

# Finnish national parliament elections 2015

- Are all parties the same?

## Scope

The aim of the analysis is to find out if one can deduct the party of a candidate, based on what the candidate answers he/she wants to work for in an election machine.

## Findings

The model could make predictions with approx. 43% accuracy on the test data.

As seen on the classification report below, Vihreä liitto was the easiest one to spot, while Suomen Sosialidemokraattinen Puolue was the most difficult one. All parties, however, got f1 scores in quite similar ranges. While most parties have similar numbers in precision and recall, Vasemmistoliitto shows a larger dispersion with 50% of the correct candidates being identified but only 28% of the ones labelled as Vasemmistoliitto were correct.

	precision	recall	f1-score	support
Kansallinen Kokoomus	0.3913	0.439	0.4138	41.0
Perussuomalaiset	0.5517	0.4	0.4638	40.0
Suomen Keskusta	0.4545	0.5	0.4762	40.0
Suomen Kristillisdemokraatit (KD)	0.4286	0.4167	0.4225	36.0
Suomen Sosialidemokraattinen Puolue	0.3571	0.3659	0.3614	41.0
Vasemmistoliitto	0.5	0.2821	0.3607	39.0
Vihreä liitto	0.4138	0.6154	0.4948	39.0
accuracy	0.4312	0.4312	0.4312	0.4312
macro avg	0.4424	0.4313	0.4276	276.0
weighted avg	0.442	0.4312	0.4274	276.0

The confusion matrix below tells us that Vasemmistoliitto are often falsely classified as Vihreä liitto. Perhaps an indication that Vihreä Liitto share many of Vasemmistoliitto's themes? Suomen Sosialidemokraattinen Puolue seems to be easy to falsely classify as Vihreä Liitto.

```
[ [18  3  8  3  3  0  6]
  [ 7 16  7  3  5  1  1]
  [ 6  2 20  3  5  0  4]
  [ 6  1  2 15  5  4  3]
  [ 3  3  2  4 15  3 11]
  [ 3  4  1  5  6 11  9]
  [ 3  0  4  2  3  3 24]]
```

The six most typical words or word combinations used by candidates for each party (after stemming) are listed below. The wordings are quite well in line with what could be expected for each party.

Party	1	2	3	4	5	6
Kansallinen Kokoomus	tekemin	työnteko	työn tekemin	kannustav	satakun	työ on
Perussuomalaiset	maahanmuuto	maahanmuuttopolitiik	suome etu	ulkom	järk	pakkoruots
Suomen Keskusta	biotaloud	luonnonvar	koko suome	asu	ja työllisyys	halua kehit
Suomen Kristillisdemokraatit (KD)	kristillis	kristillis arvo	¥	miehe ja	miehe	naise
Suomen Sosialidemokraattinen Puolue	tule tuke	julkist palvelu	jokais on	reilu	maks	yhteiskun on
Vasemmistoliitto	elvyttäv	elvyttäv talouspolitiik	verotettav	talouspolitiik	investoin	tuloero
Vihreä liitto	suoje	ilmastonmuutos	ilmasto	luonto	eläint	ilmastonmuutoks

## Material

The data used for the prediction comes from [avoindata.fi](http://avoindata.fi) and consists of parliament candidates answers to yle's election machine 2015:

<https://www.avoindata.fi/data/fi/dataset/eduskuntavaalien-2015-ylen-vaalikoneen-vastaukset-ja-ehdokkaiden-taustatiedot>

In this analysis, the columns “party” and “what will you work for in the upcoming parliamentary period ?” have been used. Rows containing NA:s in these columns, parties with less than 150 candidate answers and answers in another language have been omitted.

## Methods

Special signs were removed from the “work for column” and the words were stemmed.

Bag of Words was used to find word frequencies and Multinomial Naive Bayes to create a model for party prediction.

## Limitations

A party's or politician's viewpoint in 2015 is not necessarily the same as it is today and the algorithm would need to be retrained if it was to be used with current data . Both party agendas and hot topics in society are likely to vary with time.

All possible model variations have not been tested, and more invested time in this could result in more precise predictions.