

# Building a Content Recommender with News Data

by Sim Borodach

# medium.com

Health article recommendations:  
anecdotal v. expert opinions

Medium | Health



## How a Vegan Diet Really Works

My science-based journey to a plant-based lifestyle



Paul Greenberg in Elemental  
Apr 26 · 15 min read ★

### LATEST

#### America's Self-Imposed Doctor Drought

A shortage of doctors nationwide is leading to a healthcare crisis of our own making.

Thomas Whitworth in Inquiry of the Public Sort

Apr 26 · 8 min read



# Data

Microsoft News Data



## Article Data

- Code
- Topic
- Title
- Content

## User Data

- User ID
- Article codes



powered by Microsoft News



Outlook.com



Wayfair  
ad



Microsoft Store



Deals & Offers



Face

CORONAVIRUS NEWS

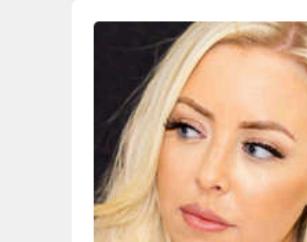
NORTH ADAMS / 49°F

NEWS

ENTERTAINMENT

SPORTS

ESPORTS



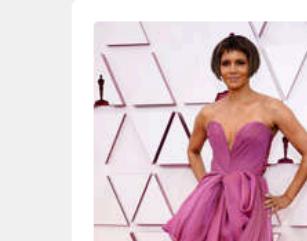
Mortgage Rates Hit 1.98%  
APR FIXED

Ad LendGo



Gun crazy: For too many  
Americans, guns are tied to  
masculinity, patriotism and...

S Salon



2021 Academy Awards  
fashion hits and misses

Ad Wonderwall



Historic Refi Rate Alert: Now  
at 2.06% APR

Ad LendingTree



Woman Victi  
While Work

N Newsweek

## TRENDING NOW ▾

Liz Cheney says Trum

Zendaya Wore A Ne

Florida woman on pi

Swimming Pool Catc

For Biden's next 100

## RECOMMENDED S

Hong Kong Singapo

Nazanin Zaghari-Rat

Discovery Cove Disc

Ridge Apartments

Camp Rock

# Tech Stack



Requests  
http for humans



BeautifulSoup



NumPy

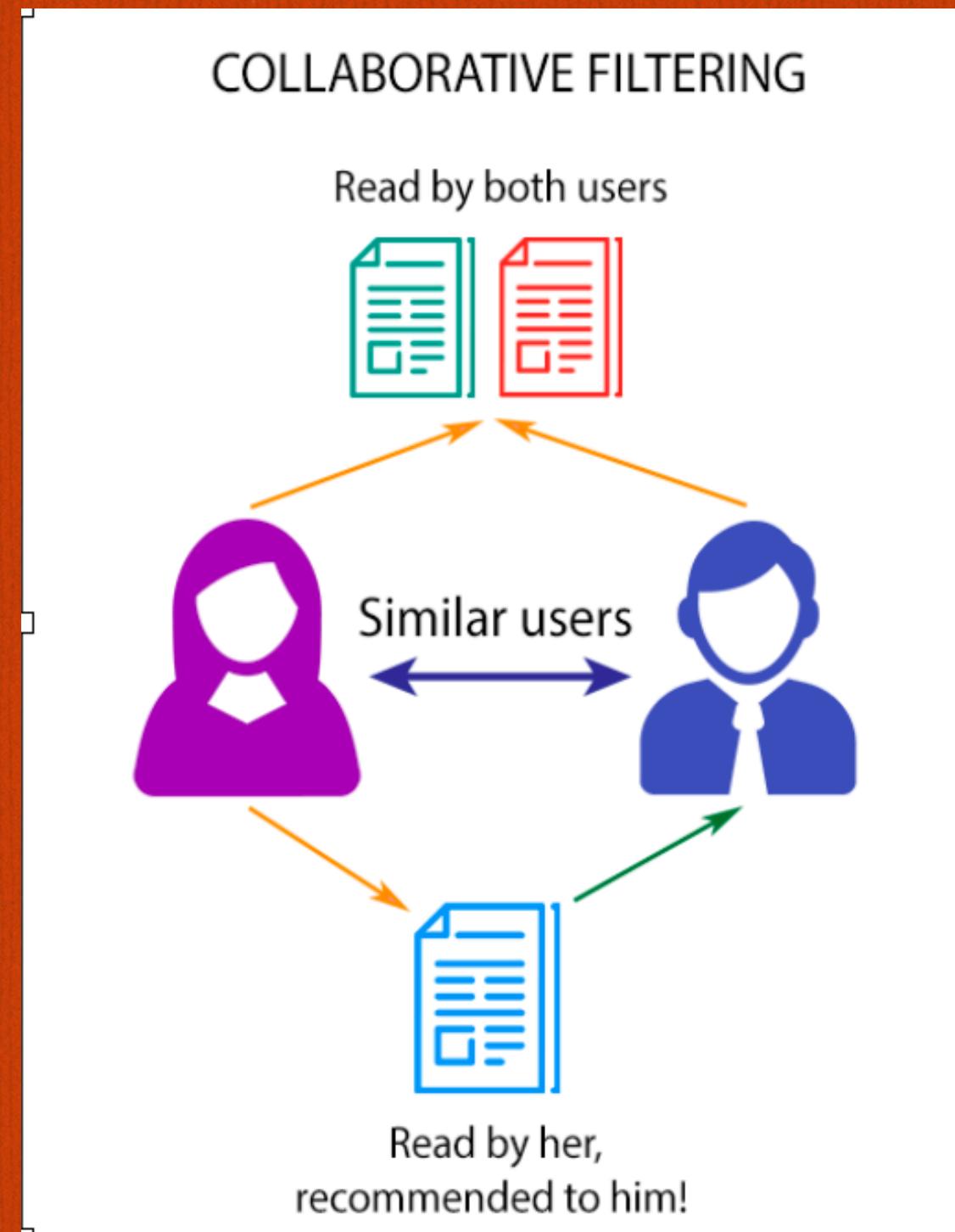


pandas

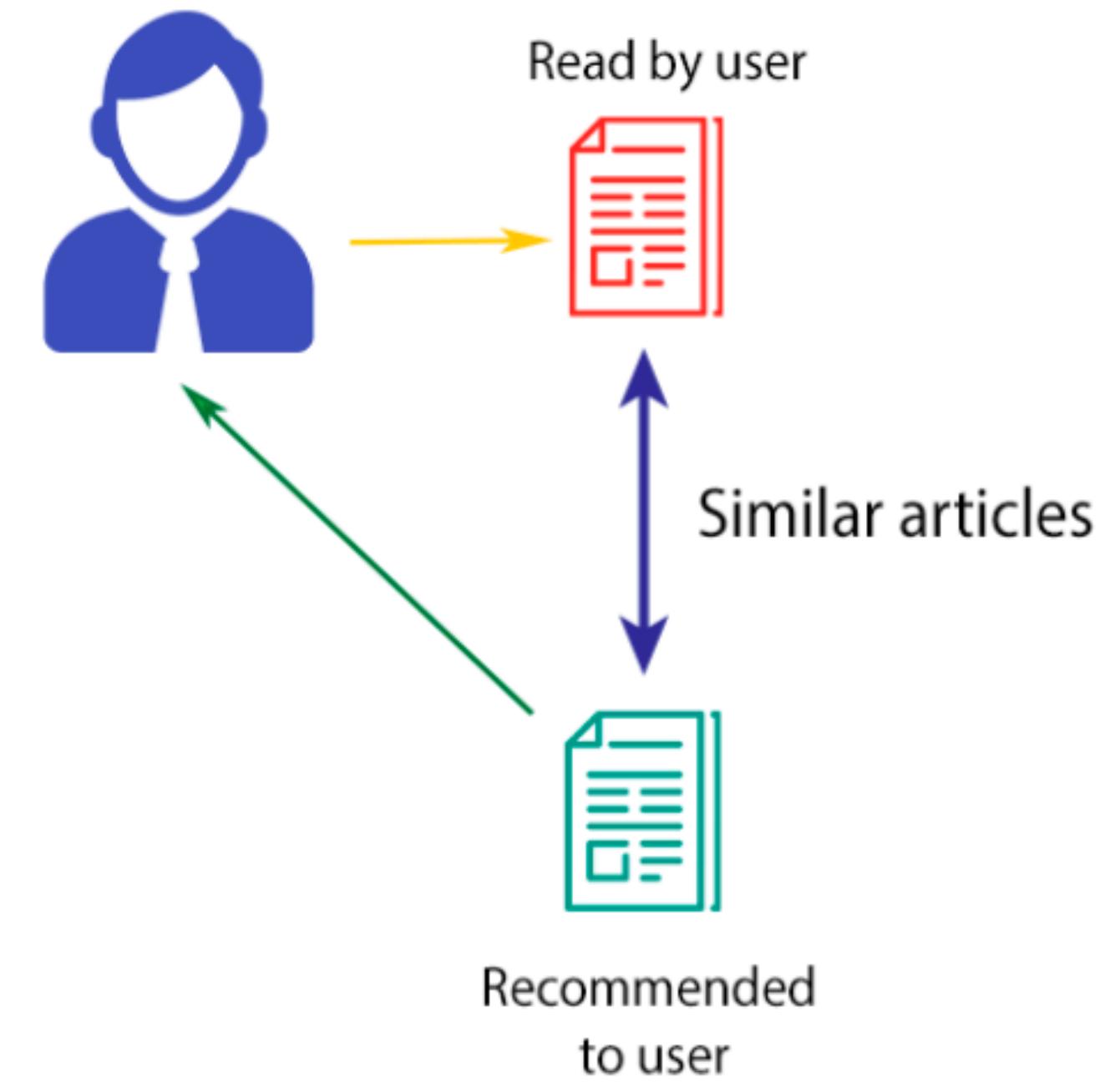
matplotlib

scikit  
learn

# Types of Recommenders



## CONTENT-BASED FILTERING



# Process

## 3 Steps

1. Reduce dimensions
2. Create user vector
3. Make recommendations:  
cosine similarity

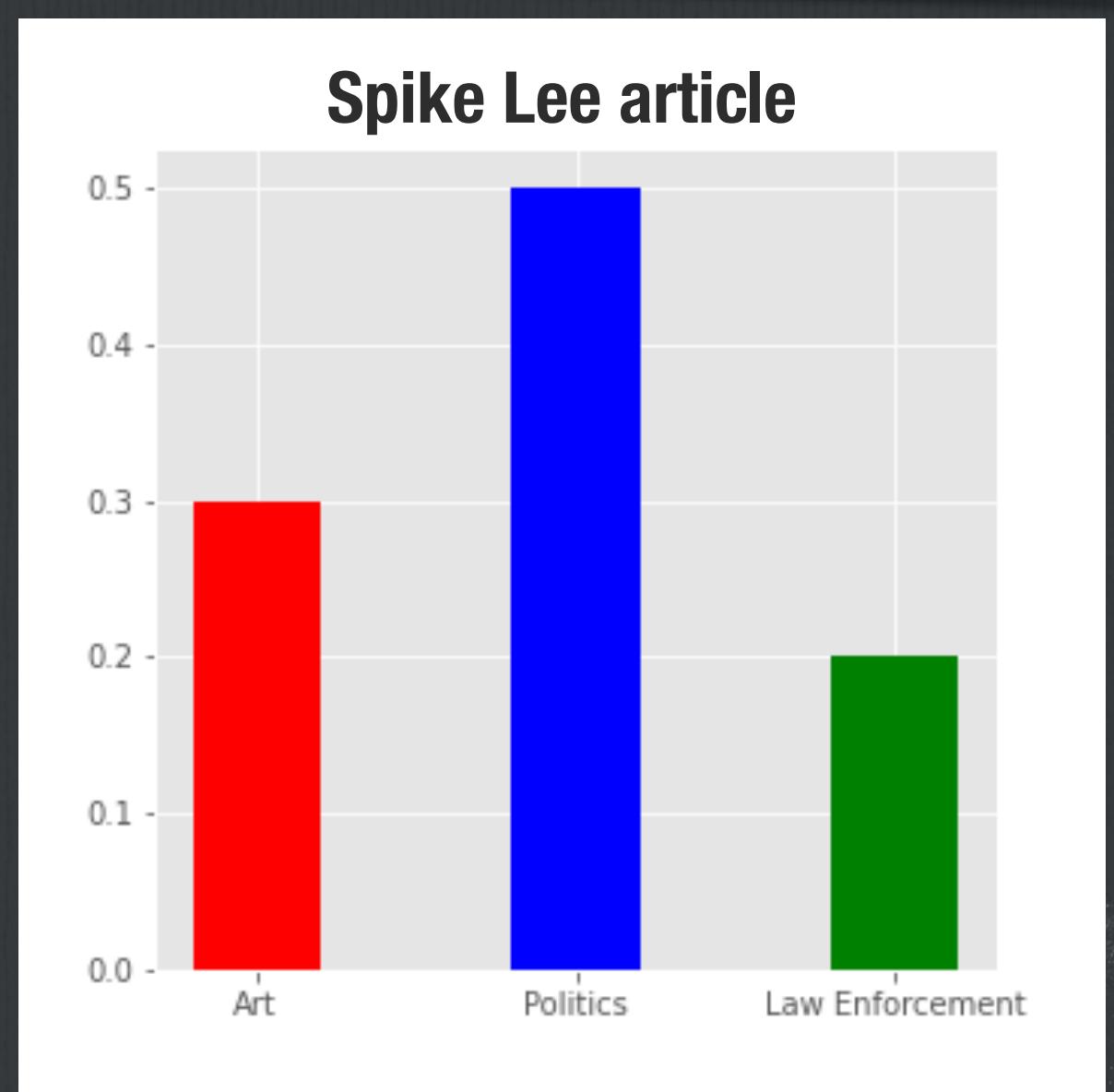
"Spike Lee tells Bushnell crowd: Wake up. This is no joke. This is the year of living dangerously"

'Spike Lee came to the Bushnell Center for the Performing **Arts** in Hartford on Saturday night to discuss his decades-long **career** as one of the country's most revered **filmmakers**. Lee also gave the **audience** strong opinions about **global warming**, **police**-black relations and the current resident of the **White House**, whom he wouldn't speak of by name, referring to him as **Agent Orange**. We have a man in the highest **office** of the **United States of America**, the so-called cradle of **democracy**, who feels he is above the **law** and also on top of that thinks that **global warming** is a crock, Lee told the audience. The EPA is dismantling all the stuff Obama put in place to save

Art  
art  
audience  
film

**Politics**  
global warming  
America  
democracy

**Law Enforcement**  
police  
office  
career



# Results

---

**About Jim:** “Jim mostly likes to read about sports and picks out a Hollywood news article every once in a while. Occasionally, he clicks an article on some other topic.”

## Previously Read

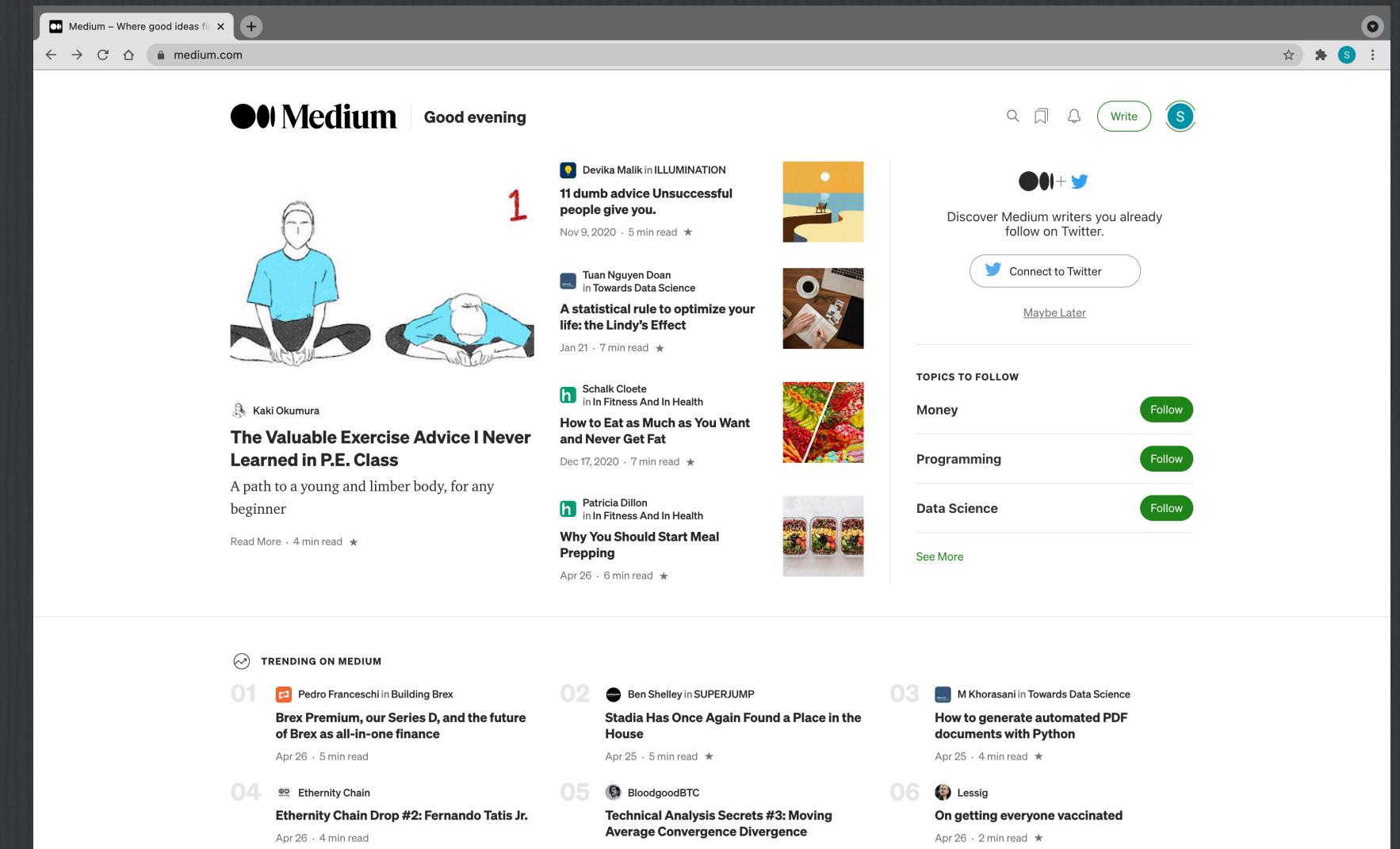
- |   |   |
|---|---|
| 1 | “High school sports with highest concession rates revealed in news study” |
| 2 | “Former NFL lineman Justin Bannan arrested for attempted murder”          |
| 3 | “Lakers and Nets players lost money on broken agreements in China”        |
| 4 | “Lindsay Lohan denies shading Miley Cyrus and Cody Simpson”               |
| 5 | “420 Pounds Of Marijuana Found In Truck At Darien Rest Stop: CSP”         |

## Recommended

- |   |   |
|---|---|
| 1 | “Half-Court Shots Earn 2 College Students Free Tuition”   |
| 2 | “The best story in the Texas high school football playoffs is Fort Worth's North Side”                |
| 3 | “‘Free Chase Young’: Criticism of the NCAA trends on social media after Ohio State star's suspension” |
| 4 | “Twitter reacts after Tom Brady tweets he and Rob Gronkowski are ‘reunited’”                          |
| 5 | “This youth football team has chance at a national title, but they need your help”                    |

# Future Work

- Get feedback from social network
- Content Recommender for [medium.com](https://medium.com)
- Deploy an app



GitHub  
[/sborodach](https://github.com/sborodach)

LinkedIn  
[/steven-borodach](https://www.linkedin.com/in/steven-borodach)

Email  
[steven.borodach@gmail.com](mailto:steven.borodach@gmail.com)



Thank you!