

NEWS ARTICLE RECOMMENDER

VANGALA SAI HEMANTH

IIT ROORKEE

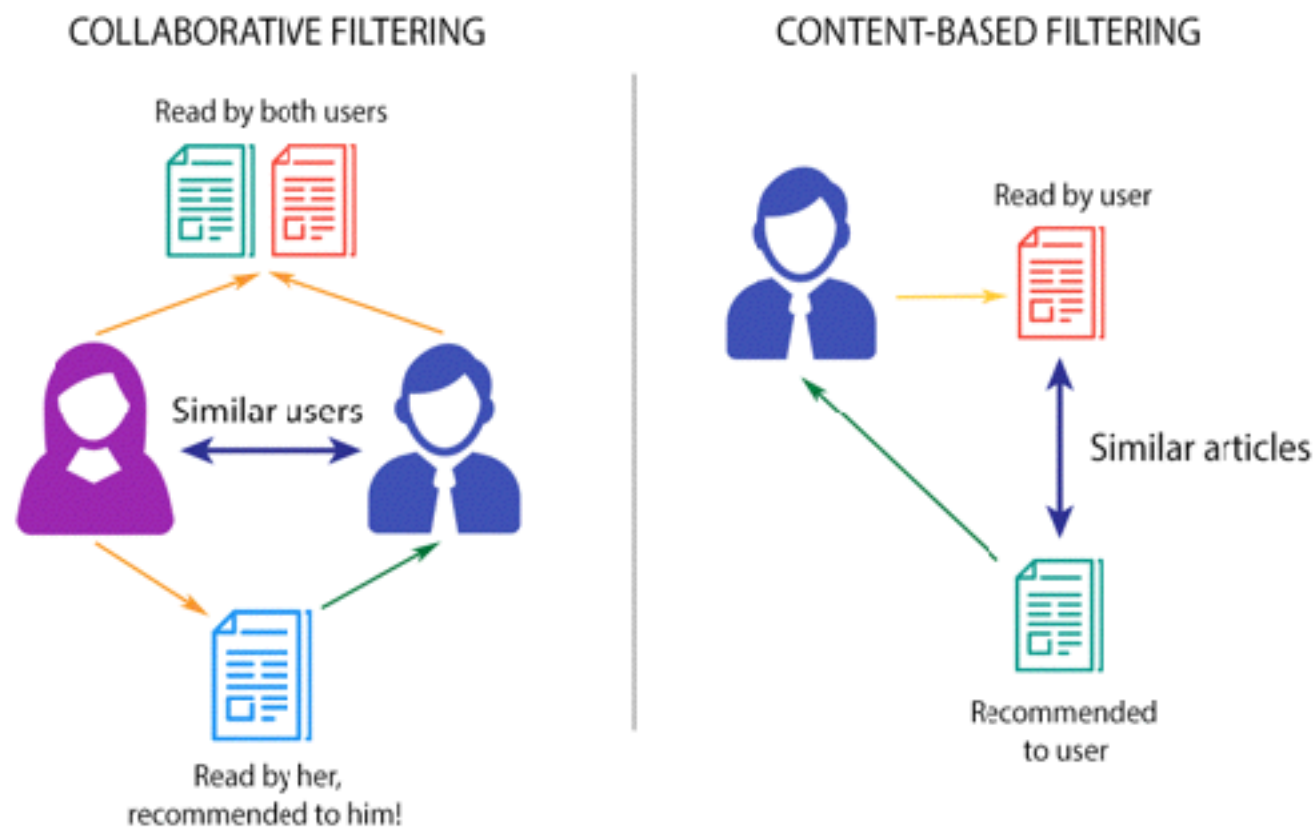
APPROACH

The factors that could influence an user to click an ad are

- ▶ Description of the Advertisement
- ▶ Position of the Ad
- ▶ Mood of the user
- ▶ Browser ,Reference url

FEATURE ENGINEERING

- ▶ **User to User** :- Collaborative filtering technique is used to calculate the probability of an ad getting clicked.
- ▶ **User to Item**:- Content based filtering technique is used to calculate the probability of an ad getting clicked.



FEATURE ENGINEERING

- ▶ **Ad_Count**:- The no of unique item_ids clicked by user including UVH
- ▶ **Item_id_visited**:- No of times the user clicked/User visited the specific item id.
- ▶ **per_item_id**:-The probability of the ad-unit to be clicked is calculated from train and impression data
- ▶ **per_adunit_id**:-The probability of the ad in a specific ad_unit to be get clicked
- ▶ **Per_Reference_url**:-The probability of an ad in a specific url to be clicked
- ▶ **Time**:- This is a categorical feature representing the time impression
- ▶ **Site**:-This is a categorical feature representing the browser of the impression

MACHINE LEARNING TECHNIQUE

- ▶ The key to selecting the meta features is to find the models which give accurate and different results.

Actual \ Preds	TRUE	<u>FALSE</u>
TRUE	782	1218
<u>FALSE</u>	667	7533

Logistic Regression

Actual \ Preds	TRUE	<u>FALSE</u>
TRUE	1753	247
<u>FALSE</u>	2978	6022

Random Forest

Actual \ Preds	TRUE	<u>FALSE</u>
TRUE	1135	865
<u>FALSE</u>	1234	6766

Gradient Boosting

Actual \ Preds	TRUE	<u>FALSE</u>
TRUE	1435	565
<u>FALSE</u>	834	7166

Meta Logistic Regression

Thank You