

# Exploring the Relationship between Complaint Narratives and Categories

**Alexander Einarsson, Sergio Servantez, Marko Sterbentz**

The Enchanted Badgers

November 19, 2020

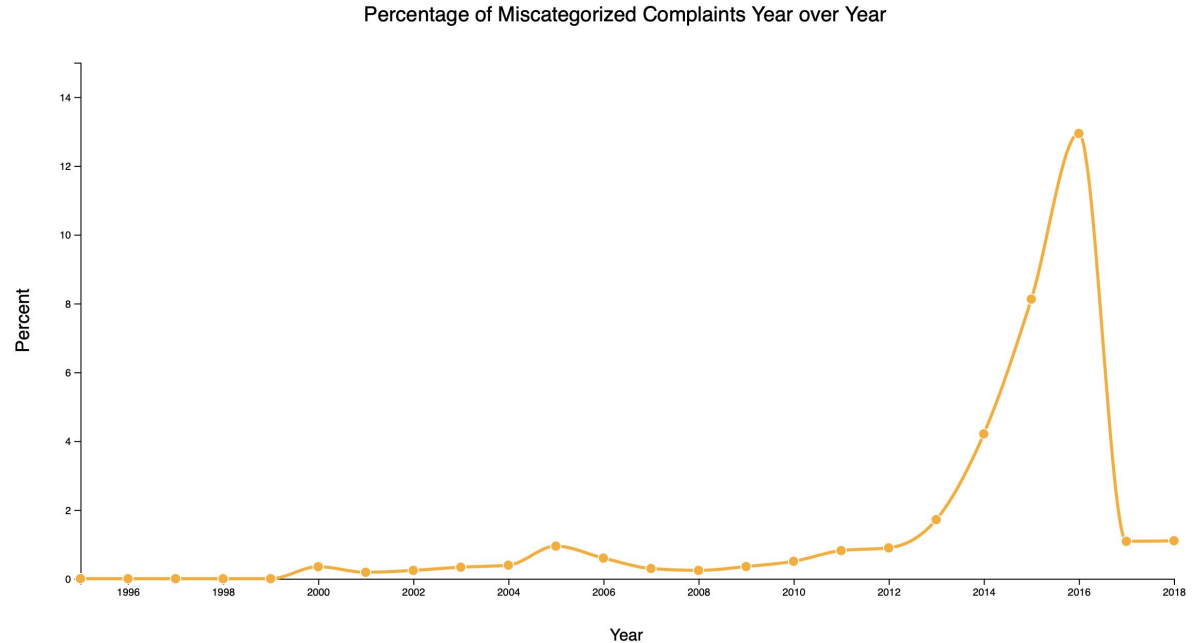
# Theme

Identify and analyze relationships that exist between the complaint category and the complaint report narrative in order to enable:

1. Classification of uncategorized complaints
2. Exploration of other meaningful groupings of complaints

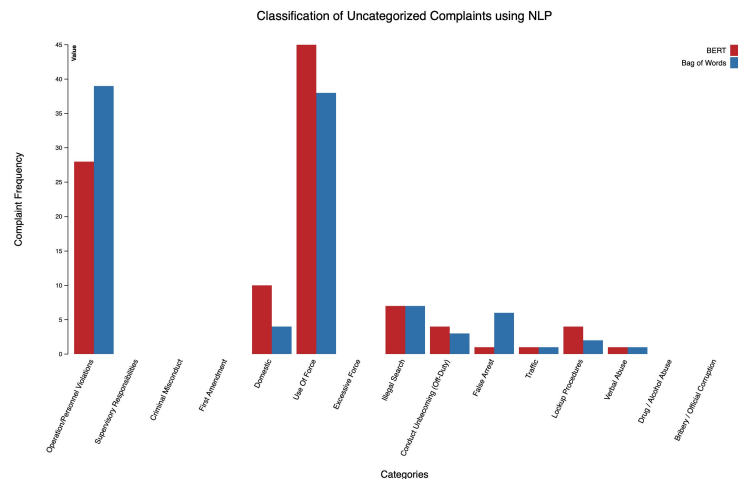
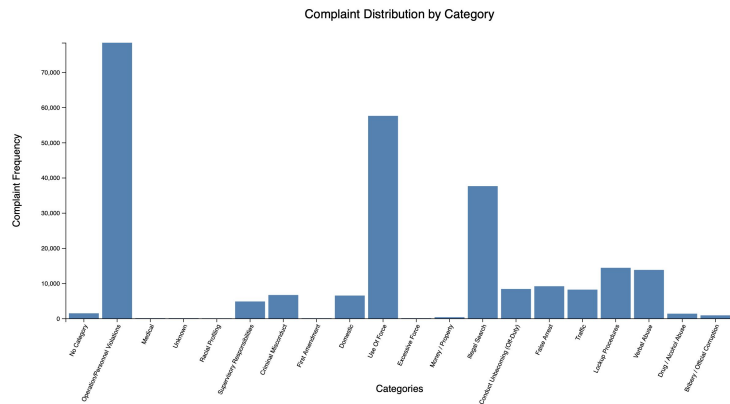
# Motivation

- Explore distribution of miscategorized complaints over time
  - Miscategorized if category = NULL or Unknown
- Rapid increase starting in 2013
- Plummet after 2016



# Predicting the Category of Complaints

- Unclear why complaints were uncategorized
- Fine-tuned state-of-the-art language models on CPDP data to classify complaints
  - 84% accuracy w/ BERT
  - 83% accuracy w/ BoW
  - Trained on ~16000 examples
- Higher proportion of “Use of Force” complaints were uncategorized

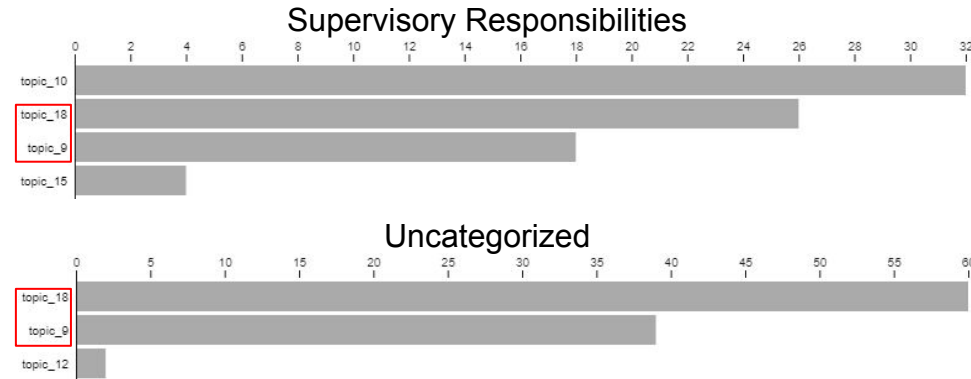


# Predicting the Outcome of an Allegation

- Trained another BERT model to predict the outcome of a model based on complaint summary and category
  - ~70% accuracy (*training set of 2339 examples after class balancing; 60/20/20 train/valid/test split*)
- Summaries and categories are insufficient predictors
  - At least, with this relatively small dataset
- Could we do better if we had more data or better categories?

# Recategorizing the Complaints

- Succeeded in finding interpretable topics
- Uncategorized and Supervisory Responsibilities map to the same LDA-generated topic
- LDA a viable option for new categories



# Example topics

search	warrant	apartment	enter	residence	door	floor	justification	damage	execution	plaintiff	permission
justification	search	stop	vehicle	detain	traffic	handcuff	impound	hour	unidentified	remove	approximately
use	force	improper	weapon	point	unnecessarily	search	profanity	direct	false	excessive	warrant
allegation	recommend	sustain	department	fail	duty	finding	base	complaint	member	incident	register

- Topic 1 - Improper search of residence
- Topic 2 - Improper search/stop of vehicle
- Topic 3 - Improper use of force
- Topic 4 - Departmental complaint (similar finding to Classification)

# Contributions and Opportunities for Further Exploration

1. Demonstrated that complaint narratives have predictive power over categories when using transformers
  - Explore how other complaint data (e.g. geographic location) could be blended into input to improve model accuracy
2. Demonstrated that LDA topic modeling can identify novel and meaningful topics from complaints
  - Compare predictive power of LDA categories on the final outcome to the original categories'
  - Use more advanced topic modeling techniques (e.g. LDA + word embeddings) to examine complaint groupings