



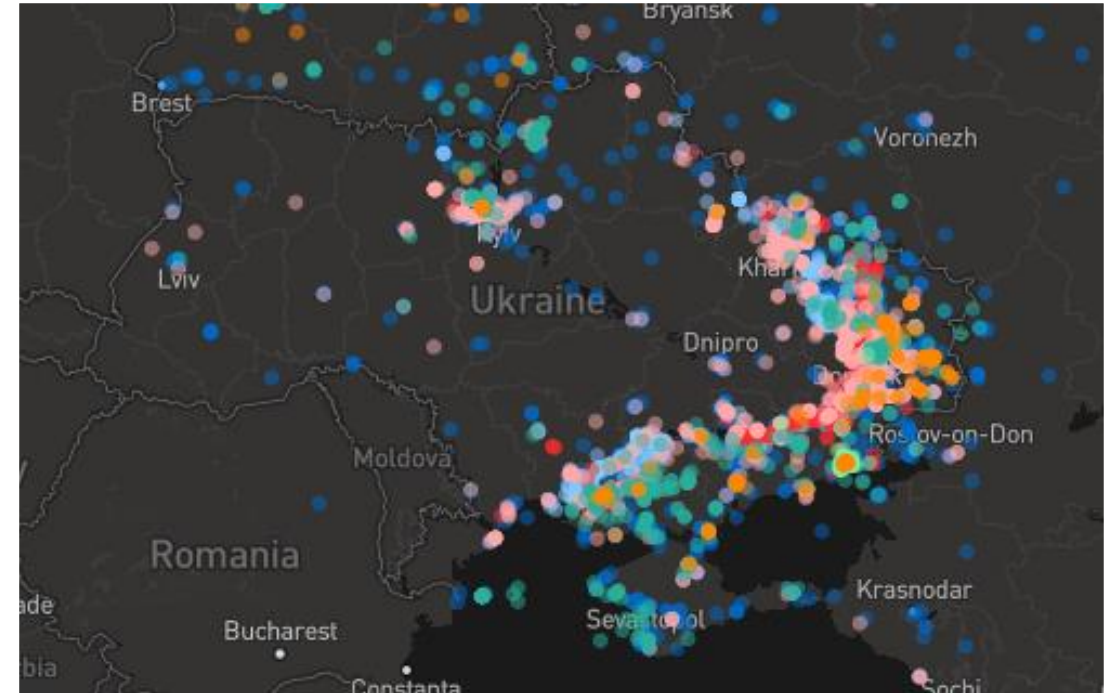
# Covering the Ukraine - Russia Conflict

Data driven, social-media oriented, and keeping you safe

---

Adam Miner

Data scientist for the Associated Press



# Russia-Ukraine Conflict Problem at hand

---

- Immense amount of social media data:
  - 60,000 useful Tweets and Telegrams since the start of conflict
  - Exponentially more raw social media out there
- Dangerous situations for journalists in Ukraine and Russia
  - Recently jailed Wall Street Journal reporter in Moscow<sup>1</sup>

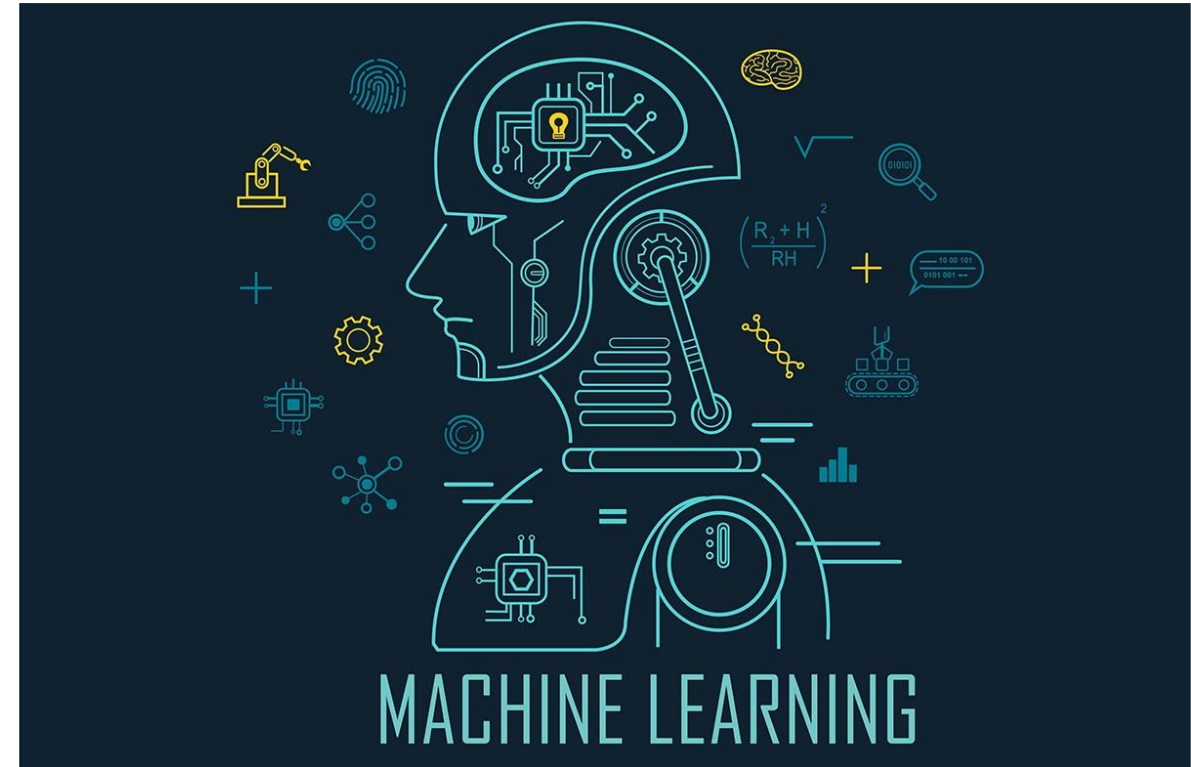


<sup>1</sup> <https://www.cnn.com/2023/03/30/business/wsj-journalist-arrested-russia/index.html>  
<https://cdn.britannica.com/18/194818-050-E7A7A993/view-Kiev-Ukraine.jpg>  
[https://lp-cms-production.imgix.net/2023-03/500pxRF\\_61786861.jpg](https://lp-cms-production.imgix.net/2023-03/500pxRF_61786861.jpg)

# A Helpful Solution

---

- Built a Machine Learning model that:
  - Takes vetted, located social media posts
  - Clusters them into seven unique clusters
  - Displays via an interactive web app
- Repeatable and customizable
- Help your coverage:
  - Saves time by filtering relevant social media posts for you
  - Narratives for each cluster have already been generated and analyzed
  - Worldwide focused, regularly updating and robust dataset



# How the Model Works (generally)

---

1. Gathers the data from five non-profit geolocation groups (Bellingcat, Ceninfores, Texty.ua, Defmon, Geoconfirmed)
2. Analyzes the text of the social media posts for relevance and importance of words within the entire dataset
3. Creates seven unique clusters to assign the social media posts to
4. Classifies new posts into one of these seven clusters

# Clusters

---

- Russian Movements and Activities
- Global Russian and Ukrainian Activities
- The Siege of Mariupol
- The Destruction Cluster
- Ukrainian Positions and Activities
- Battle for Bakhmut
- Satellite Imagery

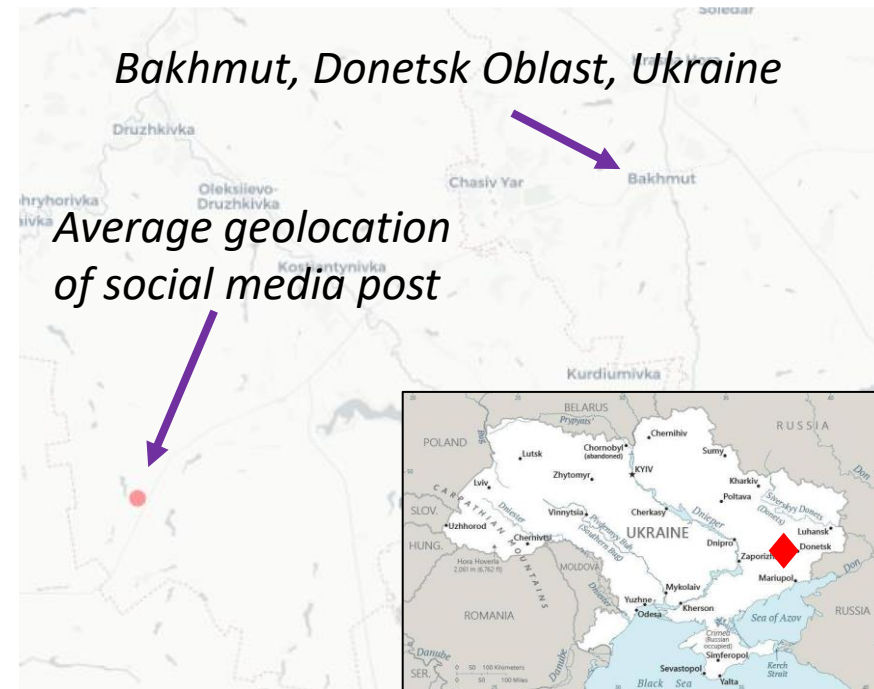


**Learn more  
about the  
clusters at this  
site here!**

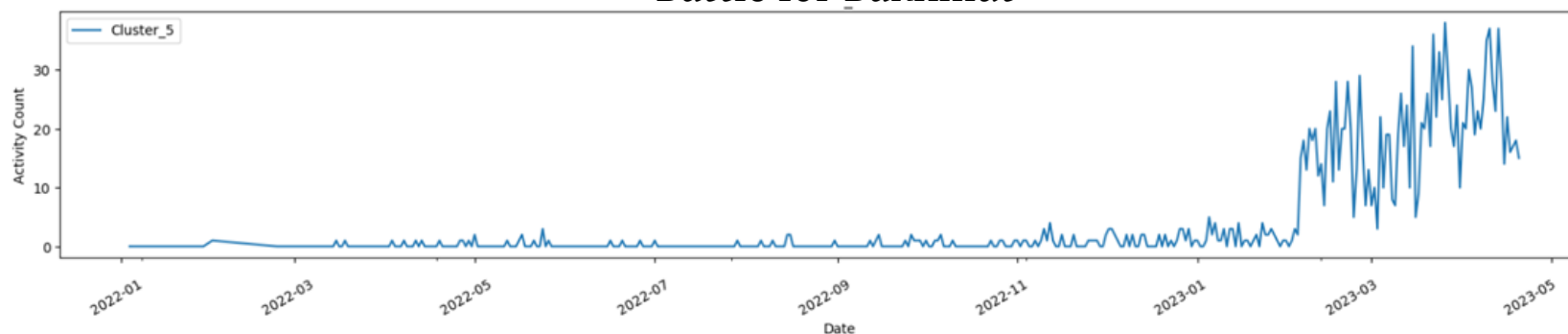


# An Example: Battle for Bakhmut

- Not much coverage from social media, until Russia focused on taking the city in January 2023
- Social media activity may be indicative of a spike in conflict
- Tipping you off for stories, with an already extracted, vetted social media dataset to use
- Potential time saver while still allowing for good journalism coverage



Battle for Bakhmut



# Application Demo

---

- Live site: <https://extracting-and-clustering-posts-webapp.onrender.com/>
- Prototype Prediction model for local demonstration

# AP

## Future Work for AP

With a partnership opportunity:

- Collect more social media data from sources
- Fine tune model using your requested inputs
- Provide method to upload file of tweet texts, for batch classification
- Generate readouts of social media activity for each cluster on a recurring basis





Thank You

---

