

# A Bilinear Model for Text Regression



The  
University  
Of  
Sheffield.

Daniel Preotiuc-Pietro

[daniel@dcs.shef.ac.uk](mailto:daniel@dcs.shef.ac.uk)

[www.preotiuc.ro](http://www.preotiuc.ro)

13.05.2013

# Linear Regression

- Standard linear regression framework:

$$f(X) = \mathbf{w}^T \mathbf{X} + \beta$$

- The optimization objective is:

$$\{\mathbf{w}^*, \beta^*\} = \operatorname{argmin} \sum_{i=1}^n (\mathbf{w}^T \mathbf{x}_i + \beta - y_i)^2 + \Psi(\mathbf{w}, \rho)$$

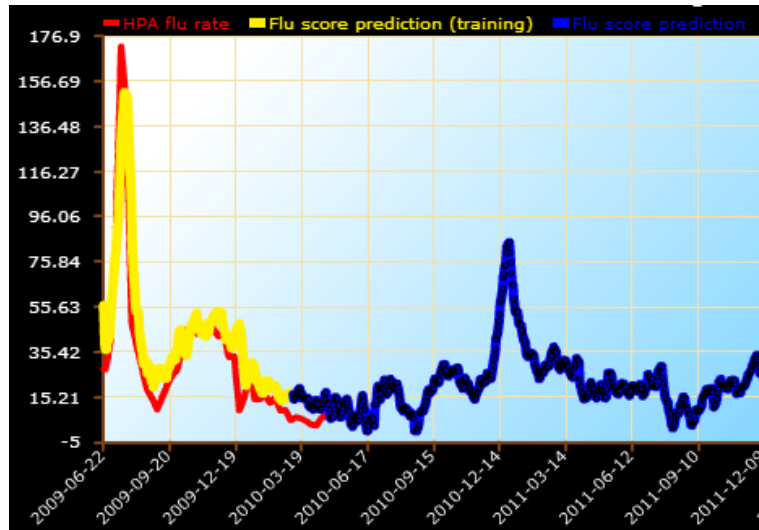
where  $\Psi$  is the regularisation (e.g. L1, L2, EN)

# Text Regression

- Task: predict real valued outputs based on textual variables (e.g. word counts)



LASSO on word counts



Lamos V., Cristianini N. (2010)

<http://geopatterns.enm.bris.ac.uk/epidemics/>

- Other examples: voting intention, financial indicators, weather, etc.

# Bilinear Regression

- Bilinear regression framework:

$$f(X) = \mathbf{u}^T X \mathbf{w} + \beta$$

- The optimization objective is:

$$\{\mathbf{w}^*, \mathbf{u}^* \beta^*\} = \operatorname{argmin} \sum_{i=1}^n (\mathbf{u}^T \mathbf{x}_i \mathbf{w} + \beta - y_i)^2 + \Psi_1(\mathbf{u}, \rho_1) + \Psi_2(\mathbf{w}, \rho_2)$$

- Goal: the model 'divides' the variables into two disjoint sets, each meant to model different aspects



# Outline

- Use case
- Motivation
- Data
- 2 models: BEN, BGL
- Learning
- Results
- Current and future work

# Trendminer project

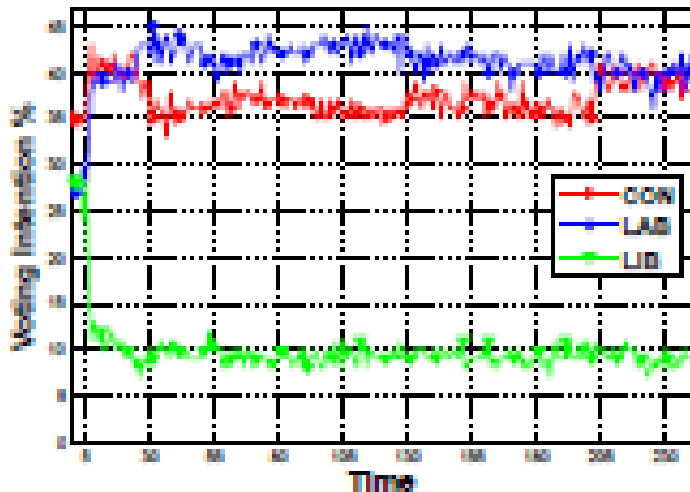
- `Large scale, cross-lingual trend mining and summarization of real time media streams`
- 7 organisations; we work with University of Southampton and SORA on machine learning
- application to predicting political polls and financial indicators

[www.trendminer-project.eu](http://www.trendminer-project.eu)

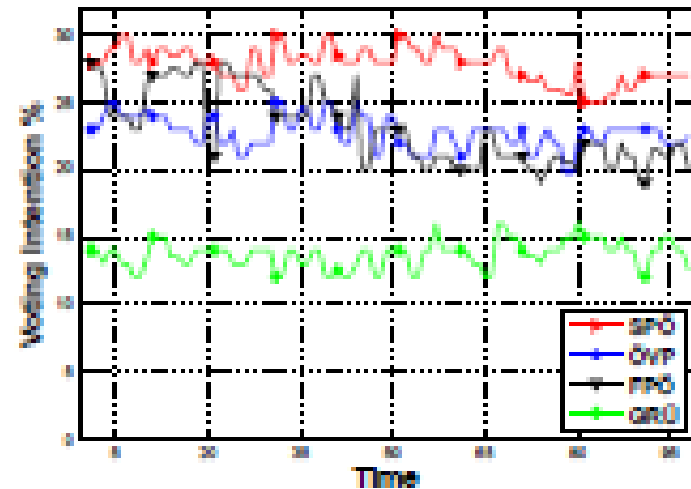


# Use case

- predicting political polls (not elections!)
- strong baselines, realistic evaluation
- 2 different use cases (U.K. and Austria)



UK polls, 04/2010 – 02/2012



Ö. polls, 01/2012 – 12/2012



# Motivation

- Twitter and real population demographics are different
- social media has biased opinions, not the most mentioned/positive sentiment party is indicative of real world trends
- more similar setup to traditional polls
- most of the users are not informative for our task and all their tweets represent noise





# Motivation

- only a few words are informative of the task
- we want to obtain a model of sparse users & sparse words
- tune based on existing polls
- regression learns weights for features without using prior knowledge, making models more portable

# Data

- collection focused on **all** the data from users of Twitter
  - 40000 U.K. (random)
    - 60 m. tweets
  - 1200 Austrian (selected by pol. scientists)
    - 800k tweets

# Model

- Bilinear predictive model:

$$f(X) = \mathbf{u}^T X \mathbf{w} + \beta$$

$\mathbf{u}$  – user weights,  $\mathbf{w}$  – word weights,  $X$  – word/user counts

- The optimization objective is:

$$\{\mathbf{w}^*, \mathbf{u}^*, \beta^*\} = \underset{\mathbf{w}, \mathbf{u}, \beta}{\operatorname{argmin}} \sum_{i=1}^n (\mathbf{u}^T \mathbf{x}_i \mathbf{w} + \beta - y_i)^2 + \Psi_1(\mathbf{u}, \boldsymbol{\rho}_1) + \Psi_2(\mathbf{w}, \boldsymbol{\rho}_2)$$

$Y$  – response variable,  $\Psi_{1,2}$  - regularisers for users and words

# Model

## **BEN (Bilinear Elastic Net)**

- Regularizers are both Elastic Nets
- a BEN model for predicting each party's score

Drawback: expect shared information between the tasks (e.g. + LAB is likely to be – CON)



# Model

- build a bilinear model that learns multiple tasks and shares strength across them
- we use the Group LASSO inside the bilinear framework
- features inside a group have to be all zero/non-zero for all the tasks
- each group is the same word/user for each task

# Model

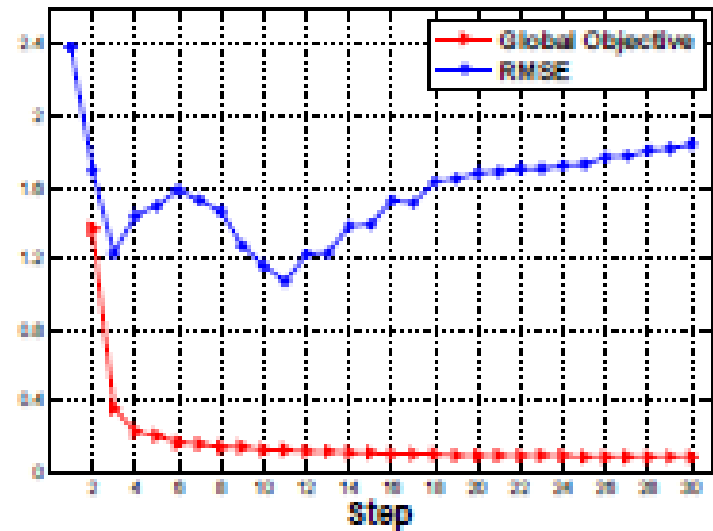
## BGL (Bilinear Group Lasso)

- the tasks are predicting each party's score
- optimisation task is:

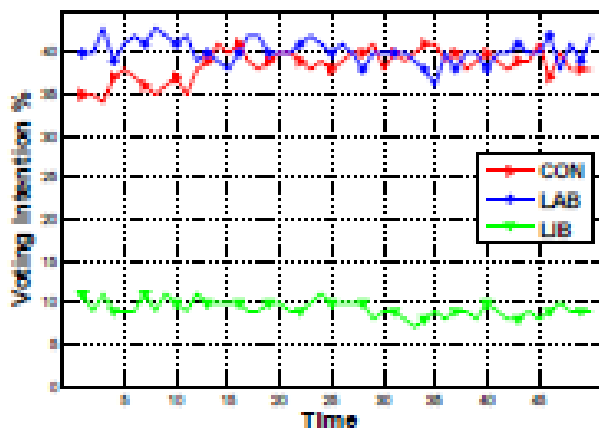
$$\{w^*, u^*, \beta^*\} = \underset{\{w, u, \beta\}}{\operatorname{argmin}} \sum_{t=1}^T \sum_{i=1}^n (\mathbf{u}_t Q_i \mathbf{w}_t + \beta_t - y_{ti})^2 + \lambda_1 \sum_{j=1}^m \|W_j\|_2 + \lambda_2 \sum_{k=1}^p \|U_k\|_2$$

# Learning

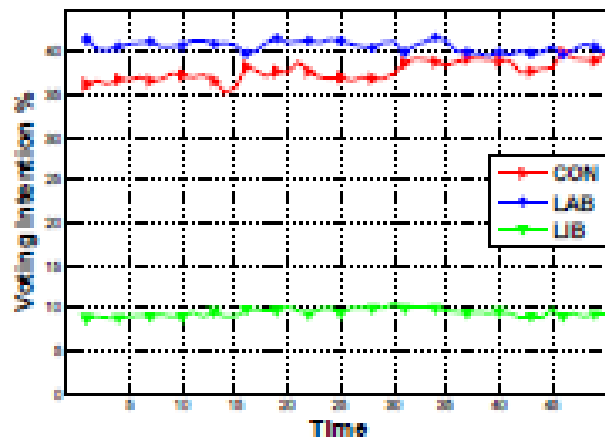
- Biconvex learning task: solved by a repeated application of 2 convex processes
- Regulariser parameters are fixed and found using grid search on validation
- Empirically choose to stop after 4 steps



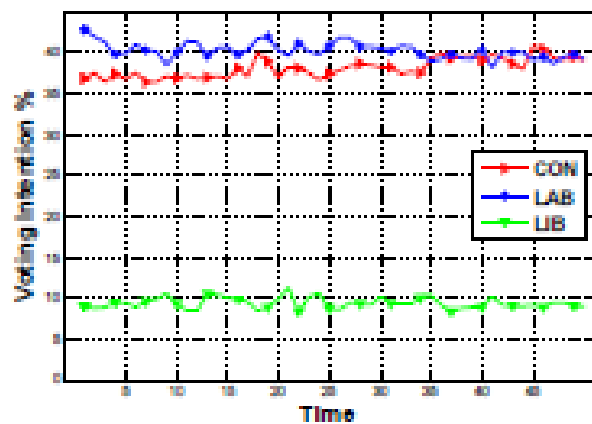
# Results – U.K.



Ground truth



BEN



BGL

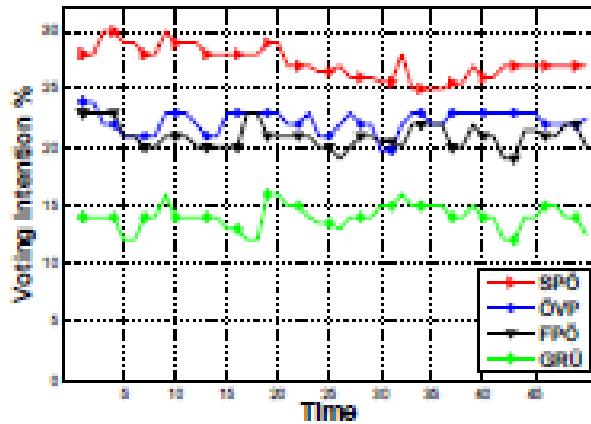
	CON	LAB	LBD	$\mu$
$B_{\mu}$	2.272	1.663	1.136	1.69
$B_{last}$	2	2.074	1.095	1.723
LEN	3.845	2.912	2.445	3.067
BEN	1.939	1.644	1.136	1.573
BGL	<b>1.785</b>	<b>1.595</b>	<b>1.054</b>	<b>1.478</b>



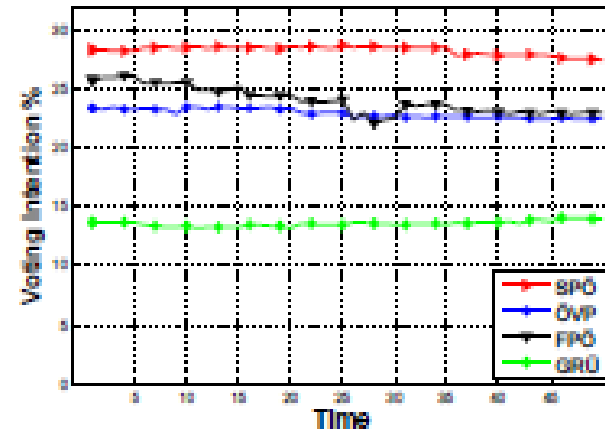
# Results – U.K.

Party	Tweet	Score	Author
<b>CON</b>	PM in friendly chat with top EU mate, Sweden's Fredrik Reinfeldt, before family photo	1.334	Journalist
	Have Liberal Democrats broken electoral rules? Blog on Labour complaint to cabinet secretary	-0.991	Journalist
<b>LAB</b>	Blog Post Liverpool: City of Radicals Website now Live <link> #liverpool #art	1.954	Art Fanzine
	I am so pleased to head Paul Savage who worked for the Labour group has been Appointed the Marketing manager for the baths hall GREAT NEWS	-0.552	Political (Labour)
<b>LBD</b>	RT @user: Must be awful for TV bosses to keep getting knocked back by all the women they ask to host election night (via @user)	0.874	LibDem MP
	Blog Post Liverpool: City of Radicals 2011 – More Details Announced #liverpool #art	-0.521	Art Fanzine

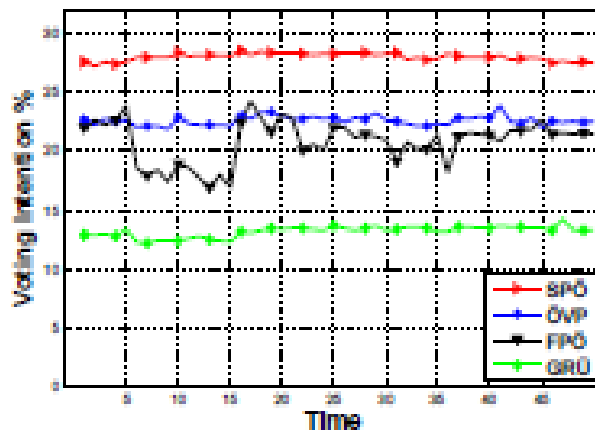
# Results – Austria



Ground truth



BEN



BGL

	SPÖ	ÖVP	FPÖ	GRÜ	$\mu$
$B_\mu$	1.535	1.373	3.3	1.197	1.851
$B_{last}$	<b>1.148</b>	1.556	<b>1.639</b>	1.536	1.47
LEN	1.291	1.286	2.039	<b>1.152</b>	1.442
BEN	1.392	1.31	2.89	1.205	1.699
BGL	1.619	<b>1.005</b>	1.757	1.374	<b>1.439</b>

# Results – Austria

Party	Tweet	Score	Author
<b>SPO</b>	Inflationsrate in Ö . im Juli leicht gesunken: von 2,2 auf 2,1%. Teurer wurde Wohnen, Wasser, Energie.	0.745	Journalist
	Hans Rauscher zu Felix #Baumgartner "A klaner Hitler" <link>	-1.711	Journalist
<b>OVP</b>	#IchPirat setze mich dafür ein, dass eine große Koalition mathematisch verhindert wird! 1.Geige: #Gruene + #FPÖe + #ÖVP	4.953	User
	kann das buch "res publica" von johannes #voggenhuber wirklich empfehlen! so zum nachdenken und so... #europa #demokratie	-2.323	User
<b>FPO</b>	Neue Kampagne der #Krone zur #Wehrpflicht: "GIB BELLO EINE STIMME!"	7.44	Political Satire
	Kampagne der Wiener SPO "zum Zusammenleben" spielt Rechtspopulisten in die Hände <link>	-3.44	Human Rights
<b>GRU</b>	Protestsong gegen die Abschaffung des Bachelor-Studiums Internationale Entwicklung: <link> #IEbleibt #unibrennt #uniwu	1.45	Student Union
	Pilz "ich will in dieser Republik weder kriminelle Asylwerber, noch kriminelle orange Politiker" - BZO" -Abschiebung ok, aber wohin? #amPunkt	-2.172	User



# Current work

- classification
- financial applications
- online implementation
- use clusters of features



# Future work

- regional analysis
- include other user features (e.g. location)
- explore other pairs of variables for different tasks
- non-stationarity

# Team



Bill Lampos  
Sheffield



Trevor Cohn  
Sheffield



Sina Samangoeei  
Southampton

# Publications

A user centric model of voting intention from  
Social Media

Lamos V., Preotiuc-Pietro D., Cohn T.  
ACL 2013, [www.preotiuc.ro](http://www.preotiuc.ro)

Regression models of trends. Tools for mining  
non-stationary data: functional prototype

Samangooei S., Lamos V., Cohn T., Gibbins N.,  
Niranjan M.

Public deliverable, [www.trendminer-project.eu](http://www.trendminer-project.eu)

Thank you !

