



## Filtering Online Social Networks Based on User Content Generation

Joshua Gisemba Okemwa

Research Scholar, Department Of Computer Science, Bhupal Nobles' University, Udaipur,  
Rajasthan, India

DOI: [10.23956/ijarcsse/V7I3/0103](https://doi.org/10.23956/ijarcsse/V7I3/0103)

*Abstract- Social media is part of computer mediated technology that helps an individual to connect with their friends, family and online society, to share information, ideas, career interests and other forms of expressions. Online social networks (OSN) are facing the problem of the people posting indecent messages on individual's wall. Machine learning (ML) intelligence is used to filter these large volumes of data. I aim to propose an automated filter wall (FW), able to filter unwanted messages from (OSN) user walls. To the fact that in OSN's there is the possibility of posting and giving comments on other posts on particular public (or) private walls (PW). Information filtering is mainly used to give user the ability to control the kind of information written on their own walls by filtering out unwanted messages from a chunk of data that a user intends to post. Content Based Filtering will give user ability to select information item based on the correlation between the content of the items and the user preferences. On the other hand Filtering will be mainly used to filter the unwanted messages and if tendency is noticed from the person posting messages, option of blacklisting is available.*

*Keyword - Online social networks, filter wall, private wall, machine learning.*

### I. INTRODUCTION

On-line Social networks helps individuals to connect with their friends, family and the society online in order to share new experiences with others, also recently used for business purpose and statistics analysis of trending stories and reports. But OSNs are facing a major problem of people posting the indecent messages on any individual's wall, and this happens sometimes unknowing because tastes and likes of friends keep changing with attitudes and emotions which is quite pre planned.

Providing a solution to this challenge I review the use of information filtering for sensitive purpose, of removing unwanted message, due to the fact that in OSNs there is the possibility of posting or commenting about other posts on particular public/private areas, called in general walls.

Filtering can therefore be used to give users the ability to automatically control the messages written on their own walls, by filtering out unwanted messages. I believe that this is a key OSN service that if fully implemented can prove to be a major boost to ensure that online users get best of what they are comfortable with. OSN's has not provided enough support to prevent unwanted messages from appearing on user walls. For example Face book allows users to accept those who are to be allowed to post messages in their walls, they include friends, friends of friends, or defined groups of friends.

But no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, that have been used for a very long time, rather it requires to design ad hoc classification strategies.

Because wall messages are constituted by short text for which traditional classification methods i.e. Boosting based, Support Vector machines, Naïve Bayesian have a few limitations. They do not provide sufficient word occurrences, which can be covered if we use an alternative method of data filtering.

In this paper I propose to evaluate an automated system, called Filtered Wall (FW), able to filter unwanted messages from OSN user walls. Machine Learning (ML) is used as text categorization techniques to automatically assign with each short text message a set of categories based on its content.

The overall short text classification strategy is based on Radial Basis Function Networks (RBFN) for their proven capabilities in acting as soft classifiers, in managing noisy data and intrinsically vague classes. Moreover, the speed in performing the learning phase creates the premise for an adequate use in OSN domains, as well as facilitates the experimental evaluation tasks.

We insert the neural model within a hierarchical two level classification strategy. In the first level, the RBFN categorizes short messages as Neutral and Non neutral. Besides classification facilities, the system provides a powerful rule layer exploiting a flexible language to specify Filtering Rules (FRs), by which users can state what contents, should not be displayed on their walls.

User-defined BlackLists (BLs) support will be provided, include lists of users that are temporarily or permanently prevented from posting any kind of messages on a user wall. this proposal of a system has ability to automatically filter unwanted messages from OSN user walls on the basis of both message content and the message creator relationships and characteristics.

The method illustrated here will bring significant changes to online social networks in that short text classification based on Machine Learning is efficient in social networks. These techniques can be applied on large documents and complied data effectively.

## II. SYSTEM OVERVIEW

It covers the prototype of working flow of system developed to illustrate the idea being implemented in this paper. It combines different modules that have been joined together to show its workability. Input is data generated by different online social media users and the result generated is ultimately filtered basing on the content inputted into the system.

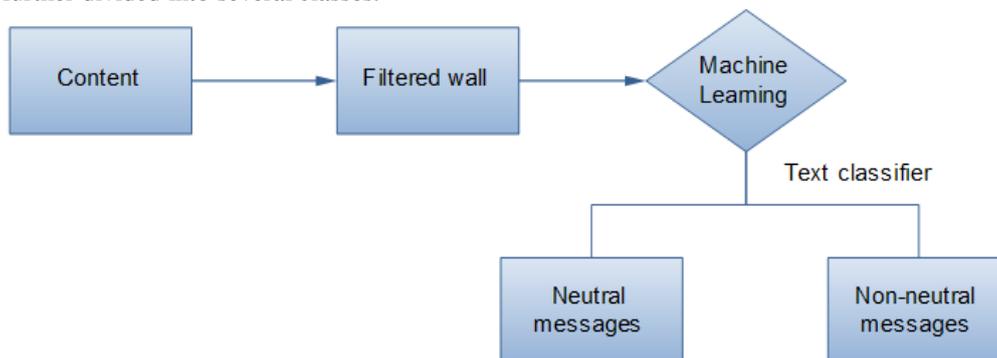
## III. DESCRIPTION OF MODULES

### User Registration and Posting Messages on User Wall.

In most of the networks it's necessary for Users of the social network give their profile details to the network when they are opening account. It is essential because of many established reasons. Information issued includes name, gender, profession, mobile number etc. After registration process user connect with their friends by friend request. With successful feedback popping up, from the backend server, it service as a first step for user to get connected to large network of online users.

### Content Based Message Filtering in Filtered Wall.

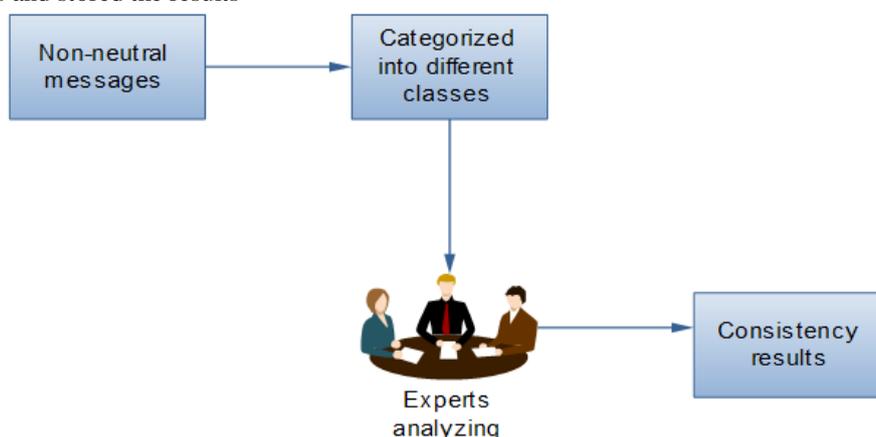
Filtered wall is designed as to filtering unwanted messages based on content received from the user. Machine Learning is the text classifier which is used to classify the messages and short texts posted by users. Documents processed in content-based filtering are mostly textual in nature and this enables content-based filtering to be close text classification. Machine learning technique classifies the message into neutral and non-neutral. After that non-neutral messages are further divided into several classes.



### Analyzing messages by experts.

Results of the Machine Learning technique are analyzed; the messages are evaluated by experts. Experts will evaluate the non-neutral messages under Violence, politics, offensive, hate, sexual harassment categories.

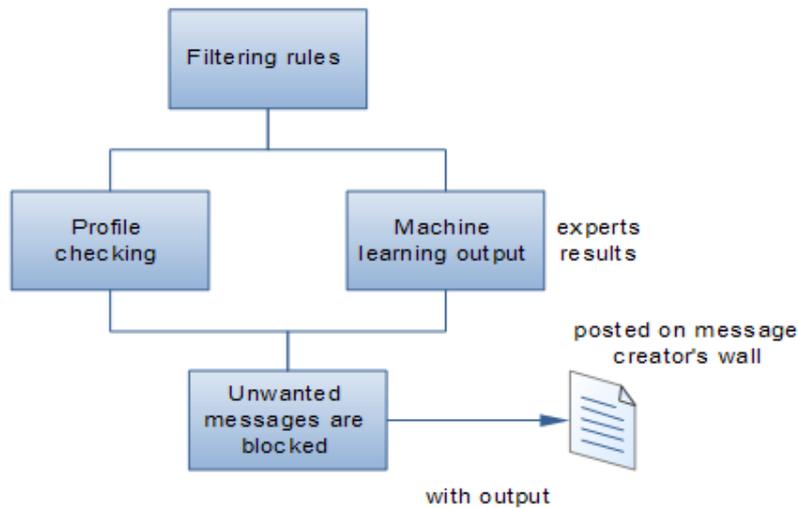
The issues regarding the consistency between the opinions of experts are considered. Bag of Words, contextual features yield good performance in text categorization. We consider the experts overall accuracy and consistency in analyzing messages and stored the results



**Applying Filtering Rules and prevent unwanted messages.**

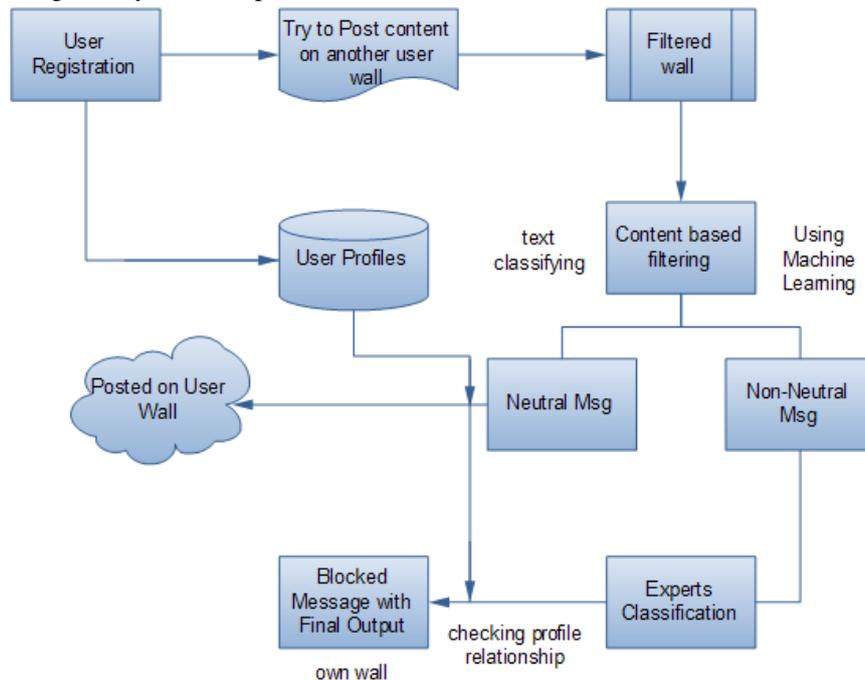
User’s details are stored in their profile database. Before posting the message filtering rules check the relation between the message creator and receiver and also the profile creation date etc. Some users are temporary or permanently blacklist because of their activities.

With the Machine learning output and experts analyzing results regarding to a user’s unwanted message, the message will be blocked and display the results in the message creator wall. Systems with the content based message filtering by Machine Learning and profile checking process we prevent unwanted messages in a social network.



**IV. DATA FLOW DIAGRAM**

The diagram below shows the overall flow of data through various modules that have been combined together to form a prototype showing how system is expected to work.



**V. CONCLUSION**

The system presented in this paper helps to filter undesired messages from OSN walls. The system exploits a machine learning soft classifier to enforce customizable content-dependent filter walls. Output from this system is content that as been filtered as per users preferences and choices

Flexibility of the system in terms of filtering options is enhanced through the management of block lists. This procedure promotes total locking of users who have tendency of writing vulgar language on the walls of users.

Improving the quality of classification of words will in a big way boost the speed of content classification thus ensuring faster filtering to cover the high number of users online. With recent emerging tools of classifying data its easy to make it possible to filter all the content no matter how large it is.

Advanced security can be introduced for further authentication before being allowed to make any comment which will further be subjected through filtering process. This further reduces the amount of data that will be available for classification and filtering.

**REFERENCES**

- [1] M. Chau and H. Chen, "A Machine Learning Approach to Web Page Filtering Using Content and Structure Analysis," *Decision Support Systems*, vol. 44, no. 2, pp. 482-494, 2008.
- [2] R.J. Mooney and L. Roy, "Content-Based Book Recommending Using Learning for Text Categorization," *Proc. Fifth ACM Conf. Digital Libraries*, pp. 195-204, 2000.
- [3] R.E. Schapire and Y. Singer, "Boostexter: A Boosting-Based System for Text Categorization," *Machine Learning*, vol. 39, nos. 2/3, pp. 135-168, 2000.
- [4] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," *Proc. European Conf. Machine Learning*, pp. 137-142, 1998.
- [5] M. Carullo, E. Binaghi, and I. Gallo, "An Online Document Clustering Technique for Short Web Contents," *Pattern Recognition Letters*, vol. 30, pp. 870-876, July 2009.