Dynamic Topic Modeling of NYT

By Liam Moran

Objective

Objective: Create a dynamic topic model that shows how news stories change over time.

Could be used by:

1. New York Times
2. Competing Newspapers
3. Independent journalists

Tools Used

Natural Language
Tool Kit (NLTK)
Basic Text Analytics

scikit
learn

pandas

GENSIM
topic modelling for humans

March vs April Word Clouds

March Word Cloud

April Word Cloud

10 Topics through Data

1.Politics

Top Words:Political,Election,Biden

2. Women

Top Words: Mother, woman,she,her

3. Pronouns

Top Words: Me, Your, You, My

4.Sports

Top Words: Player, season, tournament

5. War in Ukraine

Top Words:Military, Ukraine, War

6. Abortion Laws

Top Words: Abortion, Law, Bill, Court

10. Pandemic

Top Words: Covid, Vaccine, Health

7. Kentanji Brown Jackson

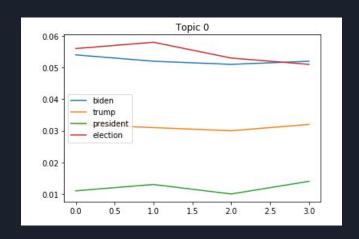Top Words: Jackson, Judge, Justice

8. Jobs
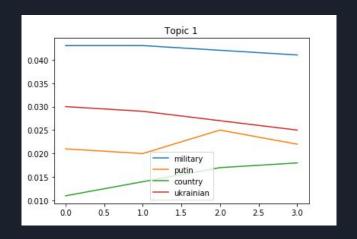
Top Words: Company, Work, People

9. Gas and Inflation

Top Words: Inflation, Oil, Price

Dynamic Topic Model

Topics do not change very much over four weeks

Future Work

1. Expand time period
2. Try other models
3. Build webapp