

Predicting Patent Rejections with NLP

John Herr



About Me:

B.S. Mechanical Engineering
M.S. Electrical Engineering

US Patent Bar
Patent agent from 2016 - 2019

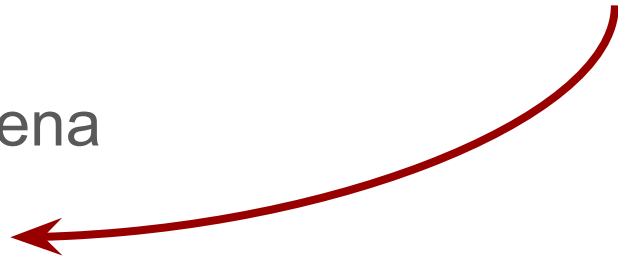


What can NOT be patented?

United States Code 101:

SUBJECTIVE!

- Laws of nature
- Natural phenomena
- **Abstract ideas**





Example Claim

1. A system for determining a direct memory access (DMA) transfer mode...*[lots of boring text]*

...the DMA transfer mode being one of **a data copy mode** and **a map/umap mode**.

ABSTRACT IDEA

Example Claim



... the DMA transfer mode being one of a ~~data copy~~ first mode and a ~~map/unmap~~ or a second mode,

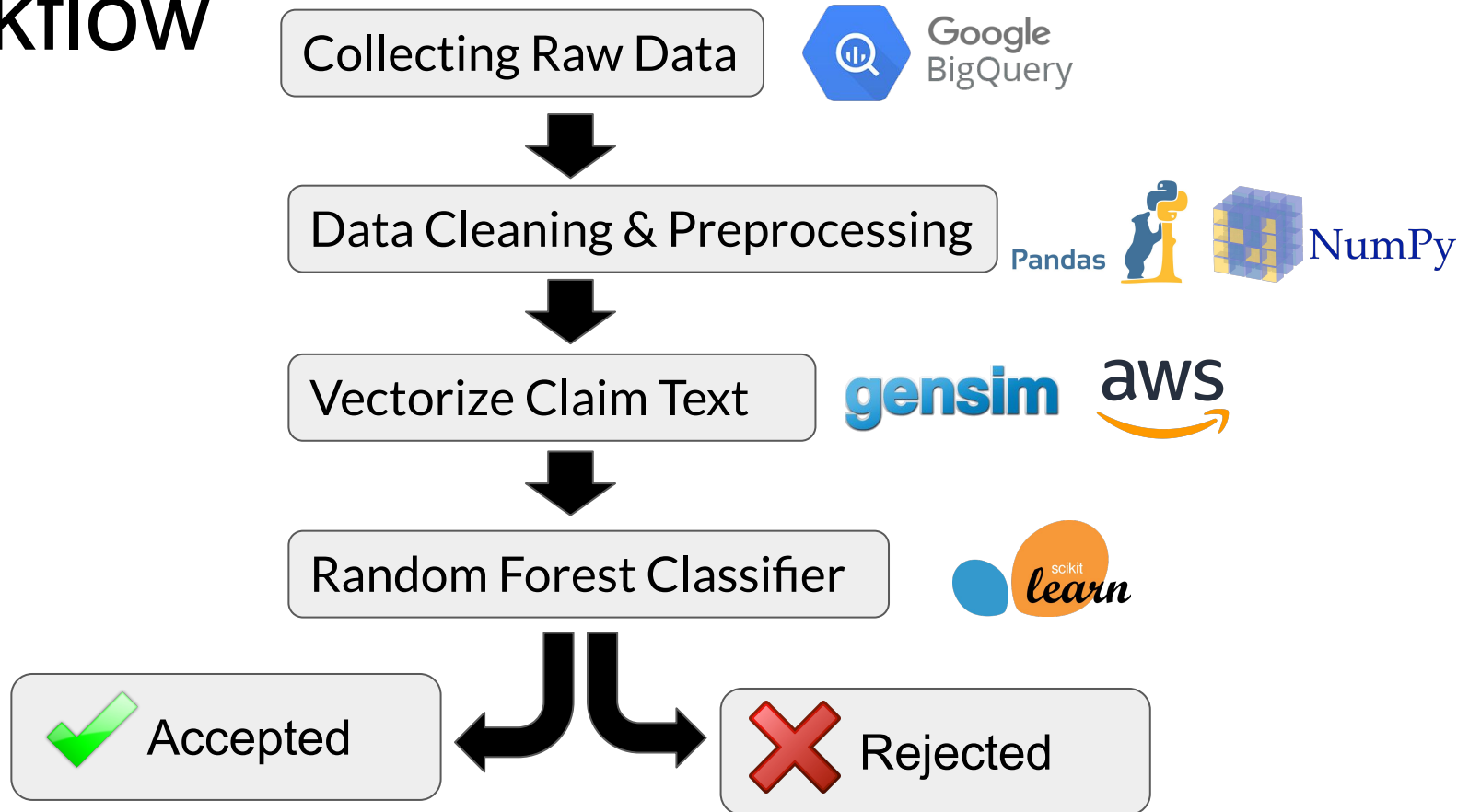
wherein, in the first mode, a first region of system memory is pre-allocated for use in DMA transfers, and

wherein, in the second mode, a second region of system memory for use in a particular DMA transfer is allocated for the particular DMA transfer and deallocated when the particular DMA transfer is complete.

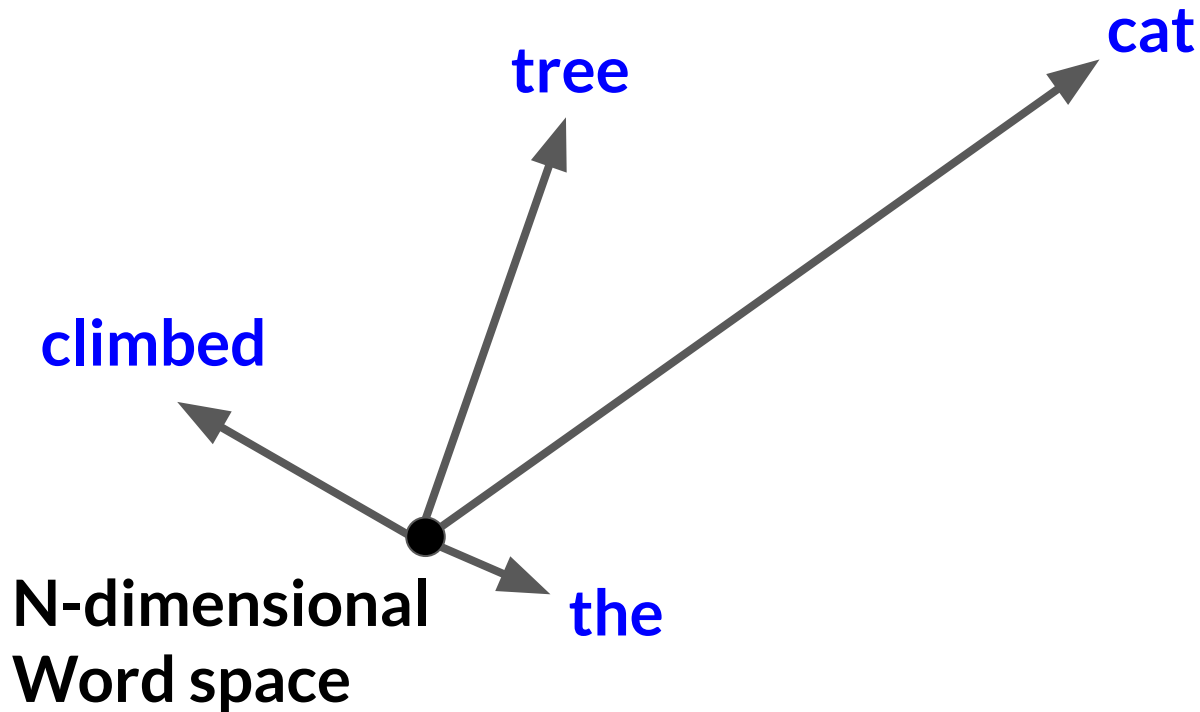
Project Goal

Build a model to automatically predict whether a claim will receive a 101 rejection

Workflow



Word Embeddings

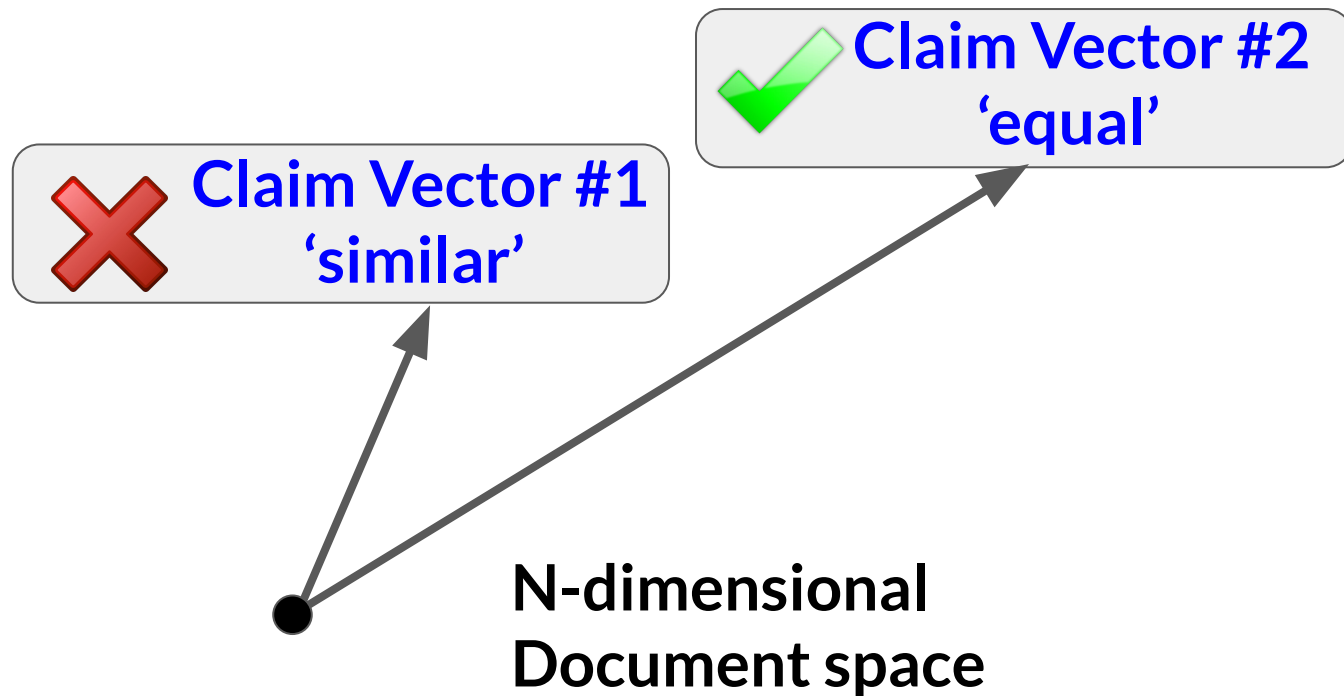


...

the
cat
climbed
a
tree

...

Claim Embeddings

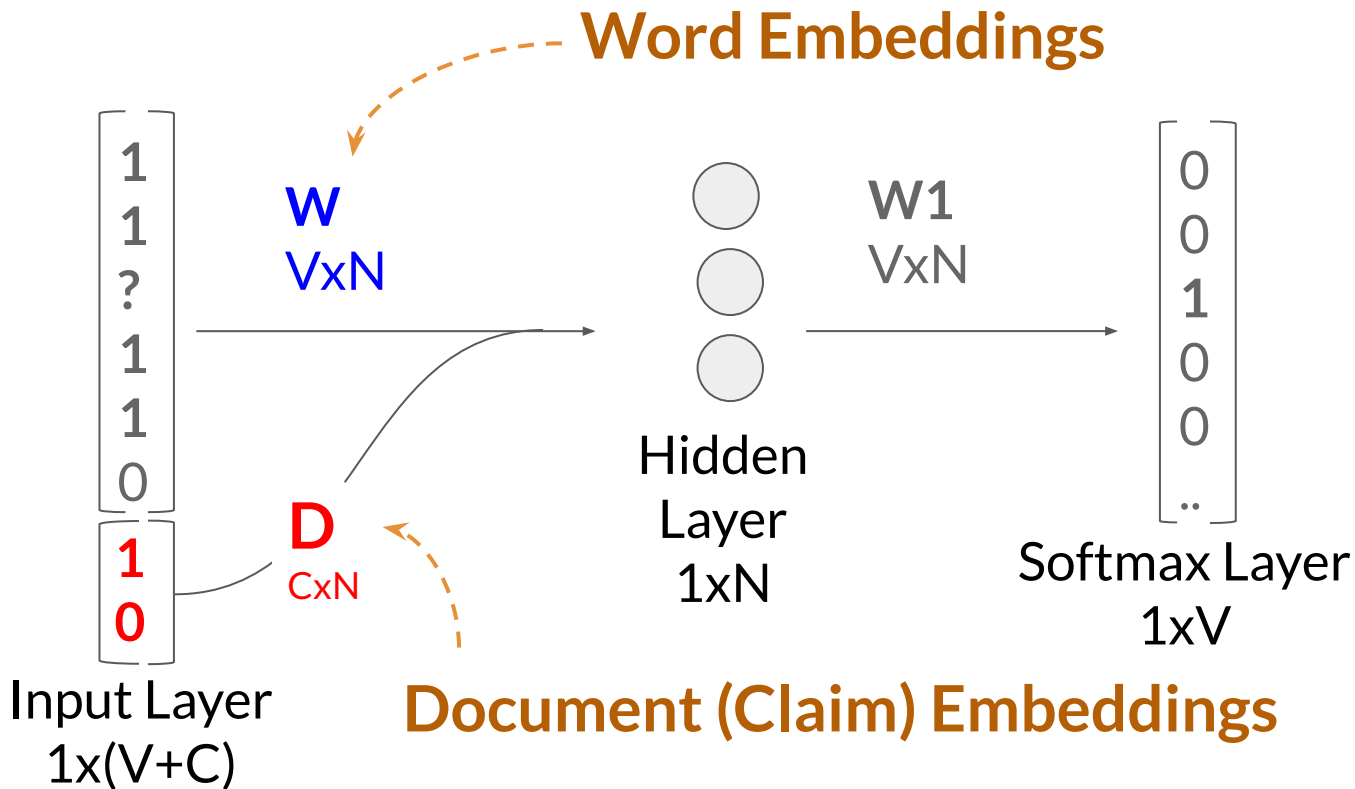


...

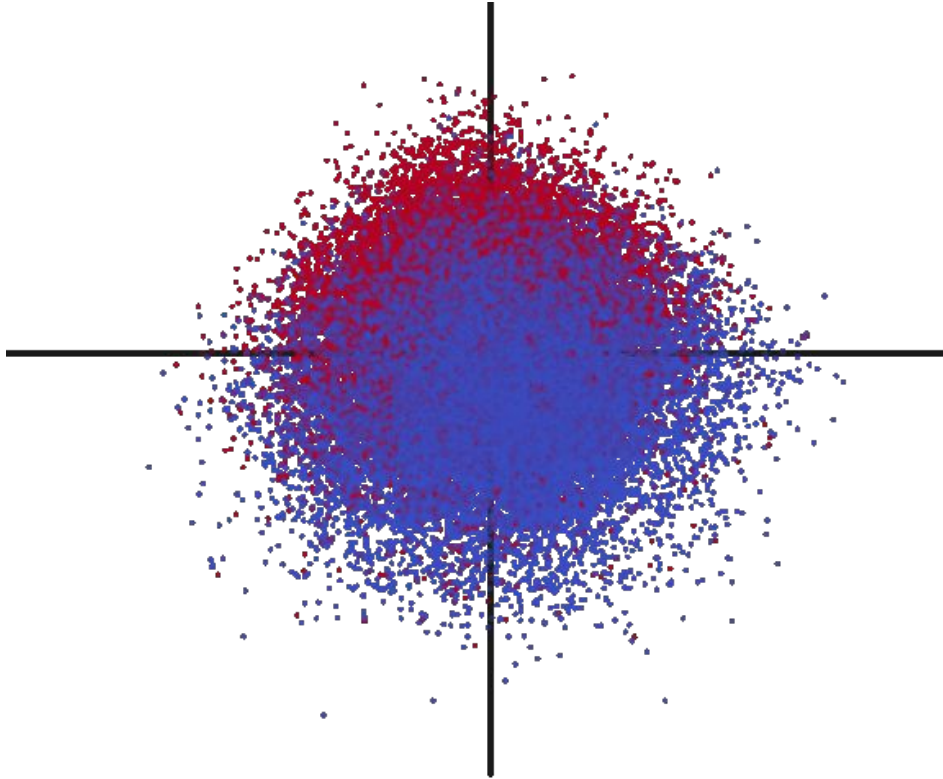
a
second
side
of
???
length
to
the
first
side

...

Claim Embeddings



Principal Component Analysis



Accepted

Rejected

**First 3 Principal
Components**

Results

Test Set: 31,916

Unseen claims
50% Rejected
50% Approved

Accuracy: 96.5%

Rejected Claims
Identified Correctly : 94.5 %

Business Value?

Samsung (in year 2016)

Low Estimate

496	(rejections, only 101)
x \$4000 ¹	(cost of rejection)
x 94.5%	(recall)
= \$1,874,880	(Savings)

Moderate estimate

3096	(total 101 rejections)
x \$4000/2	(estimate cost)
x 94.5%	(recall)
= \$5,851,440	(Savings)

¹ Average cost of responding to rejection: American Intellectual Property Lawyers Association

Thank you.



John Herr



john.r.herr@gmail.com



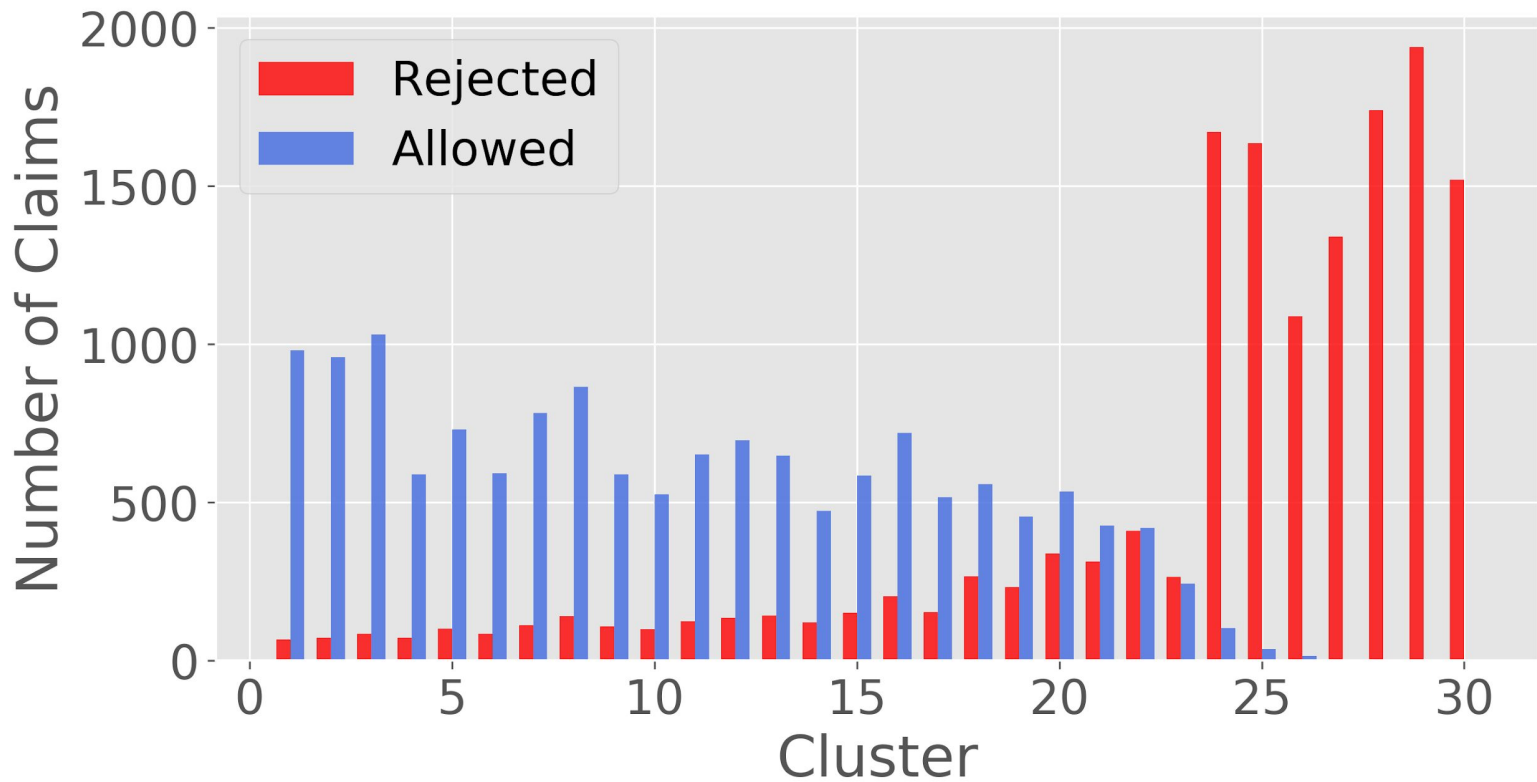
/johnrherr



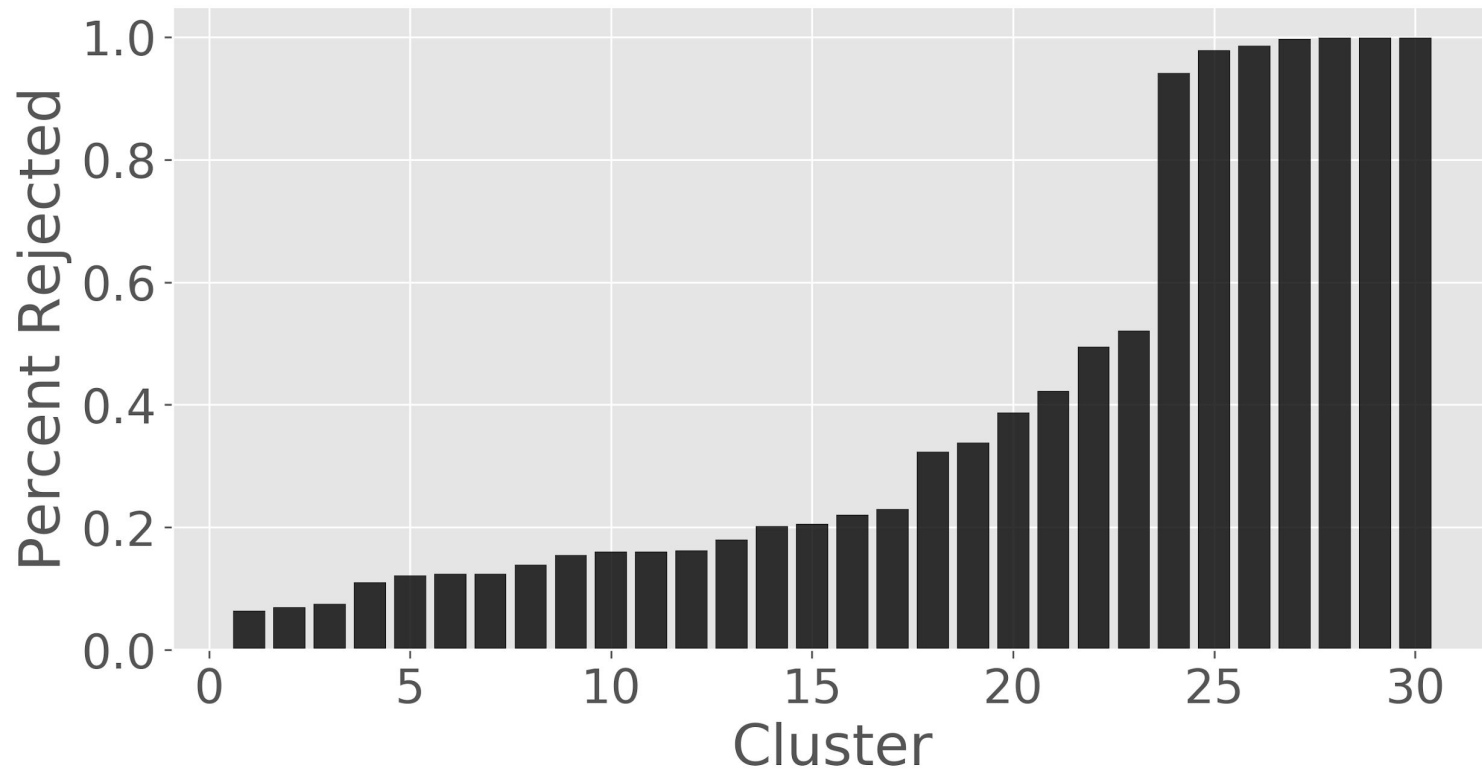
/johnherr

Appendix

K-Means Clustering



K-Means Clustering



Narrowing Words/Phrases

Examples:

'corresponding to...'

'each...'

'responsive to...'

'containing...'

'only...'

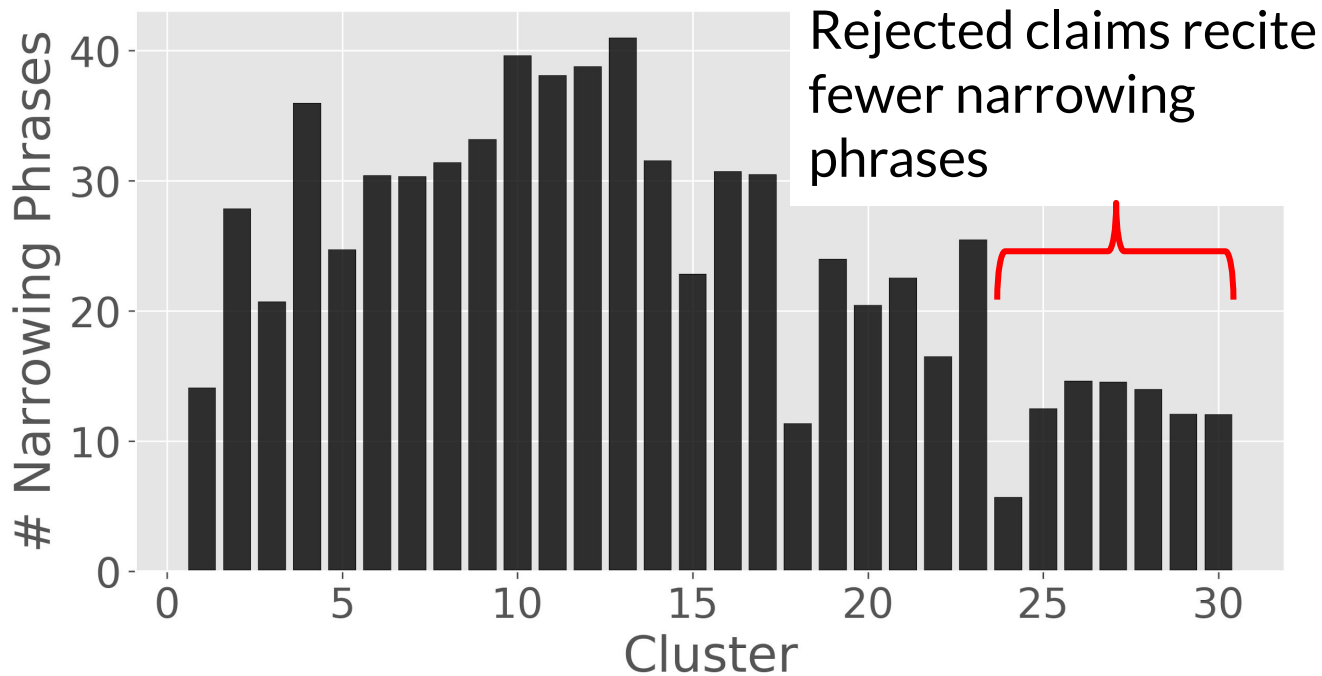
'using the...'

'a first...'

'consisting...'

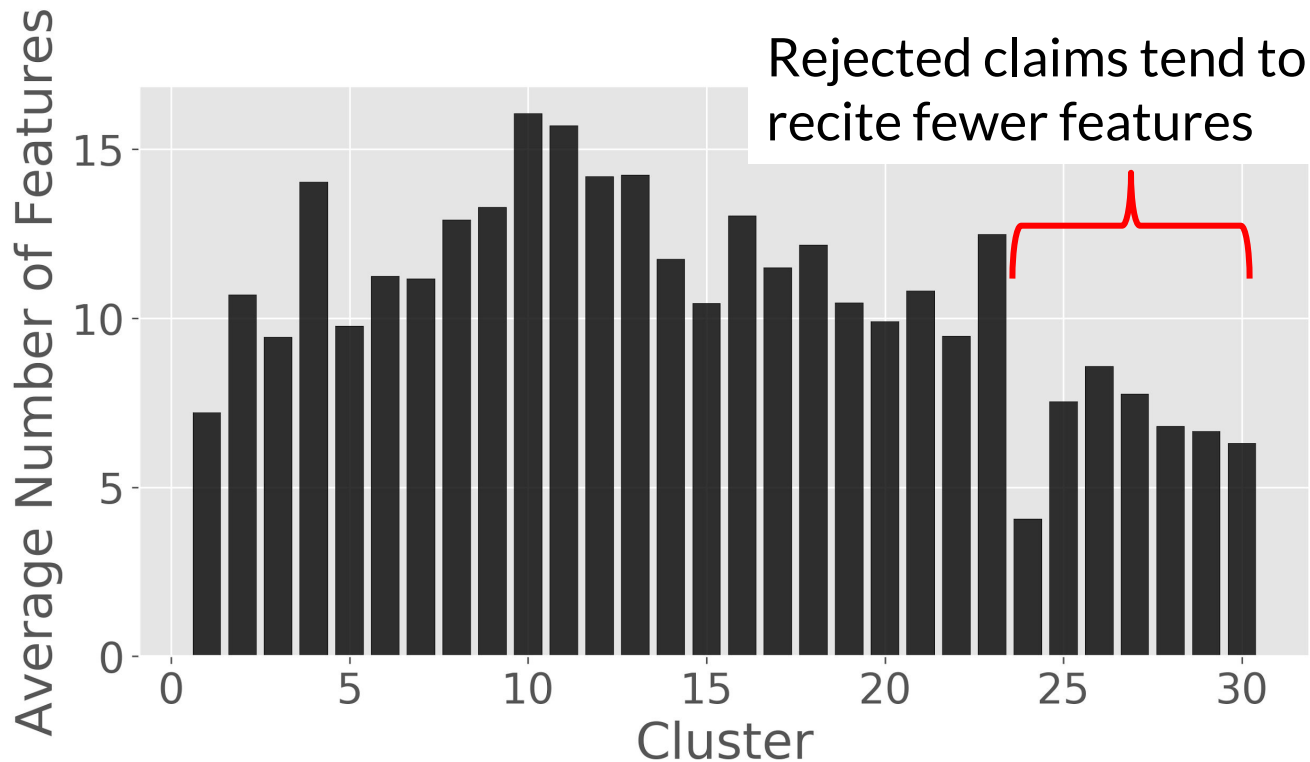
'more than...'

'between...'

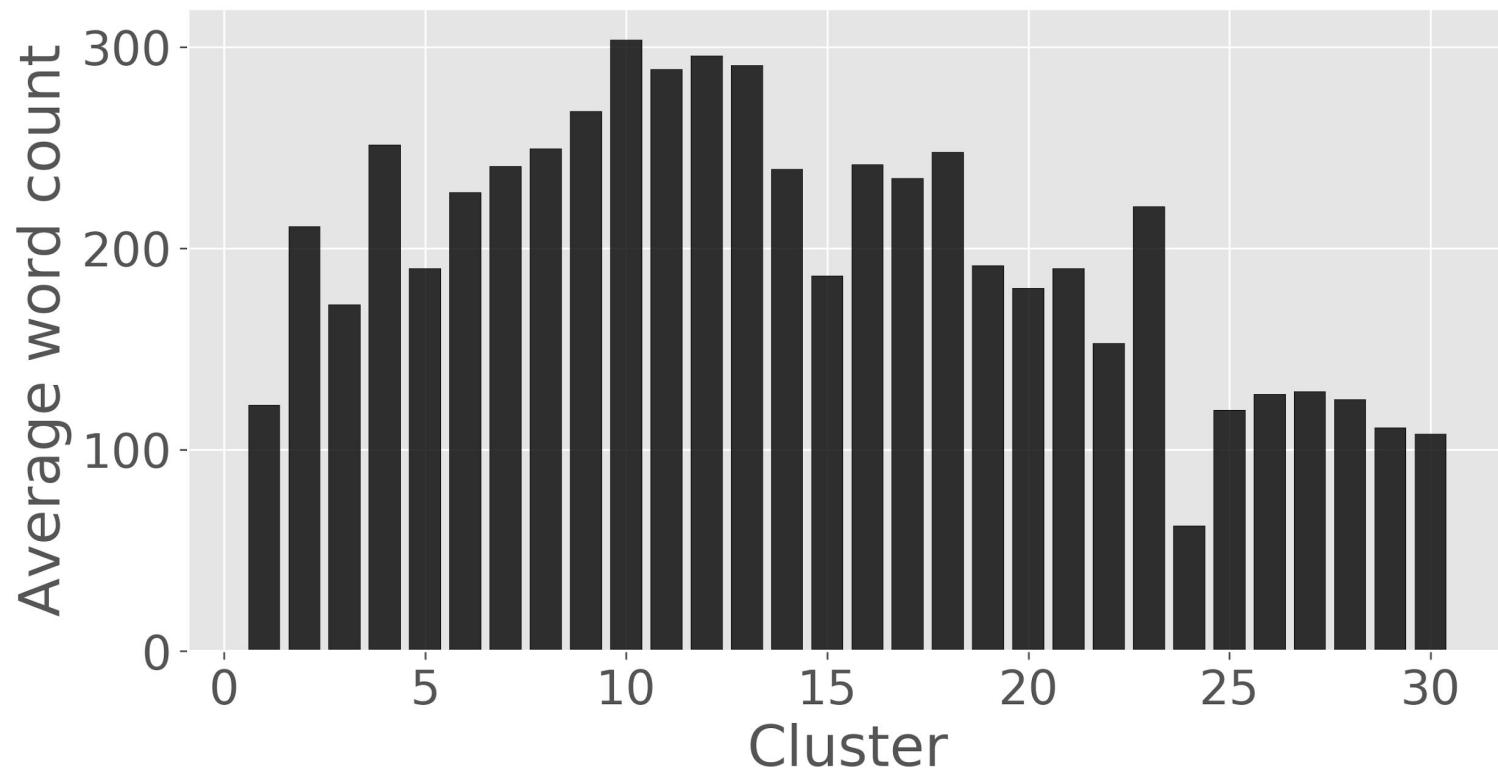


Number of Features

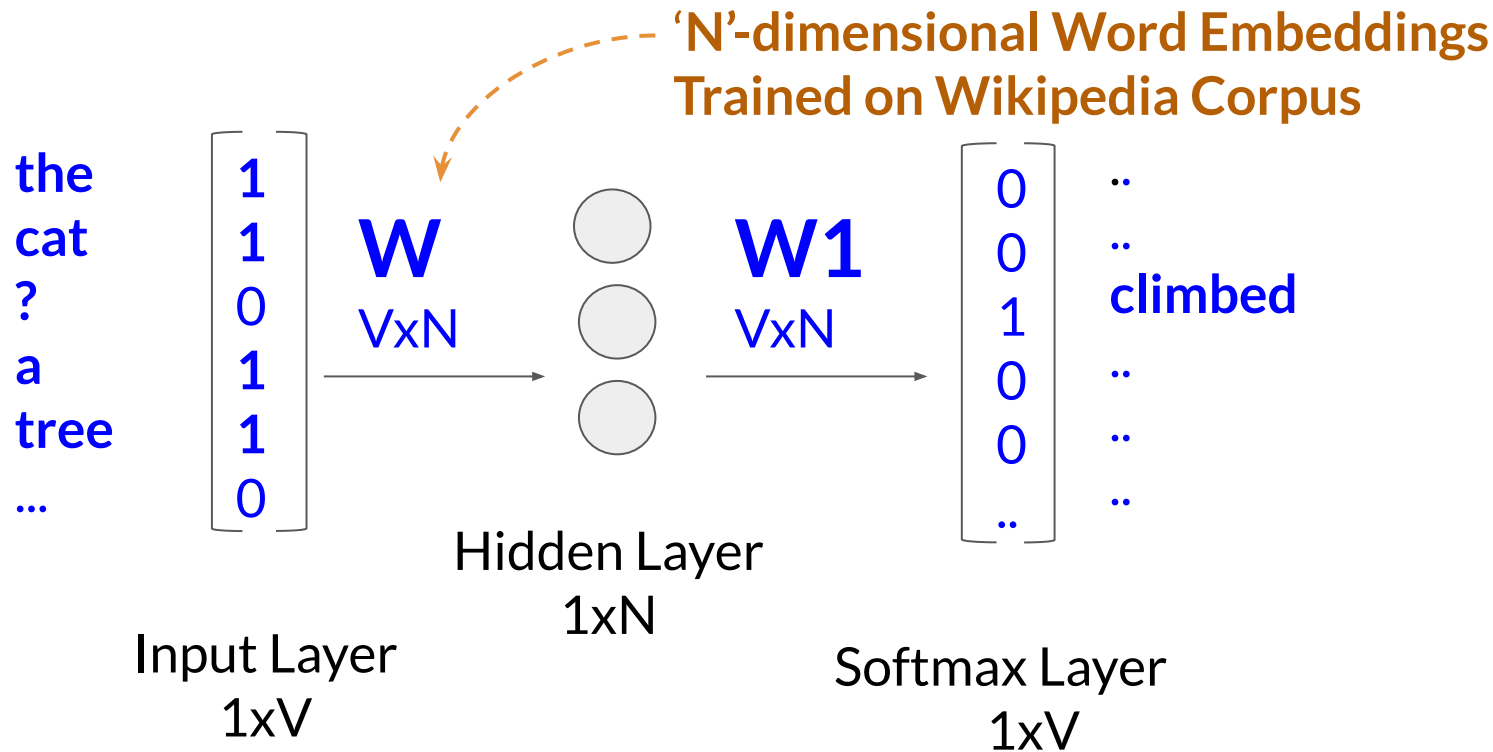
'a first CPU'
'a power supply'
'a display'
'a user interface'
'a communication
module'
...



Claim length (words)



Word Embeddings



Document Embeddings

Document 1 -

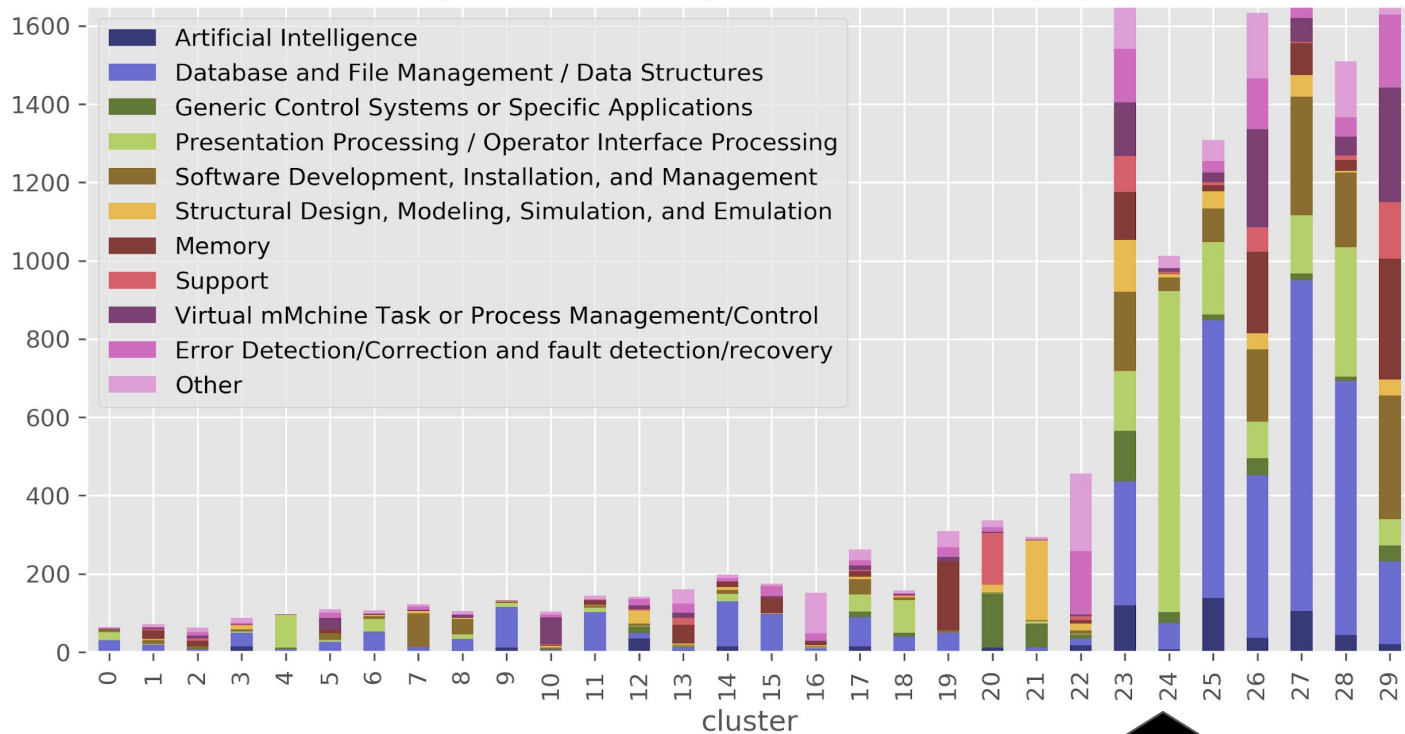
an object having a first side that is (substantially) longer than a second side.

Document 2 -

an object having a first side that is (between 120-150%) longer than a second side.

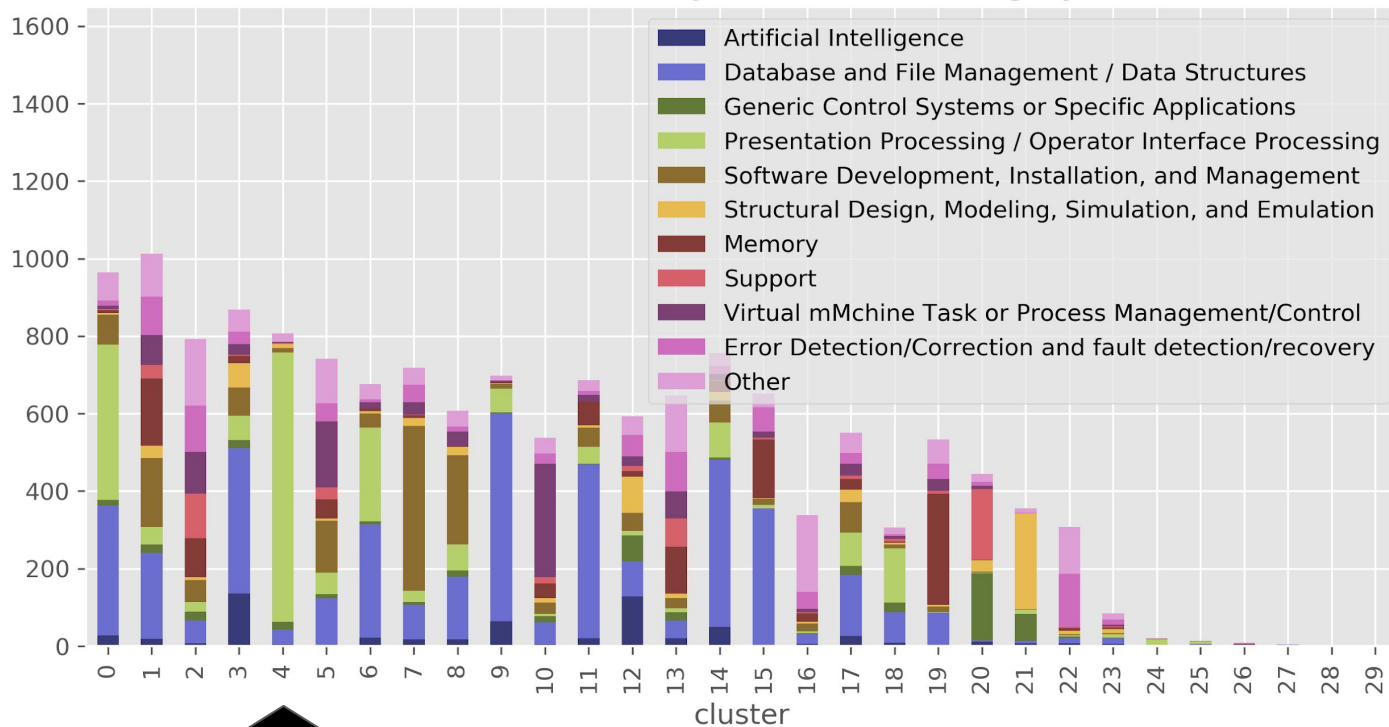
K-Means Clustering

Rejected Claims by Cluster and Category






K-Means Clustering

Allowed Claims by Cluster and Category



Results

True	Approved	Rejected
	Approved	Rejected
Approved	15793 	165 
Rejected	877 	15081 