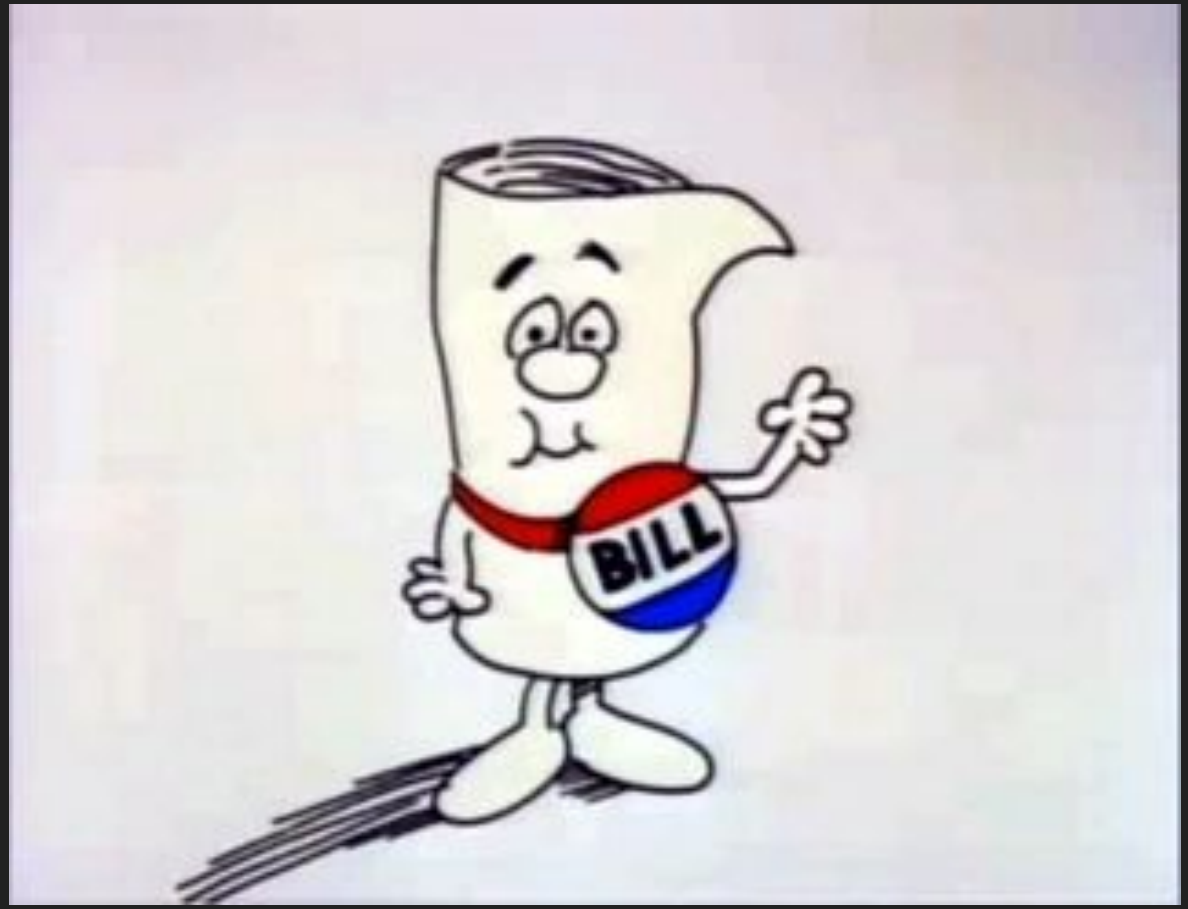
116th US Congress

Predicting the fate of a bill

Overview



Business Problem

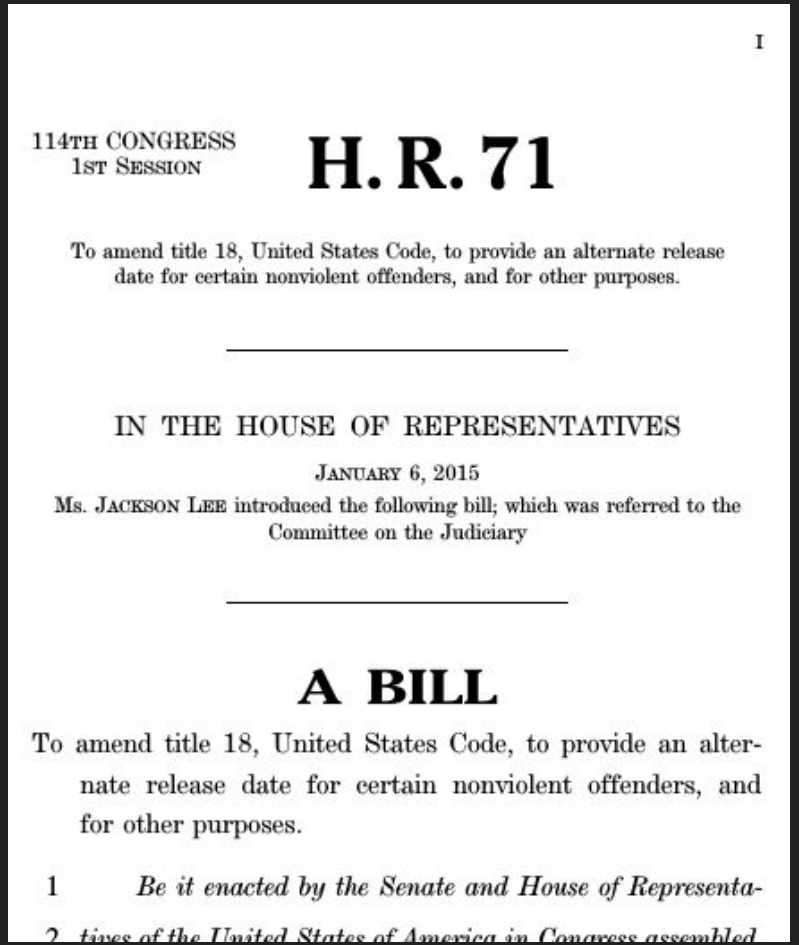
-Weeks to Write

-Limited Time

-Effects Everyone

Data Understanding

- 5,600 Proposals
- Mostly Text
- Unique Elements



I

114TH CONGRESS
1ST SESSION

H. R. 71

To amend title 18, United States Code, to provide an alternate release date for certain nonviolent offenders, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

JANUARY 6, 2015

MS. JACKSON LEE introduced the following bill; which was referred to the Committee on the Judiciary

---

A BILL

To amend title 18, United States Code, to provide an alternate release date for certain nonviolent offenders, and for other purposes.

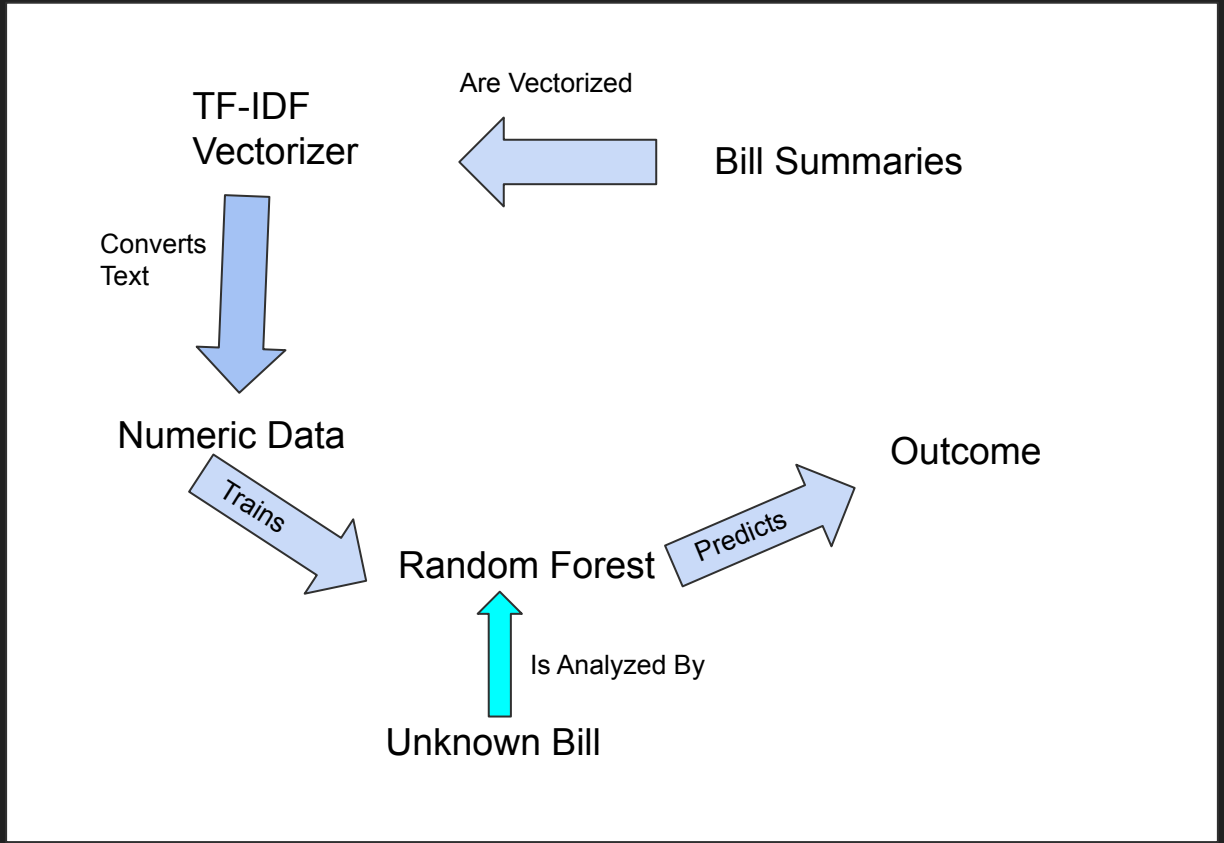
Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled

Methods

-NLP

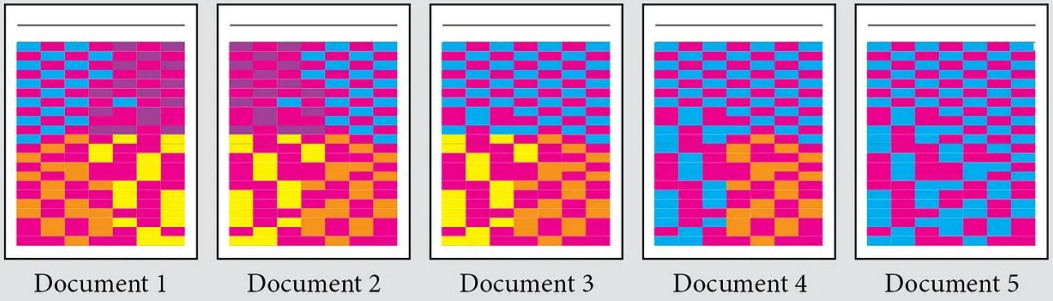
-TF-IDF Vectorizer

-Random Forest



Methods - NLP

- Simplifying of text
- Converted to Numerics
- Vectorized



Corpus TF-IDF Values						
		Terms				
Documents	1	0	.05	.027	.012	0
	2	0	.05	.027	.012	0
	3	0	0	.027	.012	0
	4	0	0	0	.012	0
	5	0	0	0	0	0

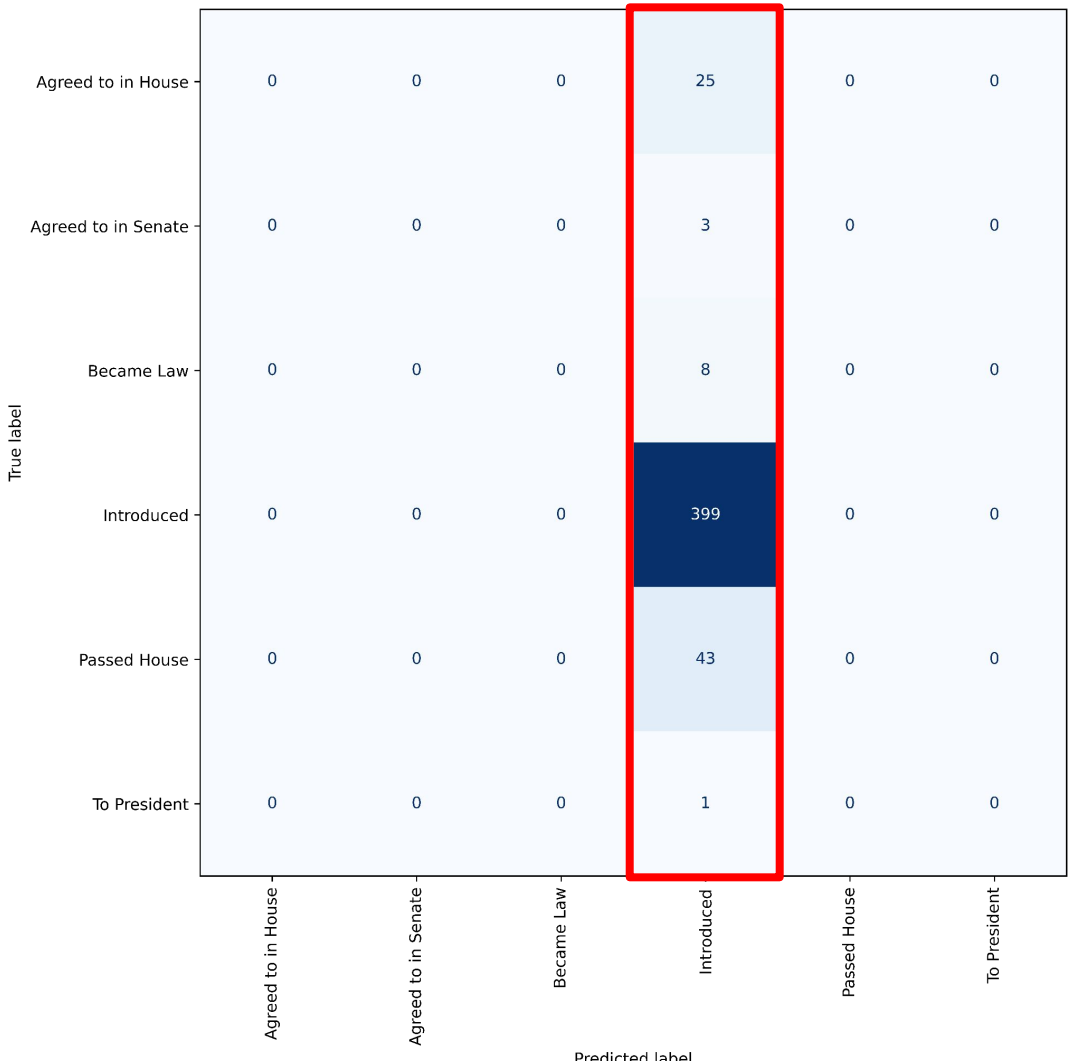
https://towardsdatascience.com/tf-idf-a-visual-explainer-and-python-implementation-on-presidential-inauguration-speeches-2a7671168550

Dummy Model

-Base Line

-Accuracy : .84

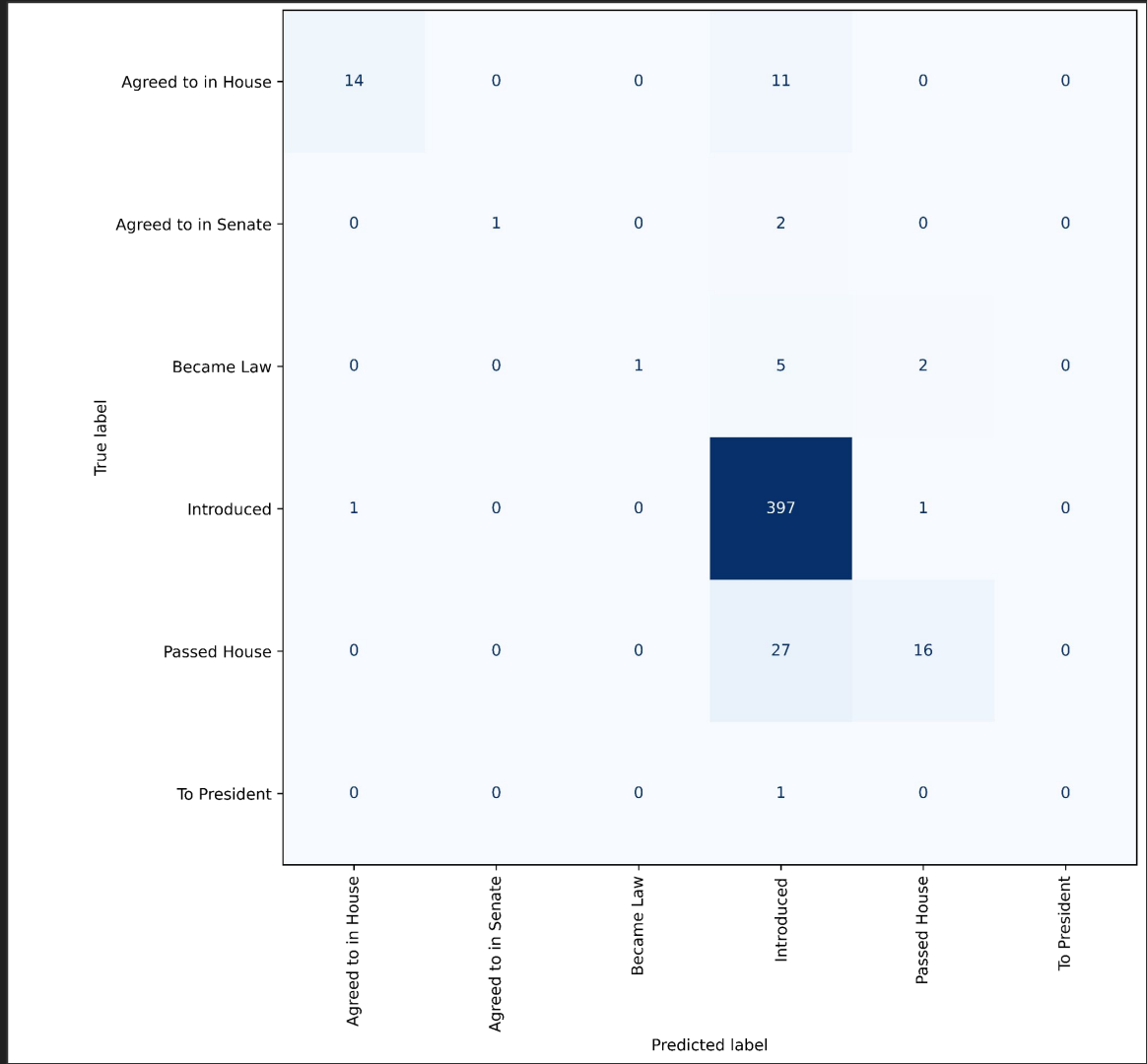
-F1 Score : .76



Final Model

Accuracy : .85

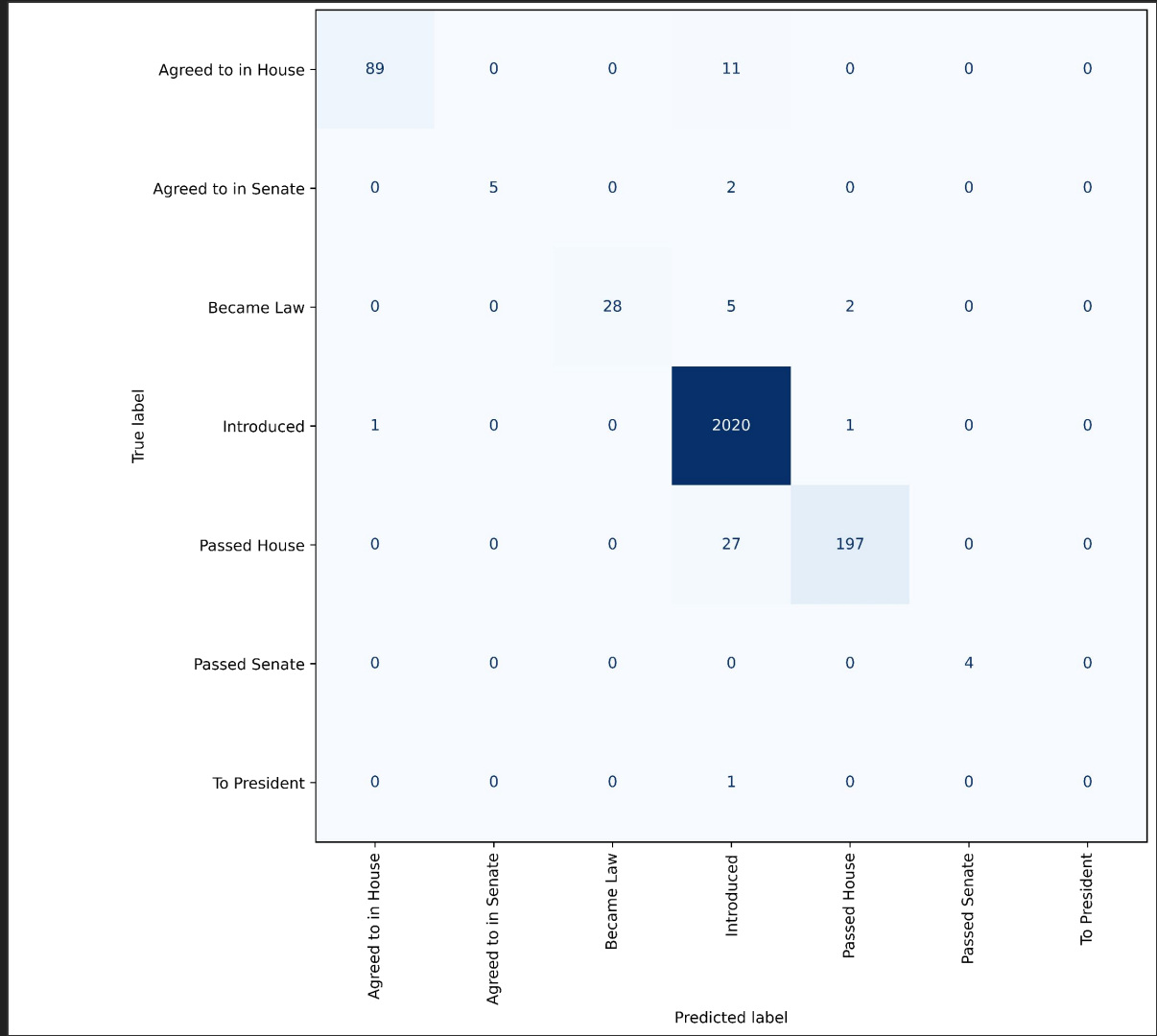
F1 Score : .85

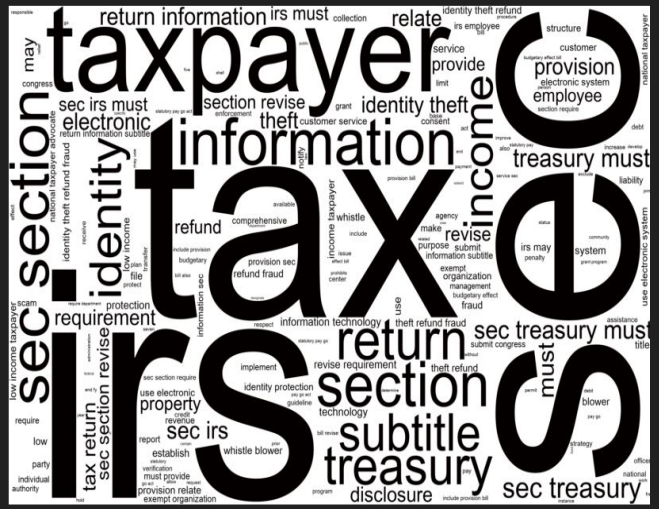


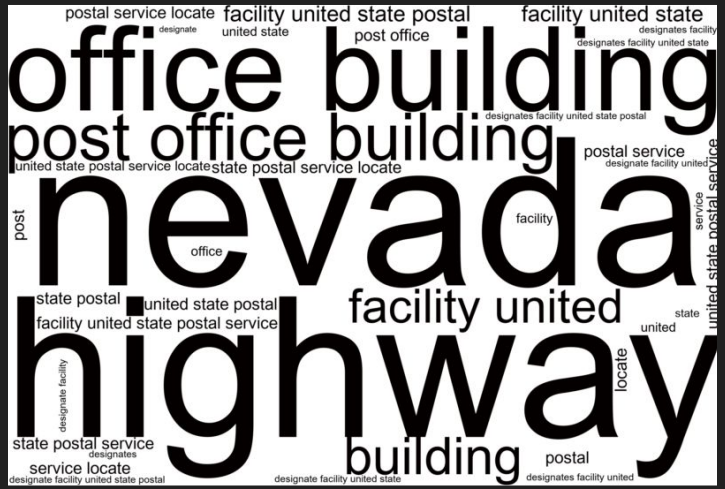
Over fitting

Accuracy : .97

F1 Score : .97




taxpayer
irs
taxes
information
sec section
identity
income
return
section
subtitle
treasury
sec treasury
provision
employee
identity theft
treasury must
sec treasury must
sec irs
electronic
refund
requirement
property
sec irs
tax return
sec section revise
sec irs must
section revise
theft
revise
must
disclosure
return information
irs must
relate
provide
structure
customer
electronic system
system
irs may
information subtitle
information technology
revise requirement
theft refund
technology
whistle blower
establish
must provide
provision relate
exempt organization
low income taxpayer
national taxpayer advocate
identity theft refund fraud
low income
national taxpayer
use electronic system
use


office building
post office building
nevada
highway
facility united
building
postal service locate
facility united state postal
facility united state
post office
designates facility united state
designates facility united state postal
postal service
designate facility united
united state postal service locate
state postal service locate
post
office
facility
state postal
united state postal
facility united state postal service
state
united
locate
state postal service
designates
service locate
designate facility united state postal
designate facility united state
postal
designates facility united

court
court must hold
court must
court must hold district
court must hold district arizona
district arizona
district
flagstaff
flagstaff yuma
flagstaff yuma list
flagstaff yuma list location
list location court must hold
yuma list location court
yuma list location court must
list location court
must hold
arizona
location
hold district
location court must hold district
hold district arizona
must hold district arizona
must
hold
location court
list location court must
must hold district
location court must holdlist location
flagstaff yuma list location court

Conclusions and Next Steps

- Not all Bad

- More Data Needed

- Social Media

Thank You

Gideon Miles
Email: milesgideon@gmail.com
Github: https://github.com/Giddybird