

## Results of the training the following classifiers on the bills documents

### SVM with TF•IDF

i. full text of the document

precision score	0.7159763313609467
recall score	0.8066666666666666
f1 score	0.7586206896551725

ii. randomly selected 10% of the lines of the document

precision score	0.6614583333333334
recall score	0.8466666666666667
f1 score	0.7426900584795322

iii. randomly selected 10 lines of the document

precision score	0.6864864864864865
recall score	0.8466666666666667
f1 score	0.7582089552238805

iv. randomly selected 1 line of the document

precision score	0.65625
recall score	0.84
f1 score	0.736842105263158

## Fasttext

i. full text of the document

precision score	0.6329113924050633
recall score	1.0
f1 score	0.7751937984496124

ii. randomly selected 10% of the lines of the document

precision score	0.6329113924050633
recall score	1.0
f1 score	0.7751937984496124

iii. randomly selected 10 lines of the document

precision score	0.6329113924050633
recall score	1.0
f1 score	0.7751937984496124

iv. randomly selected 1 line of the document

precision score	0.7816091954022989
recall score	0.4533333333333333
f1 score	0.5738396624472575

## Flair with Polish language model

- i. full text of the document (for randomly selected 1/20th of files from each category, 3 epochs)

precision score	0.6363636363636364
recall score	1.0
f1 score	0.7777777777777778

- ii. randomly selected 10% of the lines of the document (for randomly selected 1/6th of files from each category, 3 epochs)

precision score	0.6410256410256411
recall score	1.0
f1 score	0.7812500000000001

- iii. randomly selected 10 lines of the document

precision score	0.6586538461538461
recall score	0.9133333333333333
f1 score	0.7653631284916201

- iv. randomly selected 1 line of the document

precision score	0.6475770925110133
recall score	0.98
f1 score	0.7798408488063661