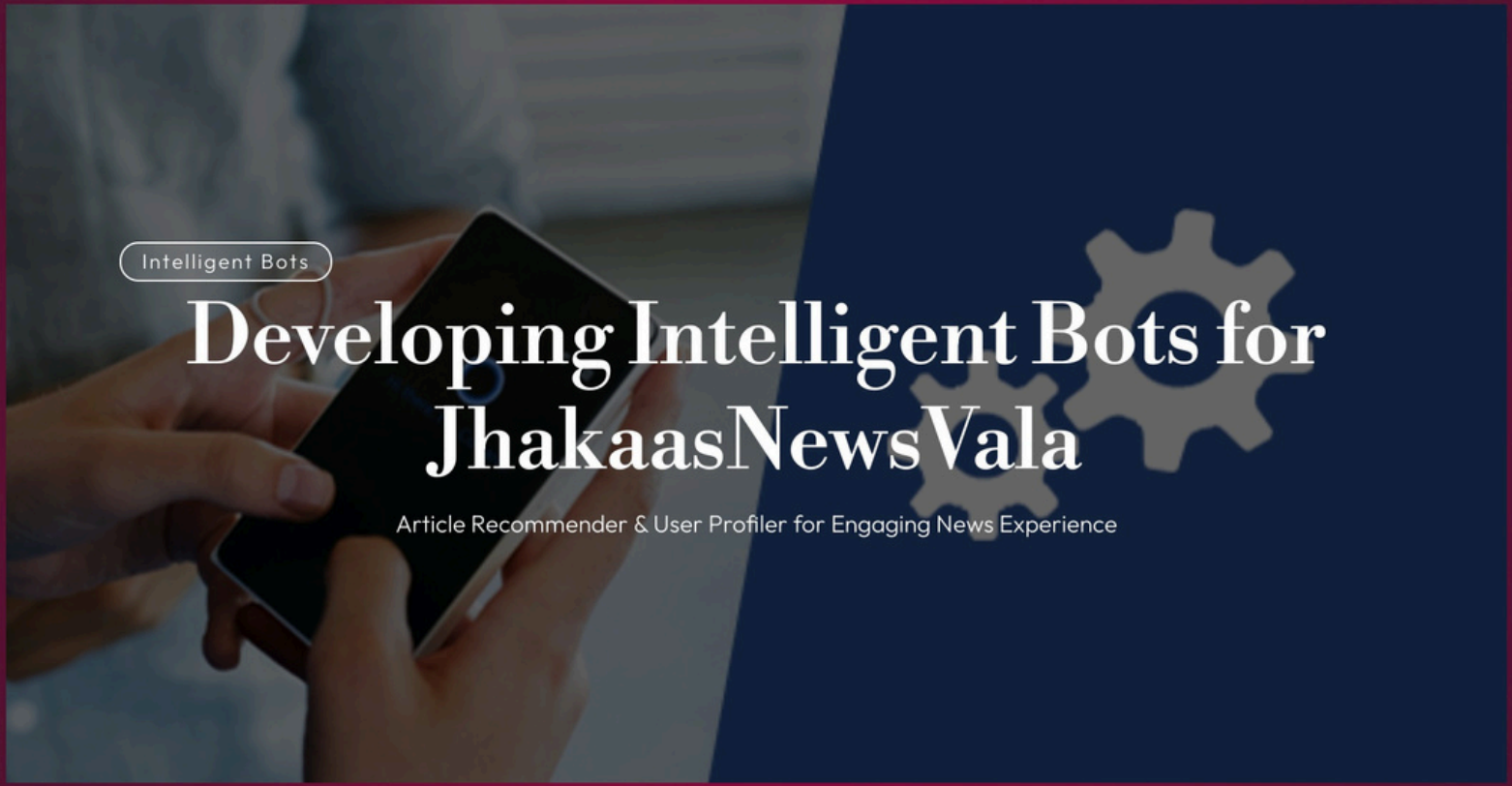
# Enhancing News Engagement with Intelligent Bots

Exploring the development of an article recommender and user profiler for JhakaasNewsVala to boost user engagement in news reading.

AYUSH DUTTA


Intelligent Bots
Developing Intelligent Bots for JhakaasNewsVala
Article Recommender & User Profiler for Engaging News Experience

Market Challenges

# Understanding the Problem

Overview of Challenges Faced by Working Professionals

### Target Market

The primary audience targeted consists of working professionals aged 21-40, who often seek efficient solutions to manage their time and information.

### Challenges

Several challenges hinder user engagement, including a lack of user data on first visits, limited screen space for displaying multiple articles, and the need to make a strong impression to encourage retention.

### User Data Limitations

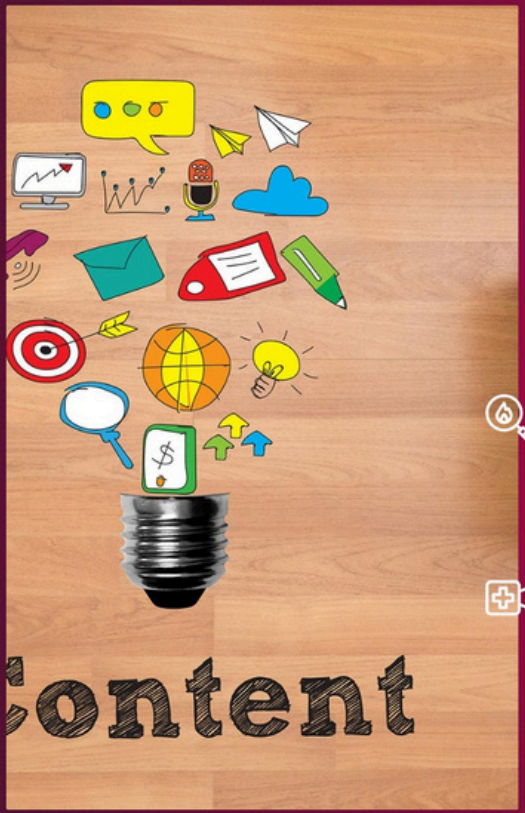
Without prior user data, understanding preferences and behaviors becomes difficult, which can lead to generic content that fails to resonate with users.

### Screen Space Constraints

With limited screen space, presenting up to 10 articles effectively poses a challenge, requiring innovative display strategies to maximize visibility and engagement.

### Retention Strategies

The objective is to reduce bias in content presentation, maximize initial engagement during the first visit, and personalize recommendations to encourage users to return.




Strategy Overview

# Key Strategy Introduction

Explore-Exploit Balance in Content Curation

## Top 5 Articles

These articles represent trending topics and high click-through rate (CTR) stories, making them crucial for immediate audience engagement and retention.

## Next 5 Articles

Focus on diverse and underrepresented stories that provide fresh perspectives, fostering exploration and broadening the audience's interests.

Bias Mitigation

# Bias Mitigation Tactics

Effective Strategies to Counteract Bias in Data Analysis

1

### Rotate Article Ranks

Implement a strategy to modify the ranking of articles to minimize positional bias, ensuring that all articles have an equal chance of being seen regardless of their position.

2

### Use Adjusted Metrics

Adopt metrics like propensity scoring that account for selection bias, providing a more accurate representation of data and improving the validity of conclusions drawn from it.

User Profiling

# User Profiling Approach

Dynamic User Profiles from Clickstream Data

1. **User Profiling**

   Developing user profiles dynamically based on interaction patterns allows for personalized experiences and better engagement.

2. **Clickstream Data**

   This approach utilizes clickstream data, capturing users' online behavior including the specific topics they click on.

3. **Topics Clicked**

   By analyzing the topics users show interest in, businesses can tailor content and recommendations to meet their preferences.

4. **Time Spent**

   Tracking the amount of time users spend on various topics helps identify what captures their attention most effectively.

5. **Skip Patterns**

   Understanding skip patterns reveals what content users find less engaging, allowing for continuous improvement of offerings.

Coverage Maximizati...

# Coverage Maximization Method

Utilizing Clustering for Enhanced Topic Diversity

## Diverse Topic Representation

Employ clustering techniques to group similar topics, ensuring a wide range of subjects are represented. This approach minimizes redundancy and promotes a richer dialogue among various ideas.

Data Preparation

# Data Preparation Process

Step 1: Data Preparation in the Implementation Plan

### Preprocess news corpus

The initial step involves preparing the news data for analysis, ensuring that it is clean and structured.

### Extract keywords, topics, and categories

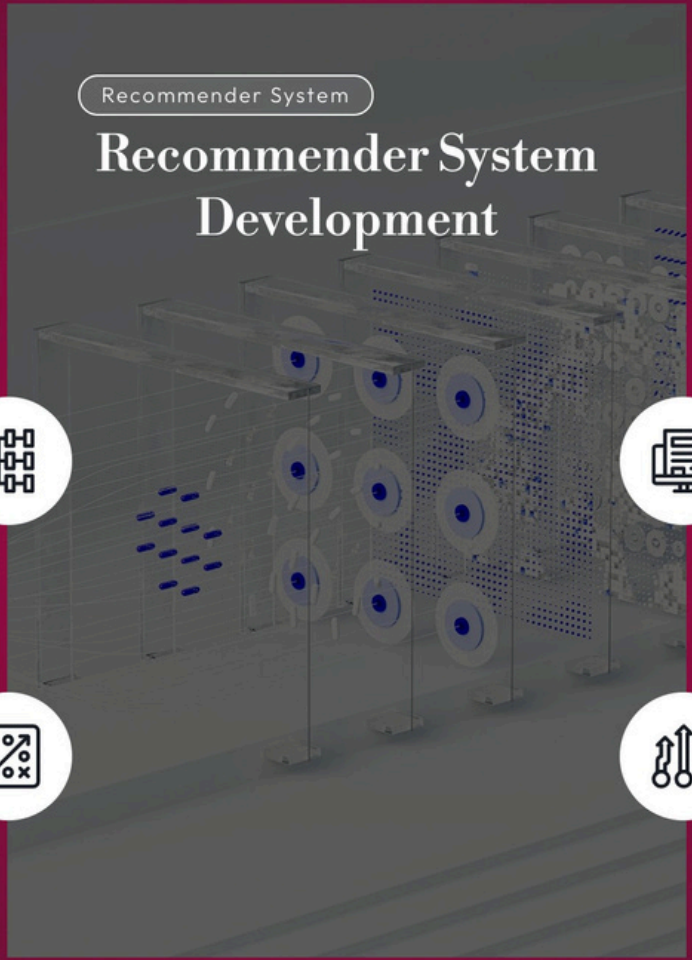
Utilize Natural Language Processing (NLP) techniques to identify and categorize essential components of the news articles.

### Cluster articles into topics

Apply k-means clustering algorithms to group articles based on similar content, facilitating easier analysis and insights.

Recommender System

# Recommender System Development



### Step 2: Recommender System

This step involves the implementation of a recommender system tailored to user preferences and behaviors.

### Apply multi-armed bandit algorithm

Utilize the multi-armed bandit algorithm to balance exploration and exploitation in content recommendations.

### Explore diverse topics

Encourage users to discover varying subjects and articles, enhancing their engagement and satisfaction.

### Exploit popular articles

Leverage trending and highly-rated articles to attract user attention and increase click-through rates.

User Profiling

# Dynamic User Profiling Techniques

Step 3: Dynamic User Profiling

### Analyze clickstream data in real-time

Monitor user interactions with content as they occur, capturing valuable insights into user behavior.

---

### Record clickthroughs, time spent, and ignored articles

Collect metrics on which content engages users, how long they stay, and what they overlook, enhancing future content strategies.

---

### Build content-based or collaborative filtering models

Utilize collected data to develop models that suggest relevant content based on user preferences or similar user behaviors.

Bias Mitigation

# Effective Bias Mitigation

Strategies for Reducing Bias in Article Ranking Systems

1

### Rotate article ranks dynamically.

Implementing dynamic rotation of article rankings helps to reduce bias by ensuring that different articles are highlighted over time, preventing any single article from dominating visibility.

2

### Normalize CTR by rank.

Normalizing click-through rates (CTR) according to rank allows for a fair evaluation of articles, ensuring that the performance of articles is assessed in a consistent and unbiased manner.

Project Timeli...

# Project Timeline

Overview of Project Phases and Activities from Week 1 to Week 8

## Week 1-2

### Data Preparation

In the first two weeks, the focus will be on cleaning the corpus and extracting relevant topics to set a solid foundation for the subsequent phases.

## Week 3-4

### Recommender System

During weeks three and four, the implementation of a multi-armed bandit algorithm will take place, enhancing the system's ability to make personalized recommendations.

## Week 5-6

### User Profiling

In weeks five and six, we will conduct clickstream analysis and develop user model to better understand user behavior and preferences.

## Week 7-8

### Testing & Refinement

The final two weeks will involve A/B testing and iterative improvements based on feedback to ensure the recommender system performs optimally.

Bias Mitigation

# Challenges in Bias Mitigation

Understanding Challenges and Solutions in Mitigating Bias

### Challenge

Articles that are ranked higher tend to receive more clicks, leading to biased visibility and potential misinformation.

1

### Solution: Rotation of Ranks

By rotating the ranks of articles across different users, we can ensure that all articles receive fair exposure, reducing the likelihood of bias.

2

### Solution: Adjusted CTR Scores

Using adjusted Click-Through Rate (CTR) scores for ranking articles can help in creating a more equitable assessment of content relevance and quality.

3

Explore-Exploit

# Explore-Exploit Balance in Recommender Systems

Understanding Multi-Armed Bandit Approach

1

### Multi-Armed Bandit

A strategic approach that balances between exploring new topics and exploiting popular articles to optimize user engagement.

2

### Exploration vs. Exploitation

Exploration involves introducing users to new content, while exploitation focuses on presenting well-received articles that users are likely to engage with.

3

### Dynamic Updates

The system continually adapts to user behavior through clickstream data, ensuring the recommendations evolve based on real-time feedback.

User Insights

# User Profiling Insights

Understanding User Behavior through Clickstream Data

## Input Data

Clickstream data is collected, including articles clicked, time spent on each article, and articles that users chose to ignore. This data serves as the foundation for understanding user behavior.

## User Preferences: Topics

Analyzing the input data reveals user preferences for specific topics, such as technology and sports, allowing for tailored content delivery.

## User Preferences: Engagement Depth

The amount of time users spend on articles indicates their level of engagement, helping to identify which content resonates most with them.

## User Preferences: Behavioral Patterns

Patterns in user behavior, such as frequently skipping certain categories, provide insights into content that may not align with user interests.

Outcomes And Metrics

# Expected Outcomes and Evaluation

Understanding the Impact of Engagement Strategies

### First Visit Impact

Achieving higher clickthrough rates (CTR) is crucial in capturing initial user interest, supported by diverse and engaging content that attracts users effectively.

### Long-Term Engagement

Dynamic personalization from the second visit onward enhances user experience and ensures increased retention and app usage frequency, fostering a loyal user base.

### CTR (Clickthrough Rate)

This metric measures engagement levels for the top 10 articles, providing insights into content performance and user preferences.

### Diversity Score

Assessing the diversity score ensures that content covers various clusters and topics, promoting comprehensive user engagement across interests.

### Retention Rate

Monitoring retention rate helps in understanding the frequency of user returns, indicating the effectiveness of engagement strategies over time.

### Time Spent per Session

This metric indicates the depth of interest users have in the content, reflecting user engagement levels and content relevance.

Strategy Overview

# Strategy Recap and Discussion

Key Takeaways and Interactive Discussion

1. **Explore-Exploit Strategy** → This strategy focuses on balancing the exploration of new options and exploiting known successful recommendations to enhance user satisfaction.

2. **Dynamic User Profiling** → Utilizes real-time data to create personalized experiences for users, adapting to their preferences and behaviors to improve engagement.

3. **Bias Mitigation** → Ensures fair data collection by actively identifying and reducing biases, leading to more equitable treatment of all user segments.

4. **Clustering Techniques** → Employs clustering methods to maximize coverage across user demographics, ensuring diverse representation in recommendations.