Predicting Legal Needs

Hee "Peter" Hwang

2020-04-16 Thu

Outline

Motivation

Data

Classifier

Prevalence Estimation

Topic

Motivation

Data

Classifier

Prevalence Estimation

Opening Question

- Posted by u/DivineJustice 10 months ago 🙃
- Moving back: Should I worry about old parking tickets?

I have (had?) two unpaid parking tickets from late 2011 and early 2012. (After the LAZ takeover.) I was too poor at the time to pay them, so I just... didn't. I do not have the car/plates that I had back then anymore.

If it matters, at the time I got both the tickets, I would have had out of state plates. (Infact, the first ticket might have been under one state, and the second under another, but same car. Both non-IL plates.)

Both tickets never managed to show up on my credit history. It's been about 7 years, so even if they were on my report, they wouldn't be anymore. In that respect paying them would be a mistake. However I want to make sure this won't turn into a legal issue.

Does anyone know if these old tickets are going to cause any issues, or can they be ignored? Has anyone been though anything similar, and of so what happened?

■ 51 Comments A Share Save A Hide Report

62% Upvoted

Questions I

- ▶ Is it a legal issue?
- ▶ What kind of legal issue is it?
- ► How to deal with it?

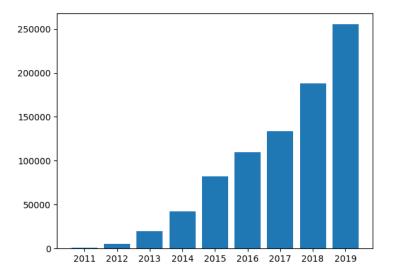
National Subject Matter Index(NSMI v2)

The NSMI Database Browse	Download	Request Addition	Search Search
Home » Browse NSMI V2			
The full NSMI V2 listing is below, click the 🖪 but	tton to see th	e subterms of the cat	ategory. The number of subterms Ⅲ can be found next to the
♣ Accidents and Torts TO-00-00-00-00 ## 5			Benefits BE-00-00-00-00 Ⅲ 8
€ Civil and Human Rights RI-00-00-00-00 ## 20			€ Courts and Lawyers co-00-00-00 ## 9
☐ Crime and Prisons CR-00-00-00-00 Ⅲ 16			Disaster Relief DI-00-00-00-00 ■8
⊞ Education ED-00-00-00-00 Ⅲ 10			Environmental Justice EN-00-00-00-00 # 7
■ Estates and Wills ES-00-00-00-00 # 7			⊕ Family FA-00-00-00 Ⅲ 18
Government Services GO-00-00-00-00 Ⅲ 3			Health HE-00-00-00-00 Ⅲ 10
Housing HO-00-00-00-0 ■8			■ Immigration IM-00-00-00 11 12
■ Money, Debt, and Consumer Issues MO-00-00	0-00-00 Ⅲ 16		■ Native American Issues and Tribal Law NA-00-00-00-00 ## 9
■ Small Business and IP BU-00-00-00-00 ## 7			Traffic and Cars TR-00-00-00 #5
Veterans and Military ∨E-00-00-00-00 10			Work and Employment Law wo-00-00-00-00 ■8

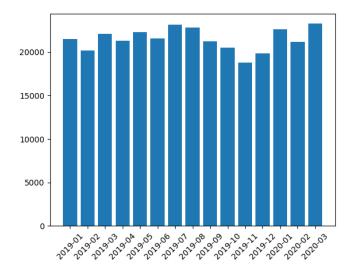
National Subject Matter Index(NSMI v2)

■ Housing HO-00-00-00-00 ## 8 Discrimination around housing HO-01-00-00-00 Eviction from a home HO-02-00-00-00 Foreclosure on a home HO-08-00-00 Housing Assistance and Subsidized Housing HO-03-00-00-00 Owning or buying a home HO-04-00-00-00 Problems with living conditions HO-05-00-00 Renting or leasing a home HO-06-00-00-00 Utilities, Energy, and Heating at home HO-07-00-00-00

Annual Submissions(2010 - 2019)



Monthly Submissions (01/2019 - 03/2020)



Questions II

- ► Can Reddit data shed a light to legal trend?
- ► Are there seasonality effects?
- ► How does society react to pandemic?

Topic

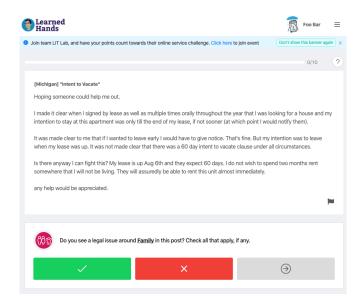
Motivation

Data

Classifier

Prevalence Estimation

Learned Hands



Learned Hands

	_id	full_text	BE-00-00-00	BE-01-00-00-00
0	5b60e59cda52255c20cff794	Will he serve time?. Long story short my broth	0.0	0.0
1	5b60e59cda52255c20cff79a	Groundwater leaking out of street 24/7. Ground	0.0	0.0
2	5b60e59cda52255c20cff7a0	How do I get my mom's license taken away. My m	0.0	0.0
3	5b60e59cda52255c20cff7bf	My boss hasn't paid me. What do i do?. I work	NaN	NaN
4	5b60e59cda52255c20cff7b8	[Texas] I signed a non-compete contract, but t	0.0	0.0
2772	5b60e66dda52255c20df433f	Do you and your parents get deported because o	0.0	0.0
2773	5b60e66dda52255c20df43ae	Wondering the legality of a minor (me) being a	0.0	0.0
2774	5b60e66dda52255c20df4462	Can I sue a billion dollar company in small cl	0.0	0.0
2775	5b60e66dda52255c20df4448	Sued by creditor and currently in settlement n	0.0	0.0
2776	5b60e66dda52255c20df4491	(CA) Sales job. Income based on performance. A	0.0	0.0

Prevalence

Class	# pos	# docs	Prevalence
HO-06	34	1662	0.0205
IM-00	36	1964	0.0183
MO-00	366	1429	0.2561
TO-00	230	1257	0.1830
TR-00	260	2006	0.1296
TR-01	22	1827	0.0120
TR-05	31	1816	0.0171
WO-00	387	1991	0.1944

Prevalence (cont.)

Class	# pos	# docs	Prevalence
BU-00	93	1590	0.0585
CO-00	106	1164	0.0911
CR-00	302	1679	0.1799
ED-00	24	1813	0.0132
ES-00	78	1944	0.0401
FA-00	357	2042	0.1748
HE-00	122	1900	0.0642
HO-00	550	2132	0.2580

Reddit Legal Advice

- ▶ http:reddit.com/r/legal_advice
- ► First submission: April 20, 2010
- \triangleright 906,693 submissions(04/2010 03/2020)
- ► 21,490 / month (01/2019 03/2020)
- ► 5,005 / week (01/2020 03/2020)
- ightharpoonup 740 / day (02/2020 03/2020)

Pushshift API

```
import requests
# Retrieve 1000 submissions with specified fields
url = "https://api.pushshift.io/reddit/search/
       submission/?subreddit=legaladvice&
       fields=id,created_utc,title,selftext&
       size=1000&after=2020-01-01&before=2020-01-31"
subs = requests.get(url)
(http:
//github.com/heeh/subreddit_downloader/sample.py)
```

Topic

Motivation

Data

Classifier

Prevalence Estimation

TF-IDF L1

- ▶ Input representation: TF-IDF Dimension: 90k 160k
- ▶ 10-fold validation
- ► Class weight: balanced
- ► Logistic regression with cross entropy loss and L1 regularization

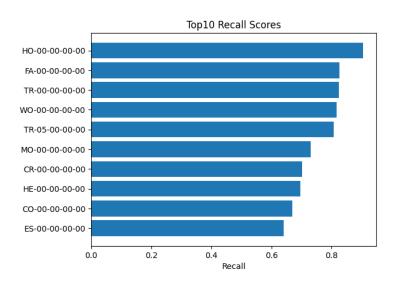
$$score(\lambda) = loss(\mathbf{x}^{de}, \mathbf{y}^{de}, \hat{\theta}) : \hat{\theta} = \underset{\theta}{arg \max} \log P_{\theta}(\mathbf{y}^{tr} | \mathbf{x}^{tr}) - \lambda |\theta|$$

• Grid search over powers of 2: $(2^{-12}, 2^{-11}, \dots, 2^{-1})$

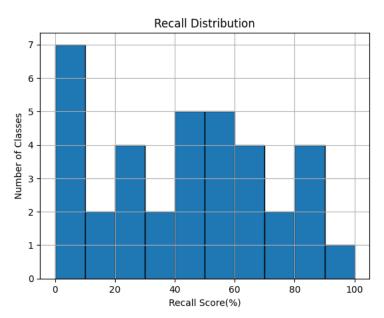
Comparison

Classifier	Acc.	Prec.	Rec.	F1	$\log_{\rm loss}$	brier
TF-IDF L1	0.97	0.52	0.41	0.46	0.0829	0.0186
TF-IDF L2	0.97	0.55	0.22	0.28	0.0759	0.0194
GloVe(50) L1	0.93	0.25	0.54	0.32	0.2049	0.0521
GloVe(50) L2	0.92	0.24	0.56	0.31	0.2081	0.0571
GloVe(300)L1	0.96	0.37	0.52	0.42	0.1086	0.0273
GloVe(300)L2	0.97	0.40	0.51	0.44	0.0968	0.0242

Recall Top 10



Recall Distribution



Input

♠ Posted by u/DivineJustice 10 months ago =

Moving back: Should I worry about old parking tickets?

I have (had?) two unpaid parking tickets from late 2011 and early 2012. (After the LAZ takeover.) I was too poor at the time to pay them, so I just... didn't. I do not have the car/plates that I had back then anymore.

If it matters, at the time I got both the tickets, I would have had out of state plates. (Infact, the first ticket might have been under one state, and the second under another, but same car. Both non-IL plates.)

Both tickets never managed to show up on my credit history. It's been about 7 years, so even if they were on my report, they wouldn't be anymore. In that respect paying them would be a mistake. However I want to make sure this won't turn into a legal issue.

Does anyone know if these old tickets are going to cause any issues, or can they be ignored? Has anyone been though anything similar, and of so what happened?

■ 51 Comments A Share Save A Hide Report

62% Upvoted

Output

Class	Prediction(%)
TR-00-00-00-00	0.9561
CO-00-00-00-00	0.6452
MO-00-00-00-00	0.3711
BU-00-00-00-00	0.0486
TO-00-00-00-00	0.0211
FA-00-00-00-00	0.0131
CR-00-00-00-00	0.0129
TR-01-00-00-00	0.0087
HO-06-00-00-00	0.0061
ED-00-00-00-00	0.0043

Topic

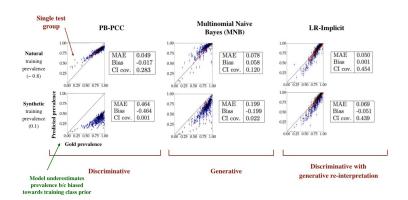
Motivation

Data

Classifier

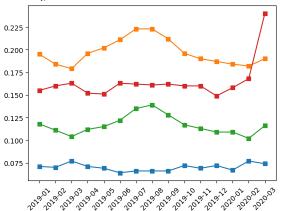
Prevalence Estimation

Freq-e



(Katherine and O'Connor, 2018)

Monthly Prevalence



HE-00-00-00-00
HO-00-00-00
HO-06-00-00-00
WO-00-00-00

▶ WO-00: Work and Employment Law

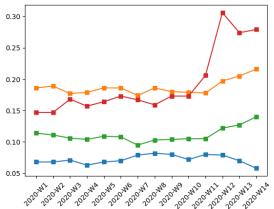
► HO-00: Housing

► HO-06: Renting or leasing a home

► HE-00: Health



Weekly Prevalence





▶ WO-00: Work and Employment Law

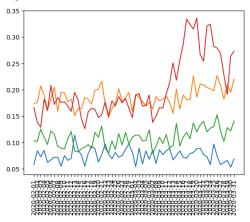
► HO-00: Housing

► HO-06: Renting or leasing a home

► HE-00: Health



Daily Prevalence



HE-00-00-00-00
HO-00-00-00
HO-06-00-00-00
WO-00-00-00-00

▶ WO-00: Work and Employment Law

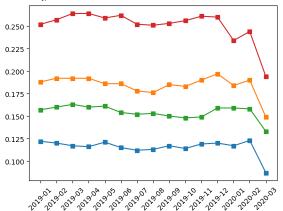
► HO-00: Housing

► HO-06: Renting or leasing a home

► HE-00: Health



Monthly Prevalence



CO-00-00-00-00
CR-00-00-00
FA-00-00-00-00
TO-00-00-00

► TO-00: Accidents and Torts

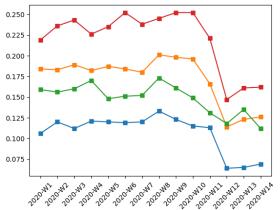
► CR-00: Crime and Prisons

► FA-00: Family

► CO-00: Courts and Lawyers



Weekly Prevalence





► TO-00: Accidents and Torts

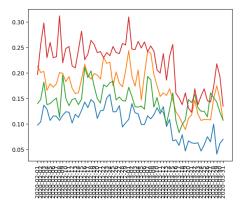
► CR-00: Crime and Prisons

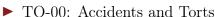
► FA-00: Family

► CO-00: Courts and Lawyers



Daily Prevalence





- ► CR-00: Crime and Prisons
- ► FA-00: Family
- ► CO-00: Courts and Lawyers





Summary

- ▶ Provide people better insight regarding their issues by classification
- ► Shows seasonality
- ► How natural disaster affects legal needs

Todo

- ► More labeled data(Currently 16 classes available)
- ▶ Dataset from other communities
- ► California statewide legal help portal (https://lawhelpca.org/)
- ► System to automate and improve the model

Reference

▶ Keith, K., & O'Connor, B. (2018). Uncertainty-aware generative models for inferring document class prevalence. In Proceedings of EMNLP.

End

Thank you for listening!