

# Predicting Party Affiliations from European Parliament Debates

**Bjørn Høyland**

Department of Political Science  
University of Oslo  
bjorn.hoyland@stv.uio.no

**Jean-François Godbout**

Department of Political Science  
University of Montreal  
godboutj@umontreal.ca

**Emanuele Lapponi**

Department of Informatics  
University of Oslo  
emanuel@ifi.uio.no

**Erik Velldal**

Department of Informatics  
University of Oslo  
erikve@ifi.uio.no

## Abstract

This paper documents an ongoing effort to assess whether party group affiliation of participants in European Parliament debates can be automatically predicted on the basis of the content of their speeches, using a support vector machine multi-class model. The work represents a joint effort between researchers within Political Science and Language Technology.

## 1 Introduction

The European Parliament (EP) is the directly elected parliamentary institution of the European Union (EU), elected once every five years by voters from across all 28 member states. An important arena for the political activity in the EP is the plenary sittings, the forum where all (currently 766) elected members of the European Parliament (MEPs) from all member states participate in plenary debates (in all represented languages, simultaneously translated). Our current work investigates to what extent the party affiliation of the legislators in the plenary debates can be predicted on the basis of their speeches. More specifically, the goal is to predict the party affiliation of plenary speakers in the 6<sup>th</sup> European Parliament (July 2004 – July 2009) on the basis of the party affiliations of plenary speakers in the 5<sup>th</sup> European Parliament (July 1999 – July 2004).<sup>1</sup> One

<sup>1</sup>The data have been collected from the official website of the European Parliament, where verbatim reports from each plenary sitting are published:

[www.europarl.europa.eu/plenary/en/debates.html](http://www.europarl.europa.eu/plenary/en/debates.html)

premise for the success of such an approach is that differences in ideology and belief systems are reflected in differences in choice of words in plenary debates. Another premise is that a shared belief system translates to the same choice of party group. As discussed below, systematic differences in prediction performance in the data can be used to reveal interesting differences in the extent to which these premises hold for various subgroups of MEPs. While this is further discussed in Section 4, we first describe the data sets in some more detail in Section 2 and present some preliminary results in Section 3.

## 2 Data sets and experimental set-up

The debates from the 5<sup>th</sup> EP are used for training an SVM(Cortes and Vapnik, 1995) multi-class classifier<sup>2</sup> which we then apply for predicting party affiliations in the 6<sup>th</sup> EP. We do 5-fold cross validation experiments on the 5<sup>th</sup> term for model tuning. Data points in the model correspond to speakers; participants in the debates in the EP labeled with their political party. All recorded speeches for a given speaker are conflated in a single vector. Although we can so far only report results for models using fairly basic feature types – various bag-of-words configurations based on lemmas, stems or full word-forms – combined with more linguistically informed ones,

<sup>2</sup>We use the freely available SVM<sup>multiclass</sup> toolkit, implementing the multi-class formulation of support vector machines described by Crammer and Singer (2001) with very fast optimization for linear kernels based on the algorithm for estimating Structural SVMs described by Tsochantaridis et al. (2004). For more information see; [www.cs.cornell.edu/people/tj/svm\\_light/svm\\_multiclass.html](http://www.cs.cornell.edu/people/tj/svm_light/svm_multiclass.html)

Party	MEPs
PSE	211
EPP-ED	272
ELDR	65
GUE/NGL	48
V/ALE	55
UEN	38
Total	689

Table 1: Distribution of members across the various political parties in the training data from the 5<sup>th</sup> European Parliament plenary debates.

like part-of-speech (PoS) tags and dependency relations; work is already in progress with respect to assessing the usefulness of e.g. class-based features drawn from unsupervised word clustering and modeling semantic phenomena such as negation, speculation and sentiment. Large-scale experimentation with different features sets and hyperparameters is made possible by running experiments on a large high-performance computing (HPC) cluster.

The main political groups in the European Parliament during these terms were the Christian Democratic / Conservative (EPP-ED), the Social-Democratic (PES), the Liberal (ELDR), the Green (V/ALE), the Socialists (GUE/NGL), and Right (UEN). Note that experiments only focus on the six largest political parties, excluding the smaller and more marginal ones which are often more unstable or ad-hoc configurations, including independent MEPs with various forms of Anti-EU ideologies. To give an idea of class distribution, the number of MEPs for all parties in our training set is listed in Table 1. Our 5<sup>th</sup> EP term training set comprises a total of 689 MEPs while our 6<sup>th</sup> term test set comprises 818. It is worth pointing out that in the 6<sup>th</sup> EP, roughly 75% of the data corresponds to MEPs from the old member states while 25% test are MEPs from the new member states. Of the members from the old member states, roughly 53% of the MEPs are incumbents while the remainders are freshmen (we return to this property in Section 4 below).

In order to facilitate reproducibility of reported results and foster further research, all data sets are available for download, including party labels and

	Baseline	stem	dep/stem
Acc	0.394	0.476	0.492
Prec	0.065	0.439	0.458
Rec	0.166	0.399	0.393
F <sub>1</sub>	0.094	0.418	0.423

Table 3: Results of the 5-fold cross-validation experiments on the training data for the majority-class baseline, a model trained on stems and one enriched with dependency-disambiguated stems.

all (automatic) linguistic annotations.<sup>3</sup>

### 3 Preliminary results

In addition to reporting results for the SVM classifier, we also include figures for a simple majority-class baseline approach, i.e., simply assigning all MEPs in the test set to the largest political party, EPP-ED. For evaluating the various approaches we will be reporting precision (Prec), recall (Rec) and F<sub>1</sub> for each individual class/party, in addition to the corresponding macro-averaged scores across all parties. Note that for one-of classification problems like the one we are dealing with here, micro-averaging would simply correspond to accuracy (with Prec = Rec = F<sub>1</sub>, since the number of false positives and false negatives would be the same). While we also report accuracy, it is worth bearing in mind that accuracy will overemphasize performance on the large classes when working with skewed class distributions like we do here.

In order to study the effects of different surface features and classifier-tuning, we conduct a number of 5-fold cross-validation experiments on the training data using different feature combinations, for each empirically determining the best value for the C-parameter (i.e., the regularization parameter of the SVM classifier, governing the trade-off between training error and margin size). In this initial experiments we trained different models with various configurations of non-normalized tokens (i.e., observed word forms), stems, lemmas and PoS- and dependency-disambiguated tokens and stems. The best performing configuration so far turns out to be the dependency disambiguated stems with the observed optimal C-value of 0.8, with F<sub>1</sub> over two percentage points higher than the model trained on stems alone at the same C-

<sup>3</sup>Downloadable at [http://emanuel.at.ifi.uio.no/debates\\_data.tar.gz](http://emanuel.at.ifi.uio.no/debates_data.tar.gz)

	PSE	EPP-ED	ELDR	GUE/NGL	V/ALE	UEN	total
PSE	111	36	15	9	6	1	178
EPP-ED	123	286	77	13	12	31	542
ELDR	3	3	7	0	1	0	14
GUE/NGL	2	0	1	18	1	0	22
V/ALE	3	2	2	5	25	1	38
UEN	7	9	3	1	8	4	32
total	249	336	105	46	53	37	826
Acc	0.445	0.851	0.066	0.391	0.471	0.108	<b>0.551</b>
Prec	0.623	0.527	0.500	0.818	0.657	0.166	<b>0.549</b>
Rec	0.445	0.851	0.066	0.391	0.555	0.108	<b>0.403</b>
F <sub>1</sub>	0.519	0.651	0.117	0.529	0.602	0.131	<b>0.464</b>

Table 2: Confusion matrix showing predicted (horizontal) and true (vertical) party affiliations, together with accuracy, precision, recall and F<sub>1</sub> scores for system predictions. Overall accuracy and macro-averaged precision, recall and F<sub>1</sub> (presented in bold text) can be compared to majority-class baseline results of Acc=0.410, Prec=0.068, Rec=0.166 and F<sub>1</sub>=0.097.

value point (see Table 3 for details). This indicates that linguistically informed features do provide the model with relevant information.

Results obtained by applying the best-performing configuration from our development experiments to the test data are presented in Table 2, together with a confusion matrix for the classifier assignments. Party-wise F<sub>1</sub> and accuracy scores in addition to overall accuracy and macro-averaged precision, recall and F<sub>1</sub> are shown in the bottom four rows; compare values in bold text to majority-class baseline results of Acc=0.407, Prec=0.069, Rec=0.166 and F<sub>1</sub>=0.097. There are two groups with comparatively poor prediction scores, the Liberal (ELDR) and the Right (UEN). In the case of the former, there are two key factors that may account for this: (1) Ideological compositions of the group and, (2) coalition-formation in the EP. Firstly, ELDR consists of delegations from national parties that tend to locate themselves between the Social-Democratic (PSE) and the Christian-Democratic / Conservative parties (EPP-ED) at the national arena. Due to differences in the ideological landscape across EU member states, some members of the ELDR may find themselves holding views that are closer to those held by PES or EPP-ED than ELDR representatives from some countries. Secondly, in the period under investigation, ELDR tended to form coalitions with the EPP-ED on some policy areas and with

the PES on others. As MEPs mainly speak on policies related to the Committees they serve on, misclassifications as PES or EPP-ED may be a reflection of the coalition-formation on the committees they served on (Hix and Høyland, 2013). When it comes to UEN, misclassification as EPP-ED may be explained in terms of shared ideology. In some cases, the membership of UEN rather than EPP-ED is due to historical events rather than ideological considerations (McElroy and Benoit, 2012).

## 4 Research questions

This section briefly outlines some of the questions we will be focusing on in the ongoing work of analyzing the predictions of the SVM classifier in more detail. In most cases this analysis will consist of comparing prediction performance across various relevant subsets of MEPs while looking for systematic differences.

**Contribution of linguistic features** Much of the work done to date in “text-as-data” approaches in social sciences has been based on relatively simple and surface oriented features, typically bag-of-words models, perhaps combined with term weighting and stemming for word normalization (for an overview of what is currently considered best practice, see Grimmer and Stewart (2013)). Much of the methodology can be seen as imports from the fields of information retrieval and data mining rather than natural language processing.

A relevant question for the current work is the extent to which more linguistically informed features can contribute to the task of predicting political affiliation, compared to “surface features” based solely on directly observable lexical information. One of our goals is to assess whether more accurate models can be built by including richer feature sets with more linguistically informed features, like part-of-speech (PoS) tags, dependency relations, class-based features drawn from unsupervised word clustering, negation scope analysis, and more. The preliminary results already demonstrates that linguistically motivated features can be useful for the current task, but there are still many more feature types and combinations to be explored.

**Differences between new and old member states** Ten countries joined the European Union in 2004. This offered a rare opportunity for the existing party groups to substantively increase their share of the seats in the European Parliament by recruiting national party delegations from the new member states. As most of the new member states have a relative short history of competitive multiparty system, there were weaker ties between parties in new and old member states when compared to previous rounds of enlargement. Since the allocation of office spoils in the EP is fairly proportional among party groups it was assumed that national parties from the new member states – less ideologically committed to any of the belief systems held by the traditional Western European party families – would shift the allocation of some offices in the EP by opting to join certain party group who controlled a larger share of ministerial portfolios. If there are large differences in classifier performance between members from new and old members states, this can provide support for the hypothesis that national party delegations from new member states joined the existing party groups for other reasons than simply shared ideological beliefs and goals. Høyland and Godbout (2008) presented similar preliminary results that already hint at this tendency. The ongoing collaboration will further explore this question by targeting it in new ways.

**Differences between incumbents and freshmen MEPs** This point is tightly connected to the previous. Given that our training and testing data are correspond to distinct consecutive terms of parlia-

ment, one should determine whether any differences in prediction performance for MEPs from new and old member states can be explained simply by the fact that the latter will include many MEPs that appear both in the training and the test data (i.e., speakers participating in the debates in both the 5<sup>th</sup> and 6<sup>th</sup> term). In order to factor out any such “incumbency effect”, we will also investigate whether any differences can be found in prediction performance between incumbents and “freshmen” (members who joined the EP after the 2004 elections) originating from old member states only.

**Differences between political domains** Another effect we would like to measure is whether there are any systematic differences in prediction performance across different political topics or domains. Among other things this could indicate that the language use is more politicized or ideologically charged in debates concerning certain political issues. Much of the work in the European Parliament is carried out in specialized committees that prepare reports that will later be debated and voted on in the plenary. By coupling the debates with information about which legislative standing committee has handled each particular case, we would be able to automatically break down our results according to political domain. This could be achieved using a resource like that described by Høyland et al. (2009). Examples of committee domains include Foreign Affairs, International Trade, Legal Affairs, Regional Development, Economic and Monetary Affairs, and Internal Market and Consumer Protection, to name a few. Another possibility here would be to train separate specialized classifiers for debates falling within the domain of each specialized committee directly.

## 5 Summary

This paper has outlined an interdisciplinary effort to explore whether the recorded speeches from the plenary debates of the European Parliament can be utilized by an SVM classifier to correctly predict the party affiliations of the participants. Preliminary experimental results already demonstrates that such predictions can indeed be made – also demonstrating the contribution of linguistically informed features – and the paper has outlined a number of related research questions currently being pursued in ongoing work.

## References

- [Cortes and Vapnik1995] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine Learning*, 20(3):273–297, September.
- [Crammer and Singer2001] Koby Crammer and Yoram Singer. 2001. On the algorithmic implementation of multiclass kernel-based vector machines. *Journal of Machine Learning Research*, 2:265–292, December.
- [Grimmer and Stewart2013] Justin Grimmer and Brandon M. Stewart. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3):267–297.
- [Hix and Høyland2013] Simon Hix and Bjørn Høyland. 2013. The empowerment of the european parliament. *Annual Review of Political Science*, 16:171–189.
- [Høyland and Godbout2008] Bjørn Høyland and Jean-François Godbout. 2008. Lost in translation? predicting party group affiliation from european parliament debates. In *On-line Proceedings of the Fourth Pan-European Conference on EU Politics*.
- [Høyland et al.2009] Bjørn Høyland, Indraneel Sircar, and Simon Hix. 2009. Forum section: an automated database of the european parliament. *European Union Politics*, 10(1):143–152.
- [McElroy and Benoit2012] Gail McElroy and Kenneth Benoit. 2012. Policy positioning in the european parliament. *European Union Politics*, 13(1):150–167.
- [Tsochantaridis et al.2004] Ioannis Tsochantaridis, Thomas Hofmann, Thorsten Joachims, and Yasemin Altun. 2004. Support vector machine learning for interdependent and structured output spaces. In *Proceedings of the 21st International Conference on Machine Learning*.