



L&T
NEUROHACK

Tactical Trends

AI based news aggregator

NEURAL KNIGHTS (PICT)

SWARUP POKHARKAR

VIRAJ MANE

PRAJWAL PADOLE

PORNIMA SADASHIV



INTRODUCTION

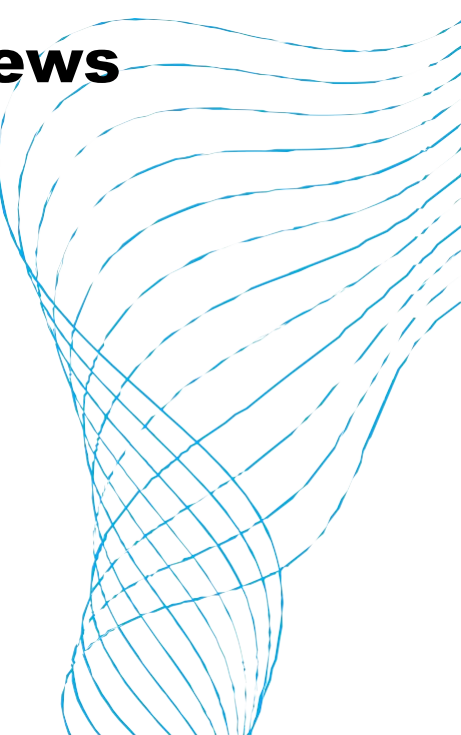


GOLD PROBLEM STATEMENT:

Develop an innovative AI based news aggregator application customized explicitly for strategic & geopolitical significance (Defence centric).

OBJECTIVE:

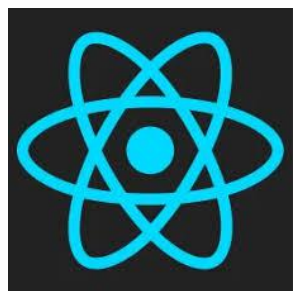
Tactical Trends curates strategic and defense-related content, providing users with personalized, relevant news from domestic and international sources. Utilizing machine learning and NLP, the platform delivers concise summaries, insightful analysis, and tailored recommendations based on user preferences, ensuring an enhanced and focused news consumption experience.



TECH-STACK



L&T
NEUROHACK



django
REST

django



BeautifulSoup



HUGGING FACE



News API

CELERY



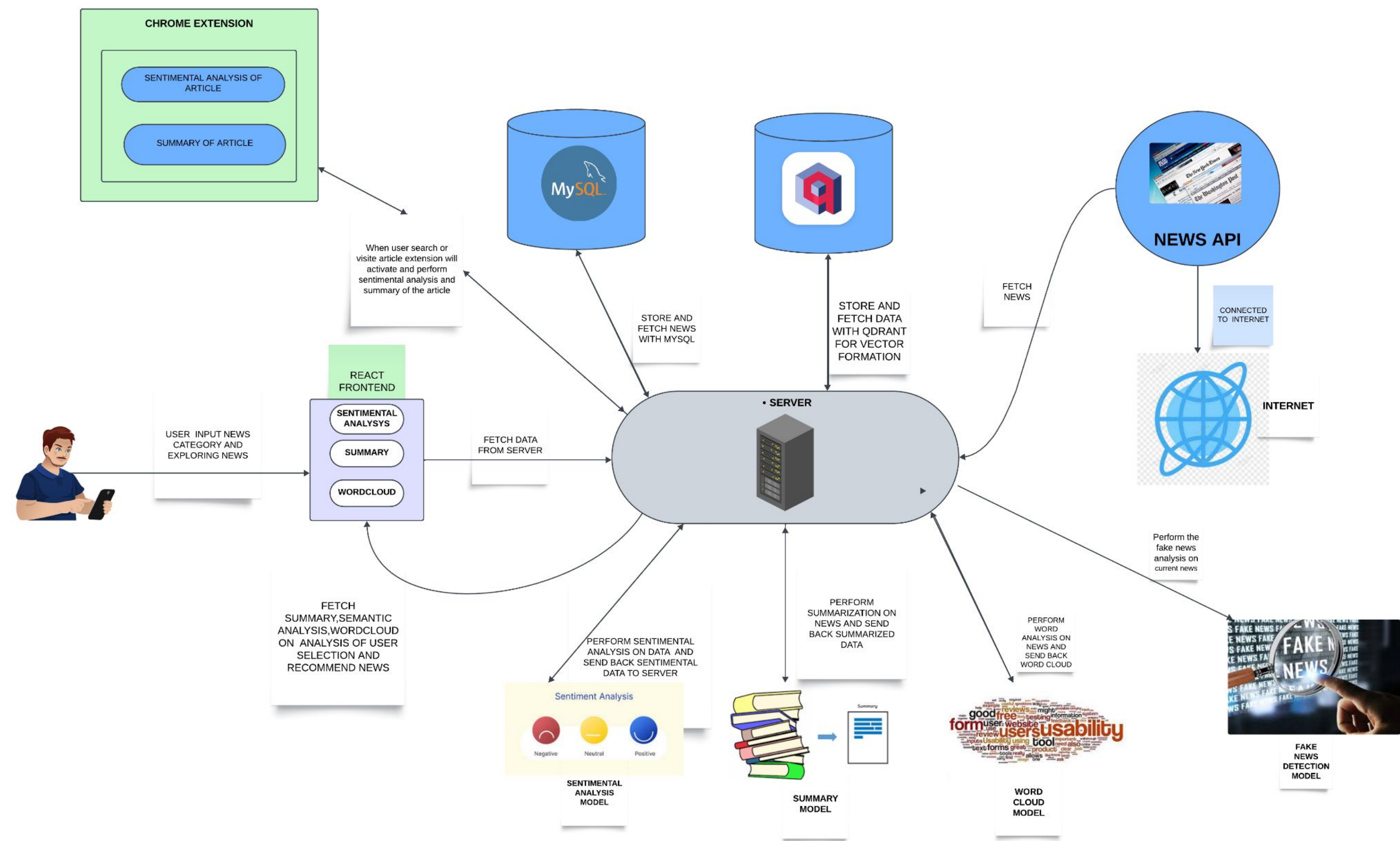


L&T
NEUROHACK

SYSTEM ARCHITECTURE



L&T
NEUROHACK



SYSTEM ARCHITECTURE FOR TACTICAL TRENDS

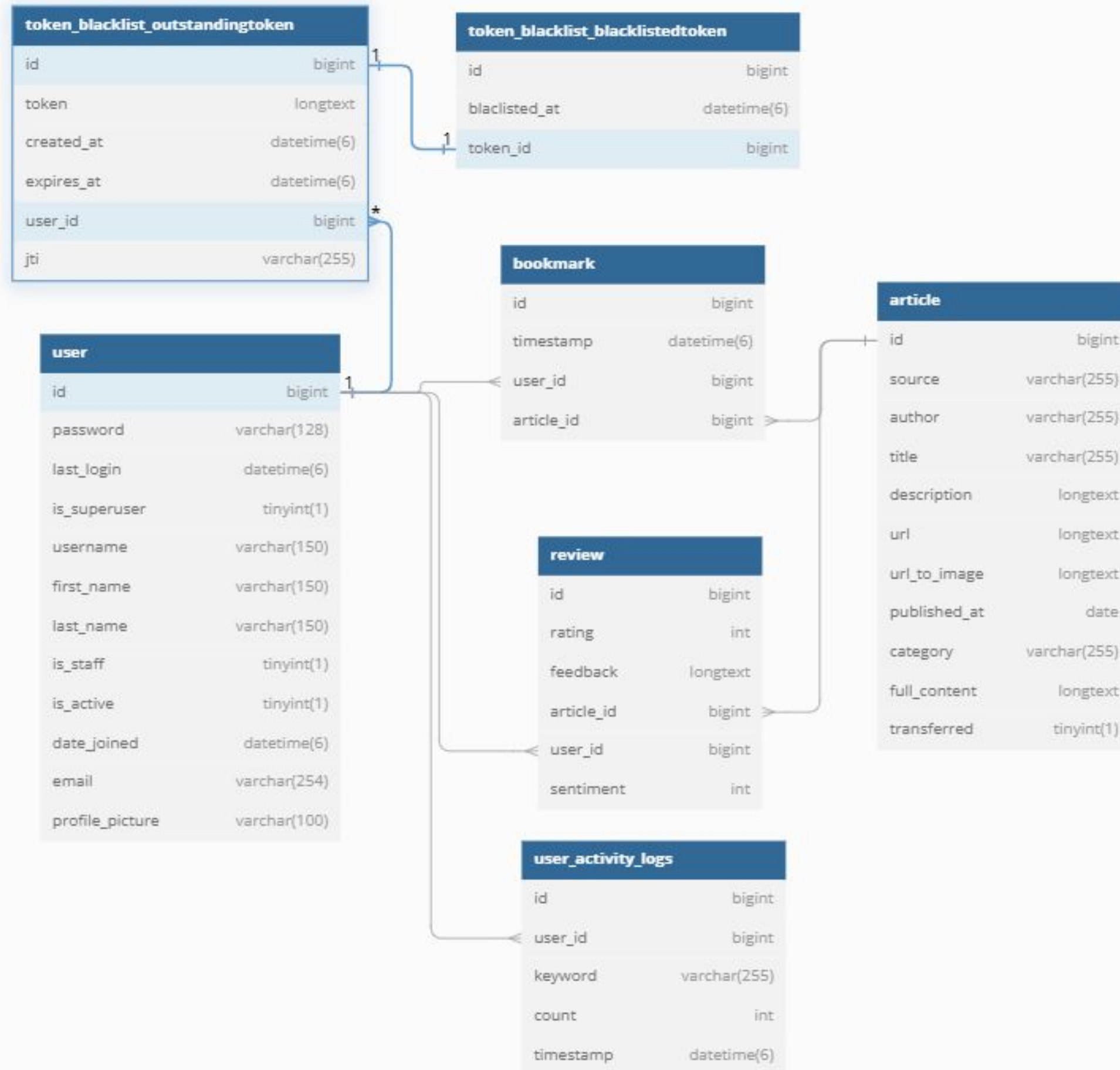


L&T
NEUROHACK

DATABASE DESIGN



L&T
NEUROHACK



DB-SCHEMA FOR RELATIONAL DATABASE



HYBRID DATABASE STRATEGY FOR RECOMMENDATION SYSTEM



CHALLENGE:

To effectively recommend articles to users based on their interests, it is essential to store and analyze the articles they have liked in the past. Traditional relational databases often struggle with performing similarity searches efficiently, which can hinder the ability to generate accurate recommendations.

SOLUTION:

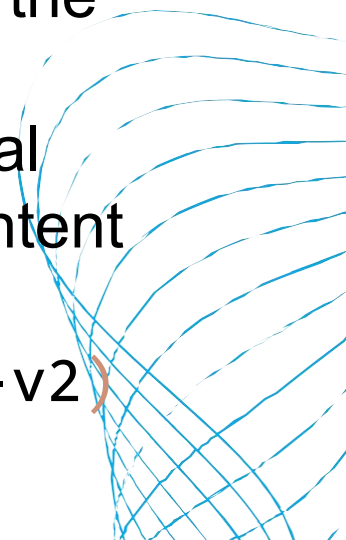
To enhance our article recommendation system, we are utilizing **Qdrant**, a vector database, to store and manage article and search activity embeddings effectively. This approach allows us to perform similarity searches based on article content and user activity, improving the relevance of recommendations for users.

Article Collection

- **Article ID:** A unique identifier for the article.
- **Vector Embeddings:** A numerical representation of the article's content generated using a pre-trained model(paraphrase-MiniLM-L6-v2)

Keyword Collection

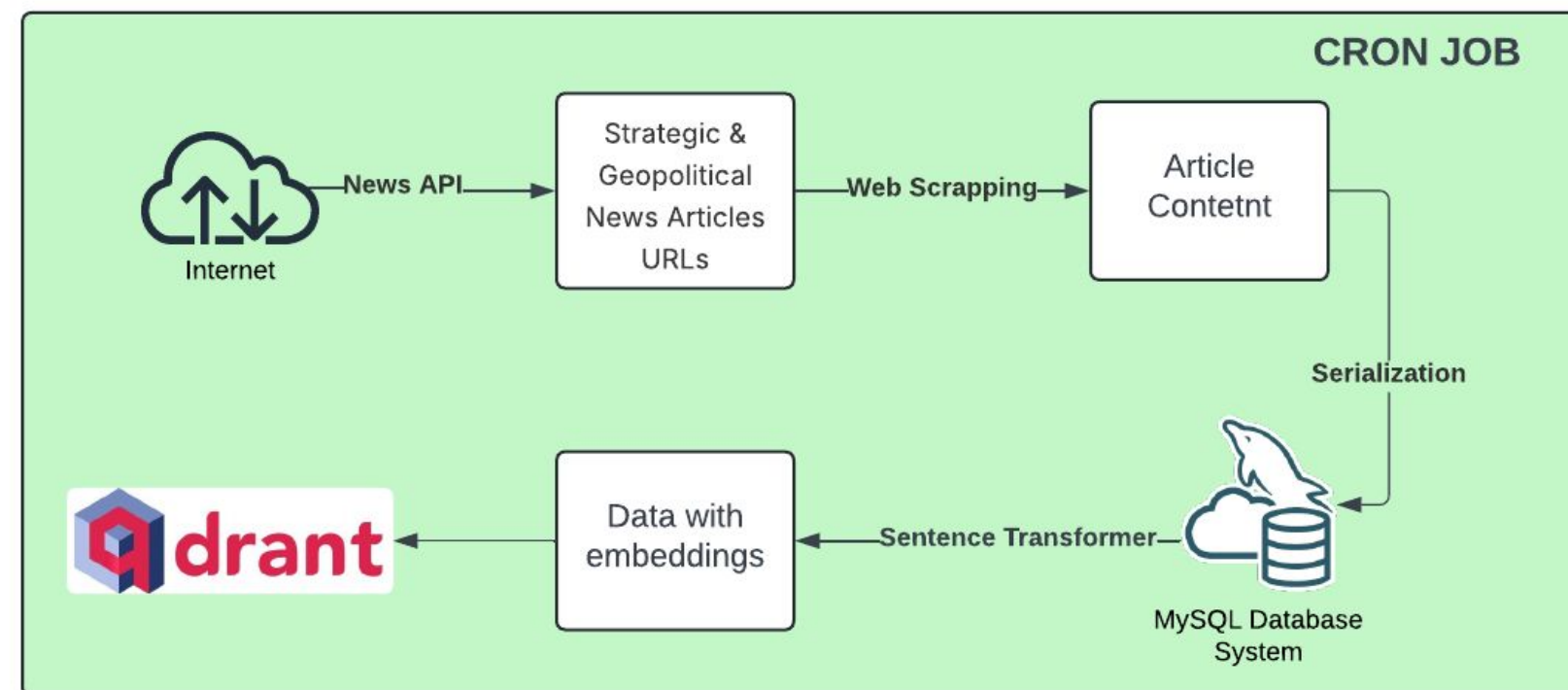
- **Article ID:** A unique identifier for the keyword.
- **Vector Embeddings:** A numerical representation of the article's content generated using a pre-trained model(paraphrase-MiniLM-L6-v2)



FEATURES

- **News Aggregation:**

News articles are fetched using the News API and stored in a MySQL database. Embeddings are generated using the Sentence Transformer (paraphrase-MiniLM-L6-v2) and stored in a Qdrant vector database for retrieval.



FEATURES



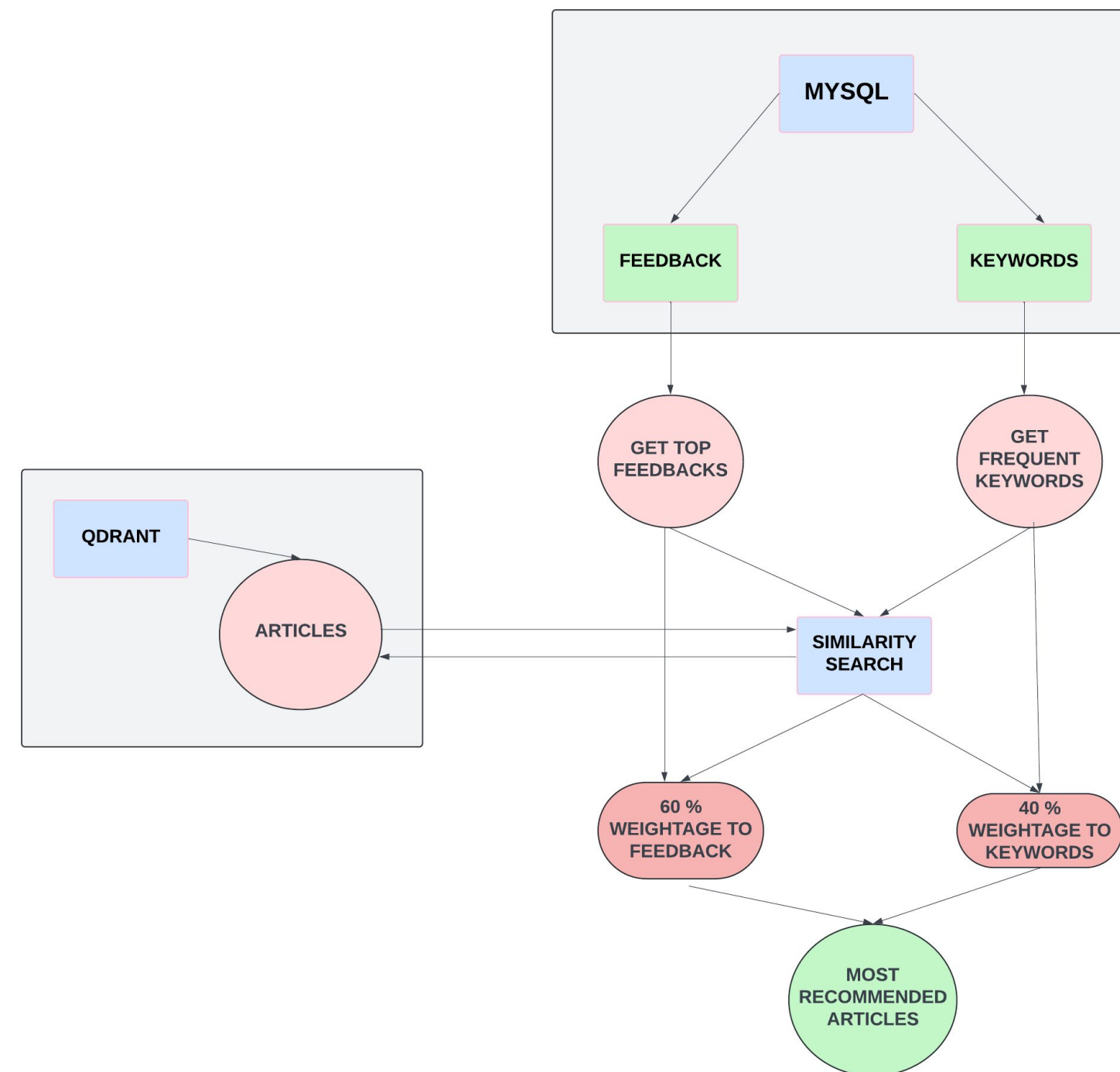
L&T
NEUROHACK

- **Keyword Search:**
Users can search for articles by keywords, retrieving relevant content from the database.
- **Word** **Cloud:**
User will get the key words from the news articles
- **Summarization:**
Users can generate quick article summaries using the sshleifer/distilbart-cnn-12-6 model for an efficient overview.

FEATURES

- **Content-Based Recommendations:**

Based on user interactions, similar articles are recommended through content-based filtering using article embeddings stored in Qdrant.



FEATURES



- **Sentiment Analysis:**

Users can analyze an article's sentiment using the distilbert-base-uncased-finetuned-sst-2-english model, which classifies the content as either positive or negative, providing a quick insight into its emotional tone.

- **Continuous Feedback:**

User interactions (bookmarks, reading habits) provide feedback to improve the recommendation system, ensuring personalized, relevant content discovery.



EXTRA FEATURES



- **Fake News Detection:**

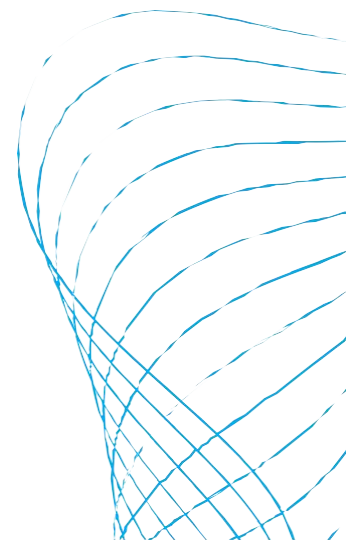
Users can analyze whether an article is fake or not using the `vikram71198/distilroberta-base-finetuned-fake-news-detection` which classifies the content as either fake or not fake, providing a quick insight into its reliability.

- **Bookmarking System:**

Users can save and organize articles with tags for easy reference, syncing bookmarks across devices for seamless access.

- **Chrome Extension:**

User interactions (bookmarks, reading habits) provide feedback to improve the recommendation system, ensuring personalized, relevant content discovery.





L&T

NEUROHACK

Business Model

Tactical Trends Business Model

- **Value Proposition:** Personalized news aggregation with sentiment analysis, word clouds, and summaries for quick insights. Enhanced with a web extension for seamless access and real-time updates.
- **Target Market:** General public, journalists, businesses
-
- who want to stay informed and analyze trends.
- **Revenue Streams:**
 - Freemium model (basic free access + premium subscription).
 - Ad revenue from targeted ads and sponsored content.
 - Data analytics and custom reports.
 - Affiliate marketing through partnerships.
- **Key Activities:** Aggregating and analyzing news, developing features, providing customer support, and marketing.
- **Channels:** Website, web extension, social media, and potential future mobile app.



CHALLENGES:



L&T
NEUROHACK

Challenge:Model Scalability with Increasing Data Volume

Solution:

- Use distributed processing tools like Apache Spark to handle large datasets.
- Implement horizontal scaling for storage and compute using cloud platforms.

Challenge:Handling Multilingual News Articles

Solution:

- Integrate translation models or APIs (e.g., Google Translate API) to process content in different languages.
- Use multilingual NLP models such as `mBERT` or `XLM-R` for accurate sentiment and fake news detection.

Challenge: Ensuring Real-Time Updates of News

Solution:

- Optimize cron jobs to fetch articles more frequently based on user activity and interests.
 - Use event-driven architecture with streaming platforms like Kafka for real-time article ingestion.
- 



CHALLENGES:



L&T
NEUROHACK

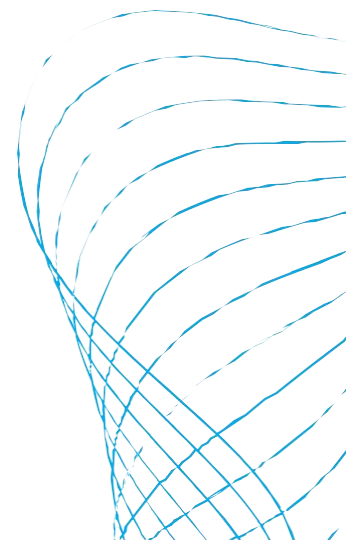
Challenge: Privacy Concerns with User Data Storage and Feedback

Solution:

- Store user data in encrypted formats and limit access through role-based authentication.
- Offer opt-in mechanisms for user feedback collection, ensuring GDPR compliance.

Challenge: Maintaining Accuracy of Sentiment and Fake News Detection Models

Solution:

- Periodically retrain models on updated datasets to minimize bias and improve accuracy.
 - Integrate human-in-the-loop systems to manually validate and correct model predictions when needed.
- 



L&T
NEUROHACK

Thank You