

AI-POWERED ARTICLE RECOMMENDATION SYSTEM

CIT-CODERELATE ROUND 3

Team Info::

- Team Name: Analysis Wizards
- Members: Sanjay, Murali, Simson
- Emails: sanjay310105@gmail.com,
muralivizard007@gmail.com,
simsonantorai2006@gmail.com
- Ph: 9443123389, 9345448797, 9940687863

PROBLEM STATEMENT

PROBLEM:

- With the ***ever-increasing amount of digital content***, users often struggle to discover relevant and engaging articles tailored to their interests
- Content overload on platforms like Medium,WordPress can overwhelm users. ***Personalized recommendations*** enhance user engagement and retention

PROPOSED SOLUTION

OVERVIEW

We developed a TF-IDF based recommendation system that suggests articles similar to the one selected by a user.

HOW IT WORKS?

- Merged title & content to form a rich representation.
- Used TF-IDF to extract important keywords
- Applied Cosine Similarity to compare articles.
- Deployed with an interactive Streamlit web app.

DATA PREPROCESSING

WORKFLOW

- Removed duplicates, irrelevant URLs, and non-English articles.
- **Cleaned the text:** lowercasing, removing punctuation, stopwords.
- Created a new column `cleaned_content = title + text`
- Handled missing values (`fillna("")` for text fields)

MODELS USED

TF-IDF VECTORIZATION

- Captures keyword relevance across the article
- Limited to top 5000 features for efficiency.

COSINE SIMILARITY

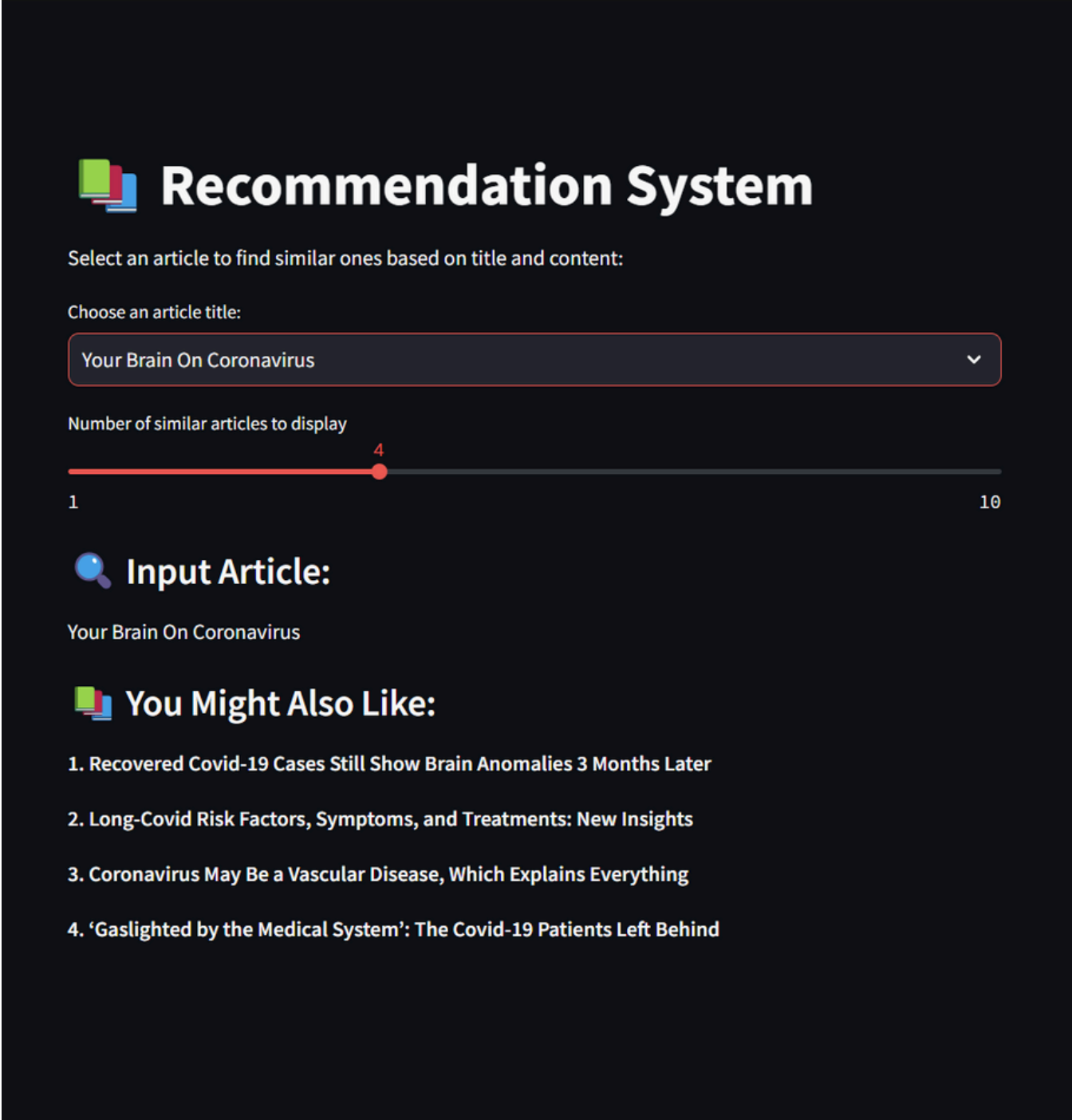
- Measures semantic similarity between article vectors.
- Higher similarity - better match for recommendations

MODEL OUTPUT

App Features:

- Select an article by title
- Choose number of recommendations (1-10)
- Displays similar article titles dynamically

UI Demo:



The UI demo shows a dark-themed interface for a 'Recommendation System'. At the top, there's a title 'Recommendation System' with a small icon of three overlapping squares. Below the title, a instruction says 'Select an article to find similar ones based on title and content:'. A dropdown menu labeled 'Choose an article title:' is set to 'Your Brain On Coronavirus'. Below this, a slider for 'Number of similar articles to display' is set to 4, with a range from 1 to 10. A section titled 'Input Article:' shows the selected article title 'Your Brain On Coronavirus'. Finally, a section titled 'You Might Also Like:' displays a list of four recommended article titles.

Recommendation System

Select an article to find similar ones based on title and content:

Choose an article title:

Your Brain On Coronavirus

Number of similar articles to display

1 4 10

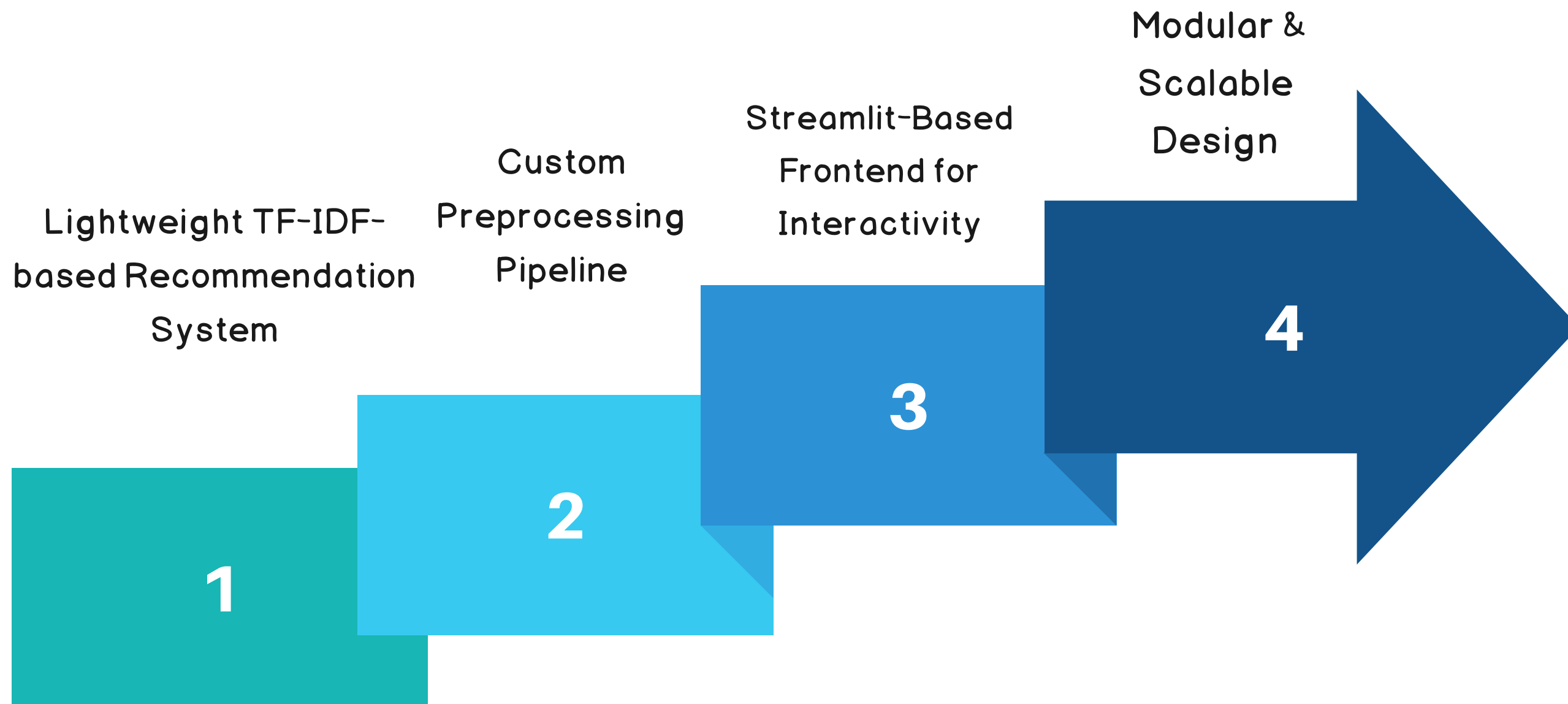
Input Article:

Your Brain On Coronavirus

You Might Also Like:


1. Recovered Covid-19 Cases Still Show Brain Anomalies 3 Months Later
2. Long-Covid Risk Factors, Symptoms, and Treatments: New Insights
3. Coronavirus May Be a Vascular Disease, Which Explains Everything
4. 'Gaslighted by the Medical System': The Covid-19 Patients Left Behind

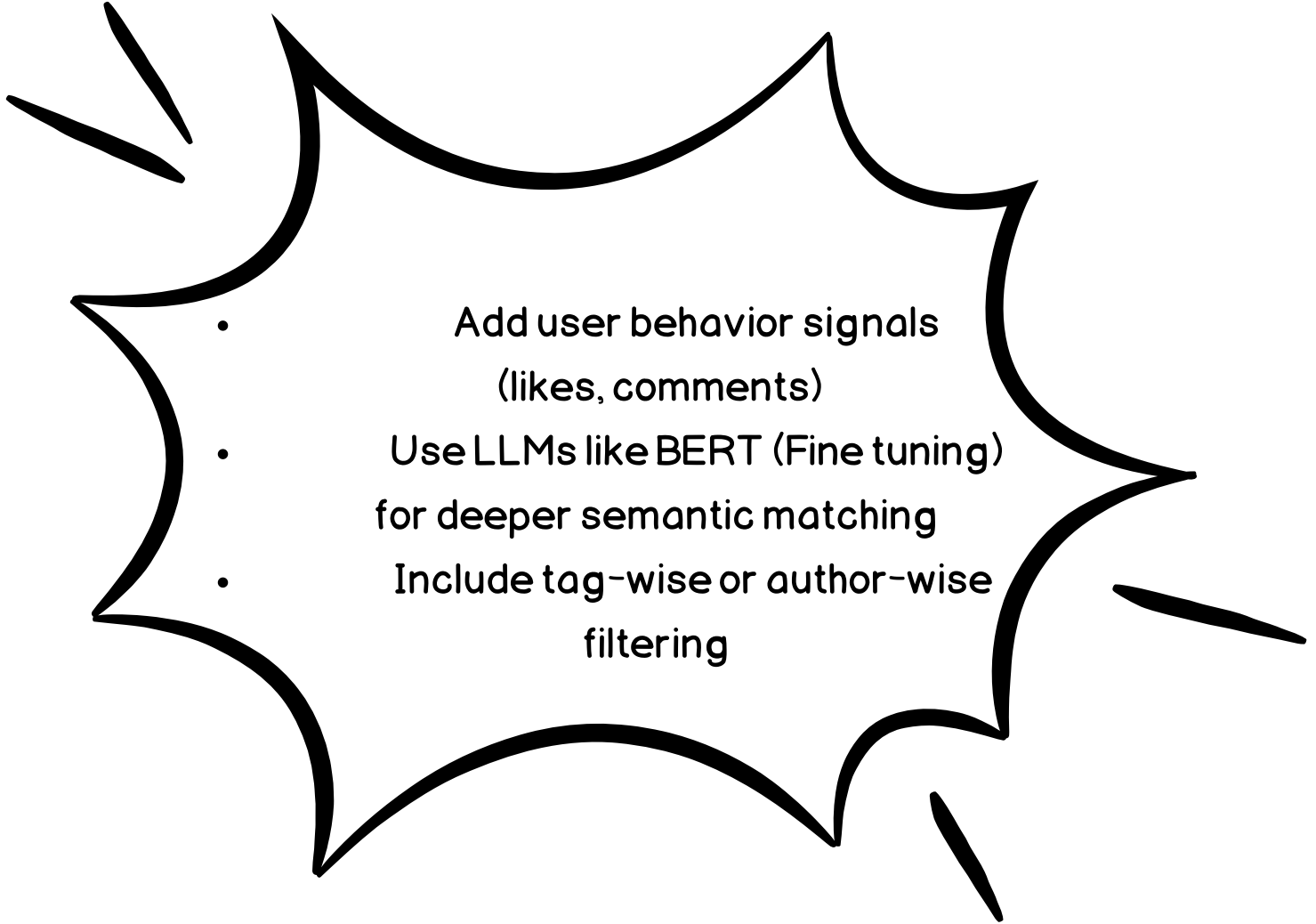
INNOVATION



FUTURE SCOPE

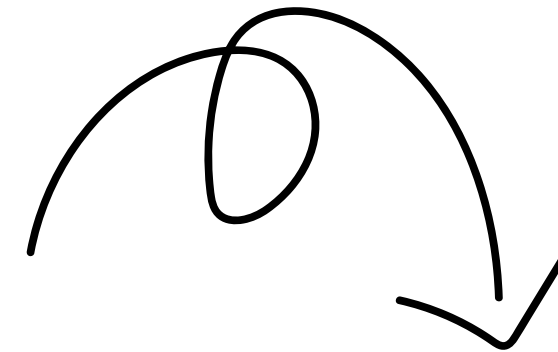


- TF-IDF is effective for short content pieces
 - Title + text combination improves relevance
 - Simpler models can still deliver impactful results
- 

- 
- Add user behavior signals (likes, comments)
 - Use LLMs like BERT (Fine tuning) for deeper semantic matching
 - Include tag-wise or author-wise filtering

CONCLUSION

- Built a personalized recommendation system
- Used explainable AI (TF-IDF + Cosine Similarity)
- Successfully deployed a working demo



IMPACT:

- Helps users find relevant content faster
- Improves content engagement

The background features abstract geometric elements in teal. In the top-left and bottom-left corners, there are nested rectangular outlines. In the top-right and bottom-right corners, there are clusters of small teal circles arranged in a grid pattern. Diagonal lines also cross the background from corner to corner.

THANK YOU

for your time