**Objective:**

**Catchment in investigating user motivations versus exploiting trends to maximize user engagement at the first encounter**

**Strategy**

1. Explore-Exploit with Article Diversity:

* Top 5 Articles: Select trending/popular stories to maximize immediate clickthrough (exploit).
* Next 5 Articles: Include articles from underrepresented disciplines to find out what users want (explore).

2. Reduce Bias:

* Restructure article ranks in the community are used to quantify positional bias.
* Per the use of inverse CTR metrics (propensity score) to perform balanced evaluation of article click-through rates.

3. Dynamic User Profiling:

* Build a profile using the clickstream data:
* Topics clicked.
* Time spent (indicates interest depth).
* Session behavior (e.g., articles ignored).

4. Coverage Maximization:

* Aggregate articles by topics based on metadata (i.e., subject keyword mining in natural languages).
* Selects one article in each cluster at the onset of the feed.

**Implementation Plan**

1. Prepare Data:

* Preprocess the news corpus: Extract topics, keywords, and categories using NLP.
* Cluster articles into clusters using algorithms like k-means or hierarchical clustering, etc.

2. Build Recommender:

* Use a multi-armed bandit algorithm (e.g., Thompson Sampling):
* Explore less popular articles to learn preferences.
* Exploit trending articles to maintain engagement.

3. User Profiling:

* Record and analyze: - Record and analyze:
* Articles clicked (topics, categories).
* Time spent (interest depth).
* Skip patterns (disinterest signals).
* Adopt collaborative filtering under the limited data and limited contents, adopt content-based filter under the rich contents.

4. Bias Mitigation:

* Rotate article ranks within clusters for different users.
* Normalize CTR by rank position for fair evaluation.

5. Evaluate and Improve:

* A/B test recommendations to compare CTR and engagement.

**Outcome:**

1. Improved Retention Rates: Users feel valued because of their uniqueness in interests, thereby leading to the increase of repeat visits.

2. Optimized Content Discovery: Users can be presented with trending or niche articles, improving the user experience and promoting longer session times.

3. Equitable Content Exposure: Less mainstream/niche posts become visible, ones that help a variety of content creators and expand on user engagement.

4. CTR Growth Over Time: Continuous learning of user behavior provides increasingly optimal recommendations, which increases clickthrough rates.

5. Scalable Personalization Framework: A powerful system that scales to expansive data sets, allows for continued usage, and has broad appeal among a growing community of users will sustain the interest of a growing audience.