

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

# Personalized News System using Recommendation Engine

SurajVaidyanathan<sup>1</sup>, Digvijaysingh Chauhan<sup>2</sup>, SairajSaji<sup>3</sup>, Ujwala Ravale<sup>4</sup>

B.E Student, Department of Computer Engineering, SIES Graduate School of Technology, Nerul, Maharashtra, India<sup>1</sup>

B.E Student, Department of Computer Engineering, SIES Graduate School of Technology, Nerul, Maharashtra, India<sup>2</sup>

B.E Student, Department of Computer Engineering, SIES Graduate School of Technology, Nerul, Maharashtra, India<sup>3</sup>

Assistant Professor, Department of Computer Engineering, SIES Graduate School of Technology, Nerul, Maharashtra, India<sup>4</sup>

**ABSTRACT:** This paper presents a system that aggregates news from various electronic news publishers and distributors. The system collects news from HTML and RSS Web documents by using source-specific information extraction programs (wrappers) and parsers, organizes them according to pre-defined news categories and constructs personalized views via a Web-based interface. The Personalized News System (PNS) provides its users with personalized access to news items harvested from multiple Web sources. It takes input from the content sources (news agencies, news portals, electronic newspapers etc.). News items would be categorized according to the user's tastes and the clicking patterns of the User on the various categories of News articles.

**KEYWORDS:** personalized; RSS; apriori; news; recommended.

### I. INTRODUCTION

**Personalization** can be defined as the process of tailoring pages to individual users' characteristics or preferences. Personalization has become the trend of today's life with all major businesses using it to attract customers. Personalization is a key element in social media and recommender systems. Web pages can be personalized based on the characteristics (interests, social category, context, etc.) of an individual, therefore providing them with a tailored user experience. The term customization is often used when the site only uses explicit data such as product ratings or user preferences.

**Recommender system**<sup>[5]</sup>(sometimes replacing "system" with a synonym such as platform or engine) are a subclass of information filtering system that seek to predict the 'rating' or 'preference' that a user would give to an item.

Recommender systems typically produce a list of recommendations in one of two ways - through collaborative or content-based filtering <sup>[8]</sup>. Collaborative filtering approaches building a model from a user's past behaviour (items previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users. This model is then used to predict items (or ratings for items) that the user may have an interest in. Content-based filtering approaches utilize a series of discrete characteristics of an item in order to recommend additional items with similar properties.

There is a vast variety of news in the internet. Searching for something specific takes time. This is where Recommender systems could help. Recommender systems could provide Users with news they like as soon as they log in into the system. Their past history would be stored in the system and it can be used to provide the User with a customized feel. The News would be arranged in an order in which the User has previously viewed them.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

## II. NEED OF PROJECT

Nowadays, Personalization is used in all walks of life. For e-commerce websites it has become necessary to ensure continued visits from Users and thereby increasing their sales. Personalization is also used in social media websites like Facebook and LinkedIn to recommend people and friends on the basis of existing friends. Personalization is done using Recommender systems. Recommender systems help in predicting a User's future actions on the basis of his past actions. In any e-commerce website this feature is very important as what the User may be interested in purchasing may be predicted on the basis of his past purchases. Learning how such a system works would help in knowing how e-commerce websites like Flipkart and Amazon recommend products to Users.

## III. LITERATURE SURVEY

The trend of personalization is increasing day by day. More and more organizations are realizing that personalization is going to be necessary to maintain their customer base. One of the main advantages of personalization is that it helps users to get a customized feel. Dr. M. Durairaj, K. Muthu Kumar stated in their paper News Recommendation Systems Using Web Mining: A Study<sup>[1]</sup> that Online news reading has become very popular as the web provides access to news articles from millions of sources around the world. Yuqi Wang, Wenzian Shang proposed a method to mine consumers' click behaviour from website logs by using association rules in their paper Personalized News Recommendation Based on Consumers' Click Behaviour<sup>[2]</sup>. Jae-wookAhn, Peter Brusilovsky, Jonathan Grady, Daqing He, Sue YeonSyn in their paper Open User Profiles for Adaptive News Systems: Help or Harm?<sup>[3]</sup> spoke about their system called YourNews which observes a user's news-reading behaviour, constructs a user model (profile) representing user interests, and uses this model to recommend the most relevant news articles.

## IV. METHODOLOGY

### A. PROPOSED SYSTEM

The proposed system aims at providing the User with personalized News. This can be achieved by analysing the User's click behaviour. He then views the News articles he is interested in. When multiple Users access the system all of the news articles they have viewed is stored in the database in the form of the feed\_id. All the feed\_ids are then fed to the Apriori algorithm to perform collaborative filtering. With the help of association rules the most frequent news items viewed are found out and then all the Users are then showed these recommended news items. Also the system has a search feature which helps in finding out a particular news article on entering its keyword. These news items are arranged in the order of recent to oldest articles on the basis of their publishing date.

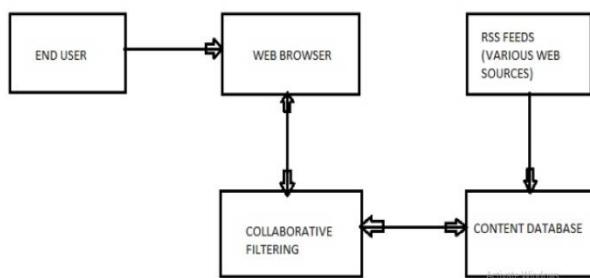


Figure 1: System Architecture

1. RSS Feeds: RSS Feeds would be used to fetch the news from various news portals and they would be stored in the database.

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

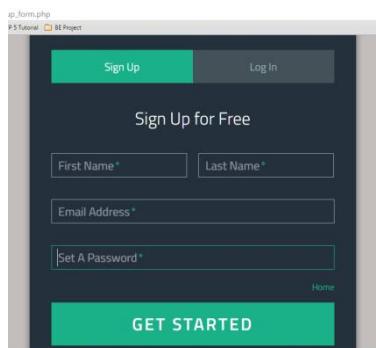
2. Content Database: This is the MySQL database present in Xampp. All the RSS feeds of the news articles are stored in this database
3. Collaborative Filtering: Collaborative Filtering is done in the system with the help of Apriori algorithm
4. Web Browser: The web browser is the front end interface for the User to access the system.
5. End Users: They are the Users that are accessing the system. Once the user has logged into the system he can use the system.

### B. APRIORI ALGORITHM

Apriori<sup>[7]</sup> is a classic algorithm for learning association rules. Apriori is designed to operate on databases containing transactions (for example, collections of items bought by customers, or details of a website frequentation). It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database: this has applications in domains such as market basket analysis. In our project we are using apriori algorithm to find the frequently viewed news articles and recommend these news articles to the registered users. The association rules are used to find out the news items which are the most frequently used.

### V. RESULTS

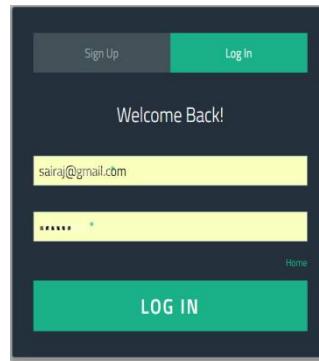
- 1.Signup page – In this page the User fills in his personal details and also sets a password for him to enter the system.



The registration screen is a web form titled 'Sign Up for Free'. It features a 'Sign Up' button at the top left, a 'Log In' button at the top right, and a 'GET STARTED' button at the bottom. The form contains four input fields: 'First Name\*' and 'Last Name\*' in a row, followed by 'Email Address\*' and 'Set A Password\*' below them. There is also a 'Home' link at the bottom right of the form area.

Figure 2: Registration screen

2. Login Page – The User logs into the system entering his Username and Password



The login screen is a web form with a dark background. It features a 'Sign Up' button at the top left, a 'Log In' button at the top right, and a 'LOG IN' button at the bottom. The middle section contains the text 'Welcome Back!' above two input fields. The top field is filled with 'sairaj@gmail.com' and the bottom field is filled with '\*\*\*\*\*'. There is also a 'Home' link at the bottom right of the form area.

Figure 3: Login screen

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

3. Home Page – The initial page the User views on entering the system.



Figure 4: Homepage

4. News Display page – This is the page where the User can view all the news items of his choice. User has an option of a search bar and also of recommended articles in the right side of the page.

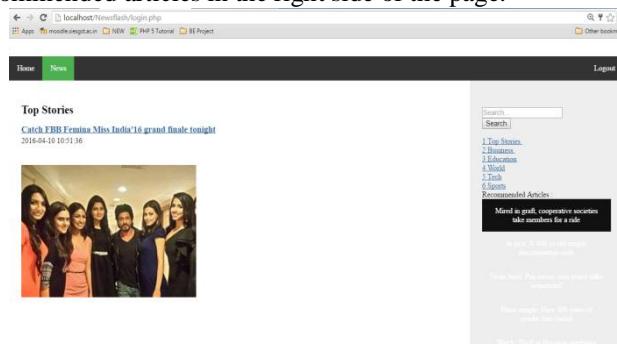


Figure 5: News Display Page

5. Feed table- Each User's click is being recorded with the timestamp to which we further apply Apriori.

phpMyAdmin				
Database: newsflash => Table: feed				
<input type="button" value="Browse"/> <input type="button" value="Structure"/> <input type="button" value="SQL"/> <input type="button" value="Search"/> <input type="button" value="Insert"/> <input type="button" value="Export"/> <input type="button" value="Import"/> <input type="button" value="Privileges"/> <input type="button" value="Operations"/>				
<small>DB: newsflash 0.0.1 8.0.11 - MySQL 5.6.24 - Localhost via TCP</small>				
+	Optimize	feed_id	parent_id	time
1	2	1	0	2016-03-27 18:31:50
2	2	1	0	2016-03-27 18:30:23
2	2	5	100	2016-03-27 18:30:23
2	2	99	2016-03-27 18:30:26	
2	2	5	98	2016-03-27 18:30:32
2	2	5	95	2016-03-27 18:30:45
3	3	72	100	2016-03-27 18:30:31
3	1	73	2016-03-27 18:30:34	
3	1	70	2016-03-27 18:30:38	
3	1	68	2016-03-27 18:30:46	
3	1	68	2016-03-27 18:30:48	
3	1	72	2016-03-27 18:30:57	
3	1	72	2016-03-27 18:37:55	
3	5	100	2016-03-27 18:38:07	
3	5	100	2016-03-27 18:38:12	
3	5	99	2016-03-27 18:38:44	
3	5	97	2016-03-27 18:38:53	
3	5	96	2016-03-27 18:38:59	
3	5	95	2016-03-27 18:39:03	
1	1	72	2016-03-27 18:40:21	
1	1	71	2016-03-27 18:40:24	
1	1	78	2016-03-27 18:40:38	
1	5	100	2016-03-27 18:40:38	
1	5	99	2016-03-27 18:40:44	
1	5	98	2016-03-27 18:40:50	

Figure 6: Feed Table where User's clicks are registered

6. RSS Feeds table – This is where all the news is fetched dynamically with each having a unique id and a category id and is stored in the database to display it in the News page.

# **International Journal of Innovative Research in Science, Engineering and Technology**

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 4, April 2016

Figure 7: RSS Feeds Storage

## VI. ACKNOWLEDGMENT

We would like to express our deepest appreciation to all those who provided us with the help required to complete this report. It gives us immense pleasure to thank Dr.AlkaMahajan, our Principal for extending her support to carry out this project.

A special gratitude given to our Head of Department Dr.RizwanaShaikh and final year project guide Prof. UjwalaRavale, whose contribution in stimulating suggestions and encouragement, helped us to coordinate our project and also to write this report.

Also, we have to appreciate the guidance given by other supervisors as well as the panel during our project presentation that has helped in the overall improvement of our project thanks to their ideas and advices.

## VII. CONCLUSION

Recommender systems can be used for a variety of purposes in today's world. Using recommender systems, the goal here is to create an application that will save people's time and also give a system that will be customized to the user likes. Predicting a User's future actions on the basis of his past actions if what recommender systems do.

Personalization is necessary in today's world in ensuring repeated visits from your customers. News is being read less and less by today's youth. By giving Users what they like from the very beginning this application intends to ensure that people stay updated with what is happening around them. One of the other Benefits to such a personalized system also is that the system has a record of the User's likes and dislikes. This can be used in the future to create a pattern on what items to recommend to a user. Such type of recommendation is used by shopping website to persuade us to buy stuff.

## REFERENCES

- [1] Dr. M. Durairaj, K. Muthu Kumar: "News Recommendation Systems Using Web Mining: A Study". International Journal of Engineering Trends and Technology (IJETT) – Volume 12 Number 6 - Jun 2014
  - [2] Yuqi Wang, Wenzian Shang: "Personalized News Recommendation Based on Consumers' Click Behaviour". 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Vol.1, pp.634-638, 2015
  - [3] Jae-wookAhn, Peter Brusilovsky, Jonathan Grady, Daqing He, Sue YeonSyn" Open User Profiles for Adaptive News Systems: Help or Harm?". WWW 2007 / Track: Browsers and User Interfaces, Vol.1, pp.11-20, 2007
  - [4] Toby Segaran (2007). Programming Collective Intelligence. United States of America: O'Reilly.
  - [5] [https://en.wikipedia.org/wiki/Recommender\\_system](https://en.wikipedia.org/wiki/Recommender_system)
  - [6] [https://en.wikipedia.org/wiki/Collaborative\\_filtering](https://en.wikipedia.org/wiki/Collaborative_filtering)
  - [7] [https://en.wikipedia.org/wiki/Apriori\\_algorithm](https://en.wikipedia.org/wiki/Apriori_algorithm)
  - [8] <http://recommender-systems.org/content-based-filtering/>
  - [9] <http://searchsalesforce.techtarget.com/personalization>
  - [10] <https://econsultancy.com/blog/62881-the-10-essentials-of-social-personalisation/personalizationperof-social-personalisation/>