

---

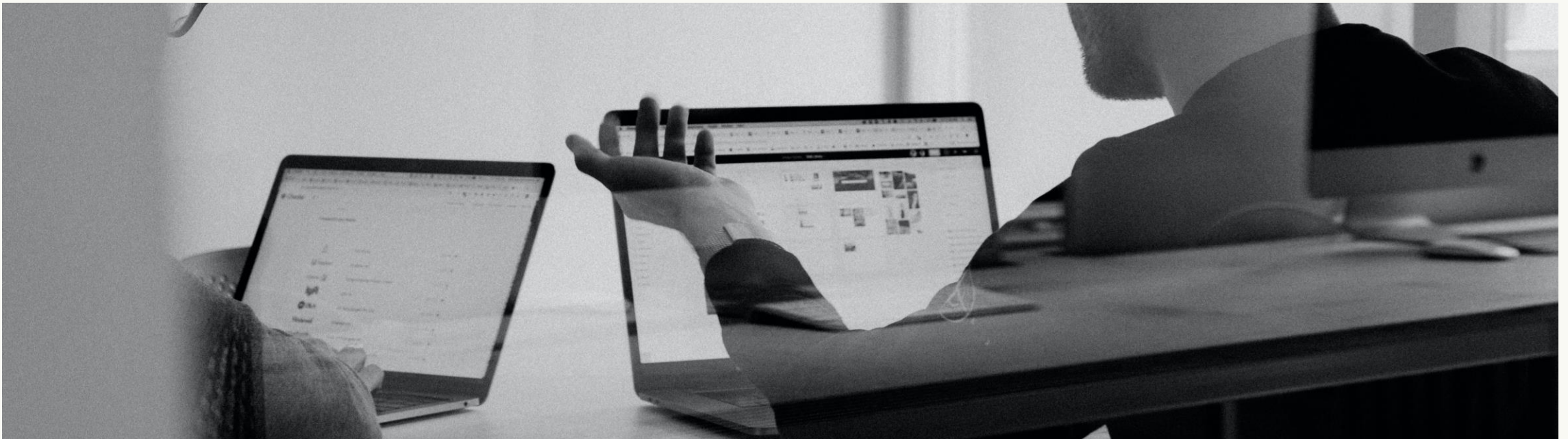
Sept 05, 2023

Adhitan Rani Kasinathan

---

# Privacy-first marketing

## Rethinking Contextual Ads through Data Science



---

# Contents

---

**01**

## Problem Statement

- Current digital advertising landscape  
- what's at stake?

**02**

## Proposal

- How Might We use Machine Learning  
to address privacy concerns in digital  
marketing?

**03**

## Impact

- What is the impact in terms of dollars ?

**04**

## Dataset

- What's our data - how does it look like?

**05**

## Pre-processing

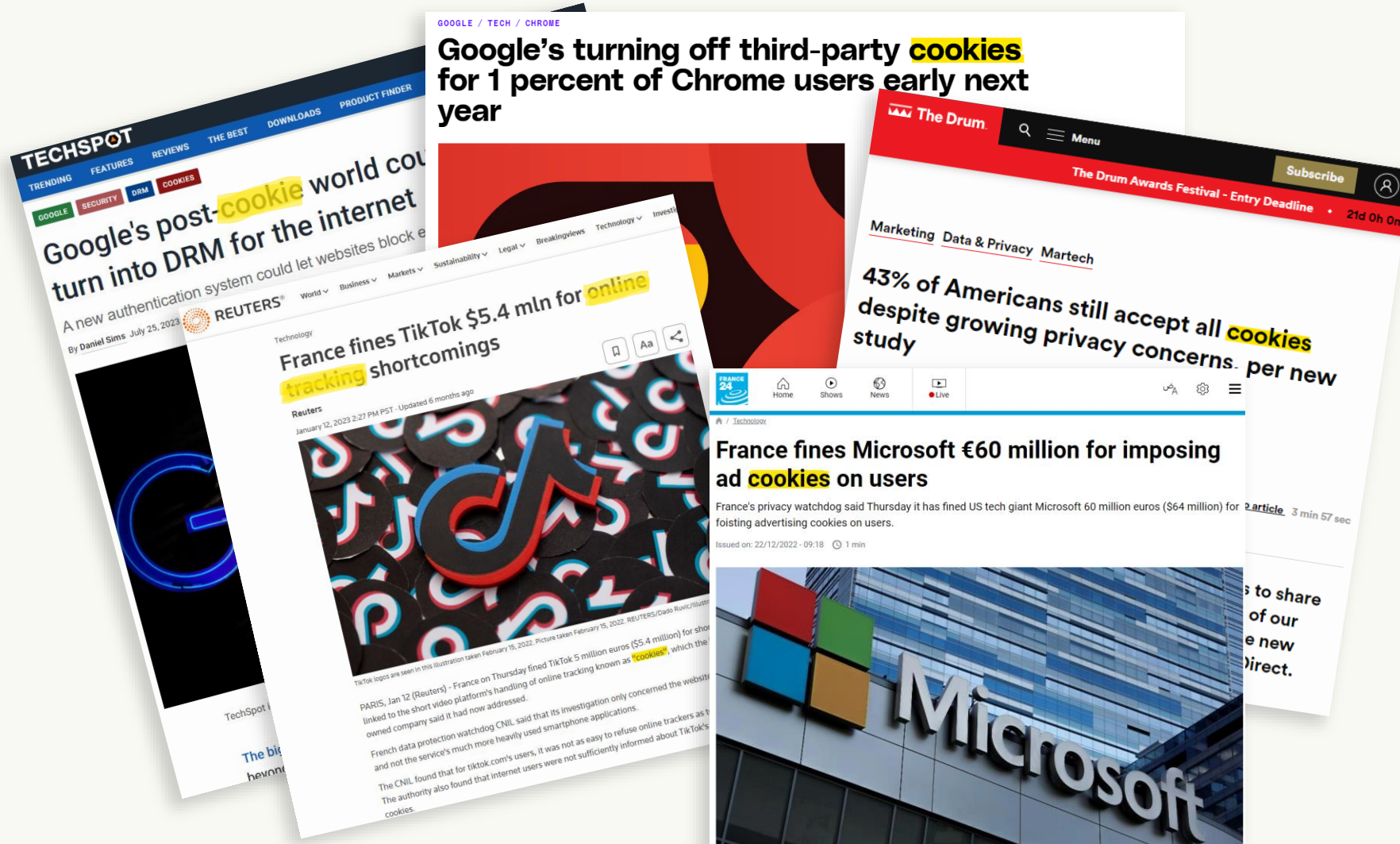
- Reduce class imbalance
- Tokenization and Vectorization

**06**

## Modeling & Evaluation

- Comparing model performances
-

# Current digital advertising landscape - what's at stake?



---

# Current digital advertising landscape - what's at stake?

---



## User Tracking:

**Cookies** track and store user website activity, including page visits, session duration, cart items, and logins.



## Ad Personalization

This tracked user behavior across websites is used to personalize ads based on interests and improve campaign effectiveness



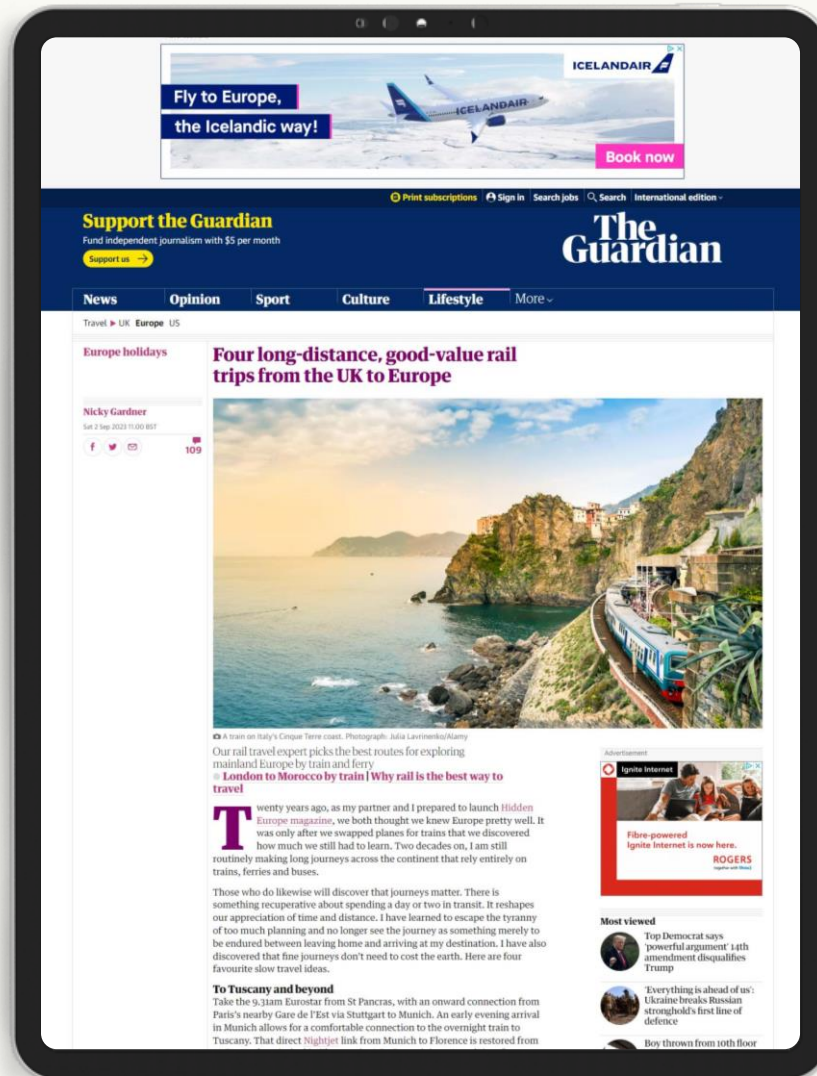
## Privacy Concerns:

However, usage of cookies has raised privacy concerns, prompting browser changes and regulations for user data protection.

---



# How Might We use Machine Learning to address privacy concerns



Travel Ad

Picks advertisers  
(IcelandAir) who wants  
to show travel ad

Ad platform receives  
the category signal  
Example: Travel

Analyze texts

ML Predicts the  
category it belongs

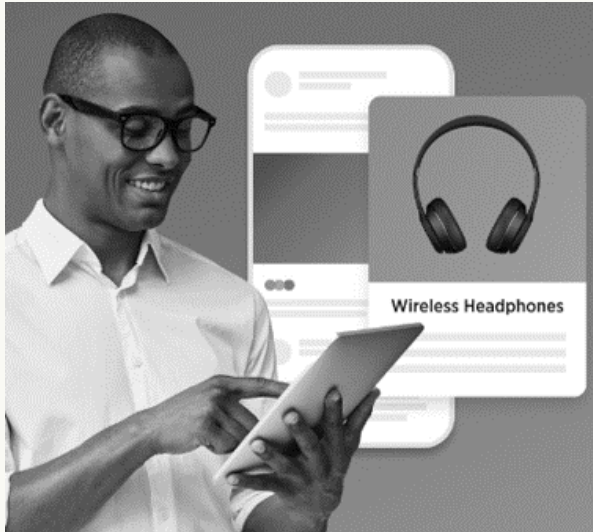
---

## What is the impact in terms of dollars ?

---

“

“..the Interactive Advertising Bureau estimates publishers' exposure at up to **\$10 billion in lost revenue** without third party cookies..” Source: [IAB](#)



Users seeing relevant ads



Advertisers will gain high ROI



Publishers sell more ads

---

# What's our data - how does it look like?



Contains news articles  
from 2012 to 2022.

Rows  
**209K**

Columns  
**6**

link	headline	category	short_description	authors	date
https://www.huffpost.com/entry/covid-boos	Over 4 Million Americans Roll Up	U.S. NEWS	Health experts said it is too early to predict whether demand would match up with the 171 million doses of the new boosters the U.S. ordered for the fall.	Carla K. Johnson, AP	2022-09-23 00:00:00
https://www.huffpost.com/entry/american-a	American Airlines Flyer Charged,	U.S. NEWS	He was subdued by passengers and crew when he fled to the back of the aircraft after the confrontation, according to the U.S. attorney's office in Los Angeles.	Mary Papenfuss	2022-09-23 00:00:00
https://www.huffpost.com/entry/funniest-tv	23 Of The Funniest Tweets About	COMEDY	"Until you have a dog you don't understand what could be eaten."	Elyse Wanshel	2022-09-23 00:00:00
https://www.huffpost.com/entry/funniest-p	The Funniest Tweets From Parent	PARENTING	"Accidentally put grown-up toothpaste on my toddler's toothbrush and he screamed like I was cleaning his teeth with a Carolina Reaper dipped in Tabasco sauce."	Caroline Bologna	2022-09-23 00:00:00
https://www.huffpost.com/entry/amy-coope	Woman Who Called Cops On Black	U.S. NEWS	Amy Cooper accused investment firm Franklin Templeton of unfairly firing her and branding her a racist after video of the Central Park encounter went viral.	Nina Golgowski	2022-09-22 00:00:00
https://www.huffpost.com/entry/belk-work	Cleaner Was Dead In Belk Bathroo	U.S. NEWS	The 63-year-old woman was seen working at the South Carolina store on Thursday. She was found dead Monday after her family reported her missing, authorities said.		2022-09-22 00:00:00
https://www.huffpost.com/entry/reporter-g	Reporter Gets Adorable Surprise I	U.S. NEWS	"Who's that behind you?" an anchor for New York's PIX11 asked journalist Michelle Ross as she finished up an interview.	Elyse Wanshel	2022-09-22 00:00:00
https://www.huffpost.com/entry/puerto-ric	Puerto Ricans Desperate For Water	WORLD NEWS	More than half a million people remained without water service three days after the storm lashed the U.S. territory.	DÁNICA COTO, AP	2022-09-22 00:00:00
https://www.huffpost.com/entry/mija-docu	How A New Documentary Capture	CULTURE & ARTS	In "Mija," director Isabel Castro combined music documentaries with the style of "Euphoria" and "Clueless" to tell a more nuanced immigration story.	Marina Fang	2022-09-22 00:00:00
https://www.huffpost.com/entry/biden-un-r	Biden At UN To Call Russian War A	WORLD NEWS	White House officials say the crux of the president's visit to the U.N. this year will be a full-throated condemnation of Russia and its brutal war.	Aamer Madhani, AP	2022-09-21 00:00:00

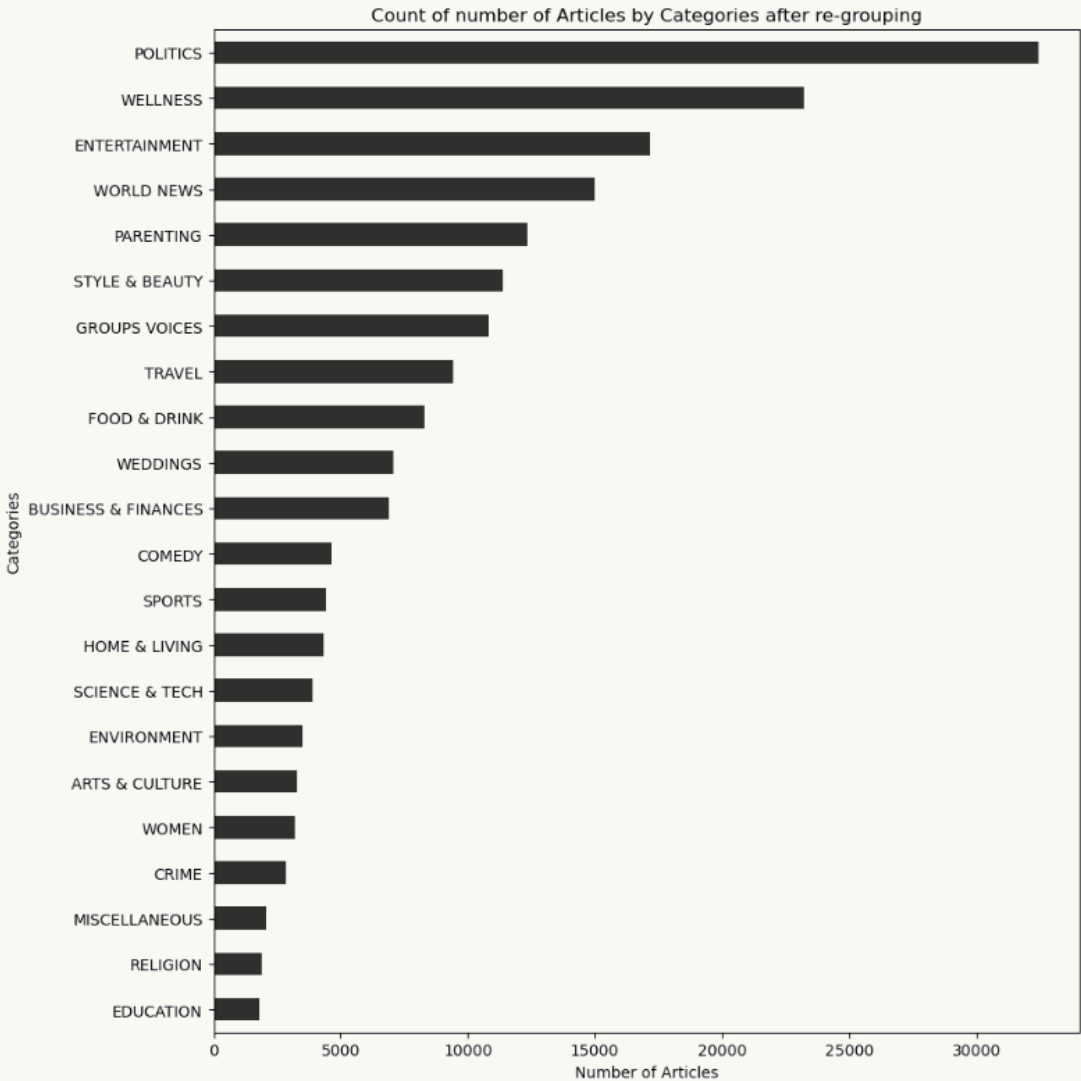
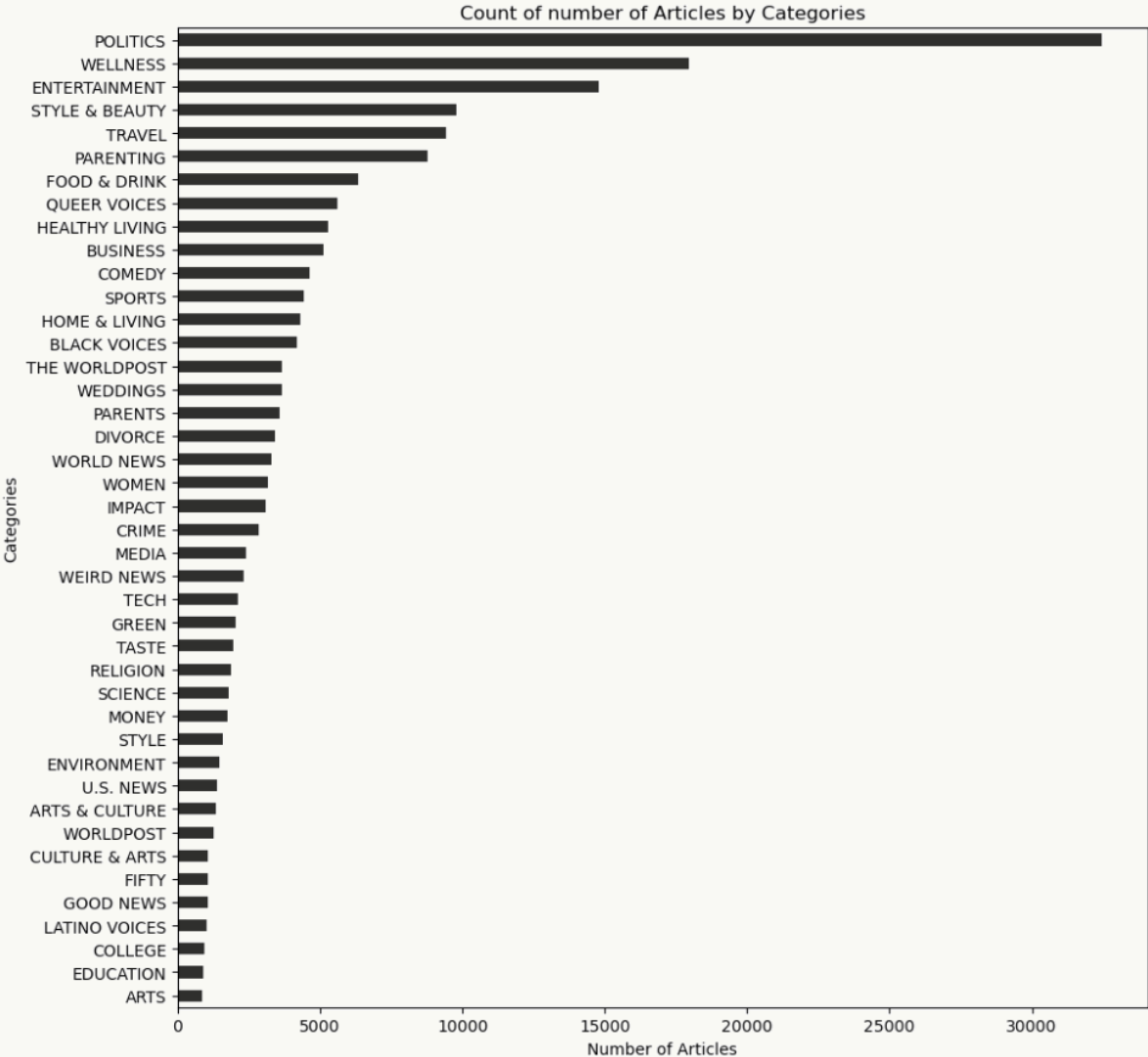
Feature

Target

Feature

# Pre-processing

Effort to reduce class imbalance: From 42 → 22 unique categories





---

# Pre-processing

---

## Custom-Tokenizer:

“

The quickest brown foxes, jumped over the lazier dog!', 'The laziest fox. Leaped Over The quicker dogs?

”

### Stemming

brown | dog | fox | jump | lazier | laziest | leap | quicker | quickest

### Lemmatization

brown | dog | fox | jump | lazy | leap | quick

---

# Modeling & Evaluation

Model	Tokenizer	Pre-processing / Variation	Train	Test	Remarks
Logistic Regression	CountVectorizer	Lemmatization	90%	66%	Overfitting
		Stemming	90%	67%	Overfitting
	TF-IDF	Lemmatization	75%	68%	High Test accuracy
		Stemming	75%	68%	High Test accuracy
		Class_weight='balanced' / Lemmatization	74%	65%	Promising!
		SMOTE / Lemmatization	91%	67%	Overfitting
		Sentence2Vec	64%	63%	Promising!
Naïve Bayes	TF-IDF	MultinomialNB() / Lemmatization	65%	61%	Promising!
		ComplementNB() / Lemmatization	69%	66%	Promising!
Support Vector Machines	TF-IDF	LinearSVC / Lemmatization	89%	68%	Overfitting
		SVC(kernel='linear') / Lemmatization	82%	69%	Overfitting / Computation heavy

# Product Mockup

×

User Input

Enter the headline:

The 'Wizard of Wharton' says stocks are on solid ground - and house prices are shaking off mortgage pain

Enter the article text:

The likelihood that the Fed will raise in September is now almost nil, and in fact it puts the November increase in doubt," he said, referring to the central bank's next two meetings.

Predict

Deploy

⋮

# Contextual Chronicle

By Adhitan Rani Kasinathan

Politics

Travel

Sports

Business & Finance

Food & Drink

Style & Beauty

Home & Living

Wellness

Parenting

## The 'Wizard of Wharton' says stocks are on solid ground - and house prices are shaking off mortgage pain


Sep 04, 2023, 10:18 PM

Jeremy Siegel says the stock market is on firm footing, and the housing market is shaking off the surge in mortgage rates for now.

"Equities can hold in here," the retired finance professor known as the "Wizard of Wharton" said on the "Behind the Markets" podcast on Friday. The benchmark S&P 500 index has gained nearly 18% this year, while the tech-heavy Nasdaq Composite has surged by 34%.

Siegel believes stocks are in good shape because the inflation threat is receding, so the Federal Reserve won't have to hike interest rates as aggressively as many feared.

"The likelihood that the Fed will raise in September is now almost nil, and in fact it puts the November increase in doubt," he said, referring to the central bank's next two meetings.



Business & Finances

---

# Next Steps

**01.**

Improve Class Imbalance

Address class imbalance by augmenting additional data

**02.**

Embedding Methods

Explore other word embeddings using techniques like Glove, SpaCy.

**03.**

Advanced Models

Implement advanced NLP models like BERT and others

---

---

**Thanks !**

---