

# Index

<b>Introduction</b>	<b>4</b>
Literature review	5
Description of the problem	6
Background	6
<b>Data</b>	<b>8</b>
Case law structure	8
The procedure	8
The facts	9
The Law	9
Dataset description	10
<b>Analysis of previous works</b>	<b>12</b>
State of the art & Limits	12
<b>Natural Language processing in the legal domain - NLP techniques</b>	<b>14</b>
N-grams model	14
Using Neural Networks	15
Doc2vec model	16
Word2vec	16
PV-DM	19
PV-DBOW	21
<b>Machine learning for legal text classification</b>	<b>22</b>
Support Vector Machine	22
K-nearest neighbors - cosine similarity	24
<b>Model</b>	<b>25</b>
Structure	26
Preprocessing	26
Classification	28
SVM	28
Normalization	28
K-NN	28
Grid search	29
Bayesian comparison	31
<b>Experimental results</b>	<b>31</b>

Model 1	31
Environment description	31
Results	32
Model 2	32
Environment description	33
Results	33
Model 3	33
Environment description	33
Results	33
Model 4	33
Environment description	33
Results	33
<b>Conclusion</b>	<b>33</b>
Measuring model goodness	34
<b>References</b>	<b>36</b>