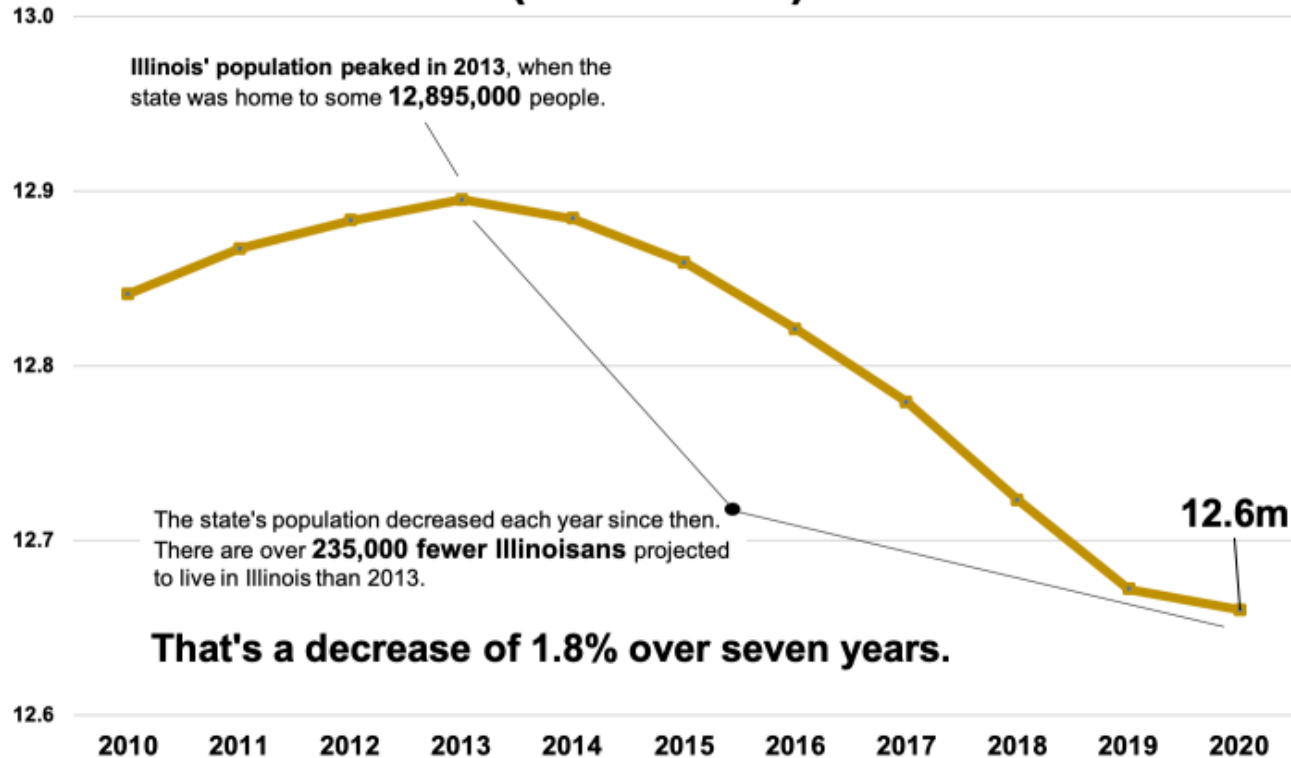


Illinois Population: 2010 - 2020 (in millions)



Source: United States Census Bureau

SHOULD I STAY OR SHOULD I GO?

ANALYSIS OF 2020 NEWS
ARTICLES ABOUT ILLINOIS
AND THE QUESTION OF
POPULATION RECESSION.

Thomas Guardi
University of Chicago
MSCA
NLP Fall 2020

EXECUTIVE SUMMARY

THE NUMBER ONE REASON PEOPLE ARE LEAVING ILLINOIS IS TAXES. PEOPLE FEEL STRONGLY NEGATIVE ABOUT TAXES, THE MEDIA COVERS TAXES EXTENSIVELY, AND POSITIVE SENTIMENT IS HARD TO IDENTIFY. POSITIVE SENTIMENT IS READILY IDENTIFIABLE IN THINGS THAT ARE PAID FOR BY...TAXES.

THERE ARE 2 OPTIONS TO CHANGE THIS DYNAMIC

1. LOWER TAXES OR ELIMINATE TAXES
2. DEVELOP A MARKETING CAMPAIGN BASED AROUND POSITIVE SENTIMENT AND TAXATION IN ORDER TO CHANGE THE POSITIVE/NEGATIVE MESSAGING RATE

IN THIS PRESENTATION WE WILL DISCUSS AND JUSTIFY THE PURSUIT OF OPTION 2



TARGETED
POSITIVE NEGATIVE
SENTIMENT RATE

16.9 %

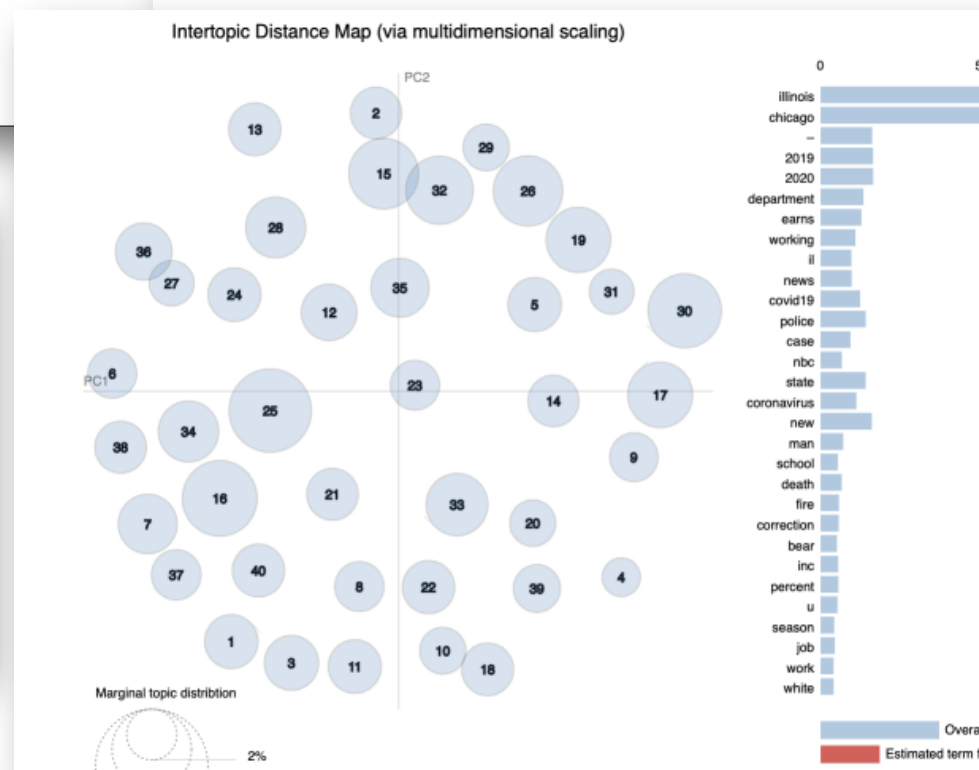
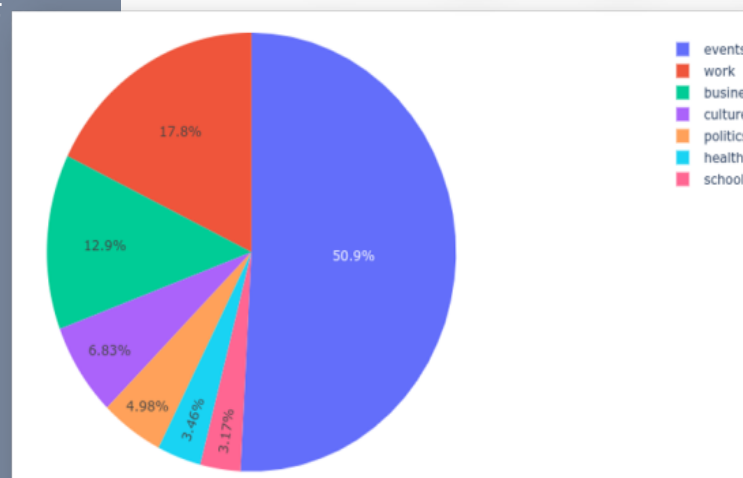
"IF I GO THERE WILL BE TROUBLE. IF I STAY, THERE WILL BE DOUBLE"

- TFIDF (TOP RIGHT)
- LDA ON 40 TOPICS
- MANY LDA ATTEMPTS, THIS IS MOST ILLUSTRATIVE OF THE TOPICS IN THE ARTICLE DROP
- ZERO SHOT EXPLORATORY CATEGORIZATION (PIE CHART)

- COVID (A NEW WRINKLE THIS YEAR)
- POLICE/CRIME
- SPORTS
- POLITICS
- BUSINESS NEWS
- ENTERTAINMENT
- EVENTS
- EPHEMERA (THERE WAS SOME EROTIC LITERATURE INVOLVING ILLINOIS)

Top30 Mentioned People

Entity (Type)	Frequency (Approx.)
(pritzker, PERSON)	3000
(lightfoot, PERSON)	2500
(lori, PERSON)	1700
(smollett, PERSON)	1400
(chicagoarea, PERSON)	1300
(jordan, PERSON)	1200
(michael, PERSON)	1100
(james, PERSON)	1000
(biden, PERSON)	1000
(johnson, PERSON)	900
(blagojevich, PERSON)	850
(robert, PERSON)	800
(george, PERSON)	800
(aaron, PERSON)	750
(kim, PERSON)	700
(floyd, PERSON)	700
(jib, PERSON)	700
(christopher, PERSON)	700
(michael', PERSON)	650
(edwardsville, PERSON)	600
(harry, PERSON)	550
(sorkin, PERSON)	500
(obama, PERSON)	450
(carmichael, PERSON)	450
(kyle, PERSON)	400
(john', PERSON)	400
(austin, PERSON)	350
(anthony, PERSON)	350
(midillinois, PERSON)	300
(trubisky, PERSON)	250



LABELING POPULATION ARTICLES

- POPULATION LOSS AND TAXES
- APPLIED LABELS BASED ON TAXES AND POPULATION USING SNORKEL
- TRANSFORMED DATA AND APPLIED MULTIPLE MACHINE LEARNING TECHNIQUES

- Utilizing transfer learning we were able to train a pre-trained BiGRU classification model on articles relating to Illinois Population loss and achieved a 98% accuracy
- 82% accuracy on Multinomial Bayes
- Important words for classifying our articles (word cloud on the right)
- TFIDF analysis of articles related to population loss were dominated by talk of taxes, followed by politicians and jobs

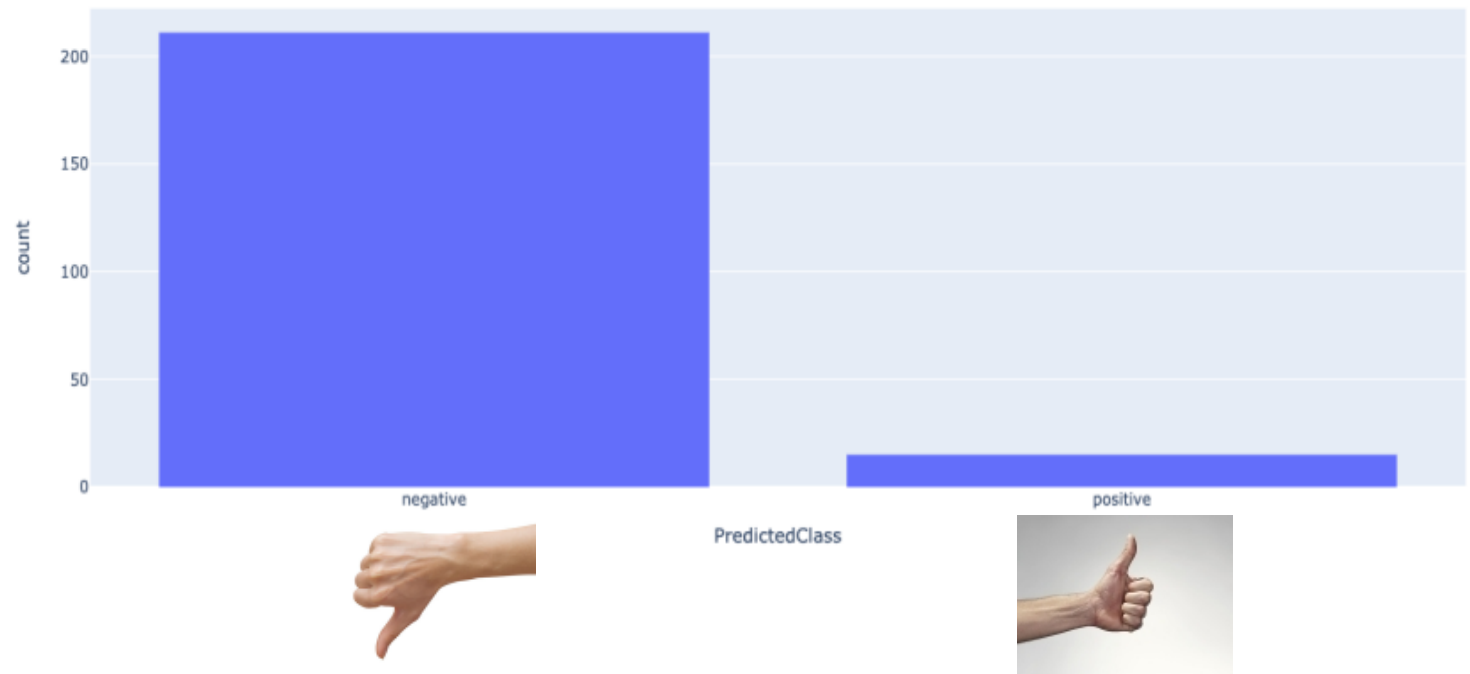


SENTIMENT ANALYSIS

"IF YOU SAY THAT YOU ARE MINE
I'LL BE HERE 'TILL THE END OF TIME"

- Example: Illinois' population declines as U.S. posts slowest growth rate in a century (negative with 98% confidence)
- Ran sentiment on population loss articles. Selected examples with 80% or higher confidence. Highly negative results overall. The positive examples are weakly related to the topics.
- Highly negative sentiment is unsurprising. The structure of the news makes it more likely to air grievances than praise competence. Illinois governance does not have a reputation of competence.
- Main targets of animosity are taxes and politicians so we examine those further.

Confident (80%) Sentiment on Target Articles



- Relevant articles selected and sentiment labeling applied using multiple methods. Vader and TextBlob sentiment were tested with good results, but a zero shot classification model that utilized a version of BERT outperformed other methods.
- Model: valhalla/distilbart-mnli-12-3
- Detail: A version of The Multi-Genre Natural Language Inference (MultiNLI) corpus is a crowd-sourced collection of 433k sentence pairs annotated with textual entailment information.



SENTIMENT ANALYSIS ON THE TARGETED ENTITIES

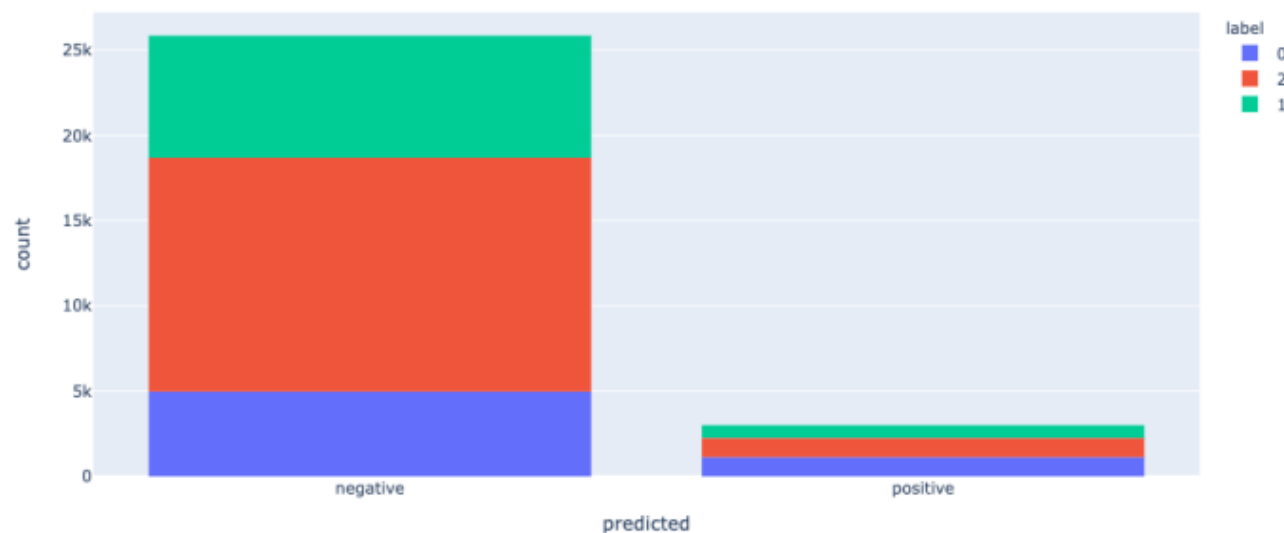
"ONE DAY IT'S FINE AND NEXT IT'S
BLACK,
SO IF YOU WANT ME OFF YOUR
BACK"

- PRE-TRAINED BERT MODEL
TRAINED ON POPULATION CENTRIC
SENTIMENT DATA

- MODEL IS USED TO LABEL ARTICLES
AS POSITIVE AND NEGATIVE
SENTIMENT



- The Targets
 - Pritzker, Lightfoot, Tax (all kinds of taxes: property, marijuana, soda, gas, etc.) came up at every level of my analysis. Deserving or not these 3 stand in for an entire host of issues, grievances, and occasionally praise.
 - Pritzker and Lightfoot are stand ins for people's complaints about the state and the city and governance in general. I was surprised Madigan wasn't more noticeable in my analysis.
- The Results
 - Our targeted entities are quite unpopular
 - Mayor Lightfoot taking the brunt of negative coverage, perhaps due to Chicago having a large media influence

Sentiment: Tax=Blue, JB=Green, Lori=Red



RESULTS

- Use our trained sentiment model on the targets -> Apply relevance classification Relevance using our trained biGru classifier that we trained earlier
- Process:
 - For relevant population predicted articles and for each of our 3 target entities:
 - Established article keywords using RAKE for both positive and negative labeled sentiment. Applied TFIDF vectorization to establish the most valuable words relating to the target and sentiment.
 - Analyzed 2-4 n-grams and manually sifted for repeat and relevant topics from top 100 weights

	Positive Rate	Negative Sentiment 	Positive Sentiment 
TAX	20%	Property, Fair, Income, Beneficiaries	Marijuana tax fund, lottery, sales, parks
LOR I	9%	Teacher contract, Taxes (property, income, etc.), marijuana, pension, sport betting, covid response	Tax amnesty, Block club (social programs), JB (be more like him)
JB	25%	Cannabis, lots of chicago related complaints, Lightfoot, census response	Housing relief, women's march, affordable housing, aarp, landlord lender solidarity, covid response

CONCLUSIONS AND RECOMMENDATIONS

- MEDIA IS DOMINATED BY COMPLAINTS. EVEN POSITIVE SENTIMENT AROUND A MARIJUANA TAX THAT FUNDS THE STATE, MEETS A NEGATIVE SENTIMENT BARRAGE. THE COMPLAINT MESSAGING IS HEAVILY TILTED TOWARD THE NEGATIVE (THIS IS AFTER REMOVING A LOT OF COVID AND POLICE RELATED NEWS). PROGRAMS THAT HELP PEOPLE ARE VIEWED POSITIVELY BUT PAYING FOR THEM IS NOT.

- RECOMMENDATION: LARGE SCALE POSITIVE MESSAGING CAMPAIGN FOCUSED ON PROGRAMS PEOPLE LIKE. MAYBE CHANGING THE SENTIMENT AND MESSAGING AROUND TAXES WILL MAKE PEOPLE STAY IN ILLINOIS BECAUSE THE ONLY 2 CERTAIN THINGS IN LIFE ARE DEATH AND TAXES.