer Science Department, University of Sheffield, UK

The University Of Sheffield.

Trend Miner



Austria
(4 parties)

$\boldsymbol{m}$ words

$$\begin{bmatrix} u_{p1} \\ u_{p2} \\ \vdots \\ u_{pk} \end{bmatrix}^{\mathrm{T}} \downarrow \boldsymbol{k} \text{ users} \quad \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & x_{ij} & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{km} \end{bmatrix} \begin{bmatrix} w_{p1} \\ w_{p2} \\ \vdots \\ w_{pm} \end{bmatrix}$$

x_{ij}: frequency of word j for user i during time interval t

parties n polls τ

$$\min_{\boldsymbol{W},\boldsymbol{U},\boldsymbol{\beta}} \sum_{p=1}^{n} \sum_{t=1}^{\tau} \left(\boldsymbol{u}_p^{\mathrm{T}} \boldsymbol{Q}_t \boldsymbol{w}_p + \beta_p - y_{tp} \right)^2 + \lambda_1 \sum_{j=1}^{m} \|\boldsymbol{W}_j\|_2 + \lambda_2 \sum_{j=1}^{k} \|\boldsymbol{U}_j\|_2$$

filtering out words & users

bias

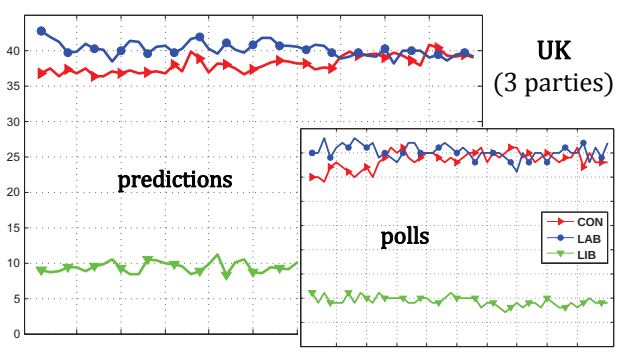
$\boldsymbol{W} = [\boldsymbol{w}_1 \dots \boldsymbol{w}_p \dots \boldsymbol{w}_n] \in \mathbb{R}^{m \times n}$
$\boldsymbol{U} = [\boldsymbol{u}_1 \dots \boldsymbol{u}_p \dots \boldsymbol{u}_n] \in \mathbb{R}^{k \times n}$
$\boldsymbol{\beta} \in \mathbb{R}^n \quad \lambda_1, \lambda_2 \in \mathbb{R}^{>0}$
$\boldsymbol{Q}_t \in \mathbb{R}^{k \times m} \quad \boldsymbol{y} \in \mathbb{R}^{\tau \times n}$

bi-linear

voting intention % for political party p during time interval t

regularisation parameter for word weights

$\ell_{2,1}$-norm
$\boldsymbol{U}_j$: j_{th} row of $\boldsymbol{U}$



UK
(3 parties)

Bi-convex iterative learning

1. Solve $\min_{\boldsymbol{W},\boldsymbol{\beta}}$ •
2. Fix $\boldsymbol{W}$ and solve $\min_{\boldsymbol{U},\boldsymbol{\beta}}$ •
3. Fix $\boldsymbol{U}$ and solve $\min_{\boldsymbol{W},\boldsymbol{\beta}}$ •
4. Validate ? Go to Step 2 : END

prediction performance

1.1K users
23K words

42K users
81K words

RMSE (%)	Method	Austria	UK
training set benchmark	mean(poll)	1.851	1.69
	Last poll	1.47	1.723
$\boldsymbol{w}$	Linear	1.442	3.067
$\boldsymbol{w}, \boldsymbol{u}$	Bilinear	1.699	1.573
$\boldsymbol{W}, \boldsymbol{U}$	Bilinear Multi-task	**1.439**	**1.478**