

## **Project Proposal**

**Team :** 7th

**Project Title :** surprise.me ( Pattern based RSS searching )

### **Problem Description :**

Today although it is easy to get the contents that a user need, it is not always possible to get the exact contents that a user really need. A user may be interested in so many things and he might want to read the articles related to the things that excites him. There are so many RSS readers available over the net such as “Google Reader”, “Feedly”, “LightRead” etc. A user can subscribe to the feeds of the website and using the tools/sites mentioned previously he can regularly visit it and can read the latest contents. This is an easy way to keep yourself updated with the latest contents.

The problem with this approach is that a user may “skip” the articles which may be really helpful to him because of so many articles in the subscription list. Currently there is no such service which gives the user the most related contents over the web. So our project is to provide a user such a platform where he can keep himself updated with the latest contents in such a way that he can see the most relevant and most rated articles on the top of his reading list. The benefit of this approach is that a user might not always have time to read the all the posted articles posted on the website so that in a rush situation he comes to know about some articles that he should not miss.

Furthermore, one another tendency that a user is having is that he only wants to see the information that is relevant to him. He does not want any useless things to be displayed on the website and He always wants to read as less as possible for information introduction. A typical case with today's information era is that there is too much informative information so that a normal user may be skipping so many information even though he wants to read it. So that our project is remedy to this situation where articles that are going to be displayed are arranged in such a way that user has to read as less as possible and also getting the information that he wants in easiest way.

### **New Learning :**

As the project that is going to be developed need a lot more work on the algorithms part, data mining and web crawling along with the social network integration. We are expecting a lot more things to be learned through this project.

The fascinating things which are going to be learned are

- Various data mining techniques and algorithms
- Crawling techniques
- Exposure to various open source tools and technologies
- Developing highly customizable UI
- Web application deployment and maintenance
- Social network integration

### **Reasons for selecting the topic :**

One major reason for selecting this project topic is that in the current information era such an application is really required where user doesn't have to dedicate so much time and get the required contents easily.

Another reason being as the application has a huge scope we can divide our work effort in such a way that every one is doing something interesting work to make the application live. For e.g Web designing, Web service development, crawling algorithms , mining algorithms and techniques etc. There are various parts to be explored in the project development.

### **Knowledge of topic :**

Currently, we are very less knowledgeable on the project that we have selected. Less knowledgeable in the sense that we haven't implemented any web crawling algorithms, mining algorithms, never made any application live and maintained it but as a part of project development process we will be working on it.

### **Preliminary Research Results :**

Project research is the part of the development process and to facilitate that process we have divided our research in the areas of Association Rule Mining, Social Network Mining, Information Retrieval, Web Searching, Crawling, Information Extraction, Classification and Clustering. Research results are yet to be discussed.

For research we have also explored some similar but not exactly same application like [www.stumbleupon.com](http://www.stumbleupon.com), [www.digg.it](http://www.digg.it) etc. From there also we got some of the interesting requirements which can attract a wide user base.

### **Features to be included in the project :**

- User can register on the website freely.
- User can mention his interests and can modify(add/drop) them later. Based on the interests he will be shown contents most related to his interests.
- He can get surprised by the most viewed article of the week.
- User can maintain his “read later” list of the articles.
- User can rate the articles, can manage history ( history will be logged ). Based on the history also the articles will be fetched to show him.
- User can surprise others by giving the articles link to other users. ( social networking integration will be there from which user can surprise other users ).
- User can search articles efficiently from his history as well as from the entire application.
- User can maintain his categories of the interests like travelling , poems etc. He can also provide his source of the information.