DOI: https://doi.org/10.31341/jios.42.1.1 UDC 004.77:316.77
Open Access Original Scientific Paper

Emotion-Based Content Personalization in Social Networks

Anthoniraj Amalanathan *aanthoniraj@vit.ac.in*
School of Computer Science and Engineering
Vellore Institute of Technology, Vellore, India

S Margret Anouncia *smargretanouncia@vit.ac.in*
School of Computer Science and Engineering
Vellore Institute of Technology, Vellore, India

Uffe Kock Wiil *ukwiil@mmmi.sdu.dk*
The Maersk Mc-Kinney Moller Institute
University of Southern Denmark, Odense, Denmark

Abstract

Personalization is the process of customizing social network pages of users according to their needs and personal interests. It can also be used for filtering unwanted information from an individual's page received from other users, in case this information is unpleasant or unacceptable. To avoid unwanted information from a particular user in current social networks, the user needs to be denied accessibility by blocking them. However, instead of blocking the user, it would be preferable to have a mechanism that prevents the undesirable content in a user's social network page. Thus, this paper presents a model that determine the emotions shared in the content of a social network page by the user. The model determines the dominant emotions for a period of time and uses these to filter the content using the user's dominant emotions. Using the developed model, a novel system based on item based collaborative filtering process to personalize the user's social network page has been developed. A user study involving 5000 Twitter messages shows that the developed system performs satisfactory with a correctness in the filtering process of 87%.

Keywords: Social Network; Emotion; Personalization; Collaborative Filtering.

1. Introduction

Generally, emotions are a complex combination of external and internal influences that reflect the feelings and moods of a human being. It forms an important component of one's behavior [1, 2]. It can emerge in any social context, but in social network setting, it has the privilege of spreading rapidly as numerous users are connected simultaneously [3].

Though the cyberspace does not provide emotions directly, it allows individuals to interact thereby sharing their emotions. Due to the revolution of social networking, in the present era, vast contents are continuously being shared. Often, the shared contents may not confirm to the likings of the users. At the same time, this may influence any individual into deviating from their basic nature, spoiling the ambiance, and negatively affecting the entire community. Emotions have great impact on on-line communication, especially in social networks. To maintain the harmony in an established relationship on the social network it is always good to share and retain the desirable content and filter the undesirable ones. This can be achieved by dynamically sieving the contents shown in the user's page based on their dominant emotions. Accordingly, a methodology was designed to filter the undesirable content of user's social network pages.

The paper is organized as follows. Section 2 relates the existing problems to the relevant literature. Section 3 demonstrates the proposed model and analyzing standard algorithms to classify user’s content for personalizing social network page. Section 4 analyses the process of personalization and section 5 deals with the concluding remarks.

2. Related Work

Several researchers have provided various dimensions of handling emotions and the extraction of the same from shared contents. The subsequent sections provide a few ideas proposed for analyzing three major facts: human emotions, emotions and social network, and personalization.

2.1. Human Emotions

Emotion is a joint function of a physiologically arousing situation [4] may be divided into positive and negative [5]. It can differ from men and women and be revealed from sound, color, smell, and taste [6]. Emotions can be derived from performing music, dance, and poetry where it is expressed in different nine forms viz. Love, Joy, Wonder, Courage, Calmness, Anger, Sadness, Fear, and Disgust [2, 7].

2.2. Emotions and Social Network

The view of emotion as recognition of non-verbal vocalization such as screams and laughs highlighted the cross cultural recognition of emotions [8]. With the advent of social networks in early 2000, emotions were considered to be a key factor for establishing contacts and relationship between the social network users [9]. Several attempts have been carried out by researchers to correlate emotions with the content being shared through social networks. More specifically, this exploration has been carried out with respect to interaction performed in Facebook and content shared in Twitter. Few researchers have considered other social networks for the scrutiny.

The following sections highlight some investigations focusing on associating emotions and social networks especially in Facebook and Twitter. Researchers have put in a lot of effort in realizing the dispersion of emotion through the content shared in social networks [8, 9, 10, 11, 12, 13,14, 15] Also, it is noted that most of the emotions are exchanged through Facebook and Twitter activities.

As a popular social network, Facebook provides a good way of emotion sharing through different activities and contents [12, 13, 14, 16, 17, 18]. According to the researchers, emotion is dispersed in Facebook in the form of topics, messages, wall content, size, and density of the connectivity of individuals. It is understood from the study that the users with a dense connectivity tend to disperse more positive emotions such as happiness and surprise.

Similarly, Twitter has been serving as a good emotion diffuser through tweets [8, 14, 19, 20, 21]. According to these researchers, content, topic, messages, type of conversation, social ties, amount of network activity of individuals, participatory patterns, and language use act as emotion carriers in Twitter.

In addition to Facebook and Twitter, on-line communities such as product review communities, chat rooms [22], and other social tagging systems like Flickr [23] have also been analyzed for emotion influence.

Kim (2012) proposed a computational framework for analyzing the emotions in micro blogging like Twitter users' conversations. The analysis was done to a maximum of eight emotions (anticipation, joy, anger, surprise, fear, sadness, disgust, and acceptance) and a set of expanded seed words in Twitter conversations [19].

Hence, there arises a scope of analysis for the rest of the emotions. Ultimately, the study of the significance of emotions in social networking leads to appropriate and careful decision making in strategic formulation to attract other users to one's content. To keep the message short, content may be in any one of the form such as text or emoticons. Hence, analyzing different emoticons with respect to information sharing in a social network may also be helpful to understand the real intention or true meaning of the content.

The categorization would necessitate the mining process of emotions using text content and emoticons in social networks [24, 25]. The results of the classification process may further be used for personalization of user content that would avoid unnecessary information in social network page of a user.

2.3. Personalization in Social Network

Social networks site Facebook had been a source of emotional content diffuser through different activities [12, 13, 14, 15, 16, 17, 18]. The ultimate objective is to examine the power of emotions in social network, to analyze the user's behavior in social network. This process of behavior analysis would be helpful in the spectrum of application starting from organization knowledge building [14], social marketing, [11, 26, 27] to personalization.

Though a variety of applications were approached for social network user behavior analysis [28], a major challenge was met in personalization of shared user content to avoid undesired content in the social network pages. To start with the personalization process, it is required to perform user content classification based on the type of emotions expressed in the content. After tremendous growth of social network activities from 2008, vigorous researches have been carried out for this purpose. Typical classification process was carried out for both content and user in the social networks.

According to research, the following algorithms as stated in Table 1 have been identified as standard methodologies to perform mining and thereby handling the classification of content. Specifically, for performance of social network content classification process, a few of the mining techniques are considered to be effective.

Algorithm	Description
Multi-nominal Naive Bayes	It chooses the most likely label for an input, under the assumption that every input value is generated by first choosing a class label for that input value, and then generating each feature, entirely independent of every other feature [29].
Support Vector Machine	Given labelled training data (supervised learning), the algorithm produces a finest hyper plane which classifies new samples [30].
Decision Trees	This classifier works by creating a tree structure, where each node resembles to a feature name, and the branches resembles to a feature values. Tracing down the branches, you get to the leaves of the tree, which are the classification labels [31, 32].
Logistic Regression	Predicts the probability of occurrence of an event by fitting data to a logistic function [33, 34].
Perceptron (Neural Network)	A classification algorithm that makes its predictions based on a linear predictor function combining a set of weights with the feature vector [35].

Table 1. Standard Text Classification Algorithms

Hence, an attempt is made to design and develop a model for personalization using emotion-based user content classification in social network.

3. Methodology to Personalize Social Network

Considering the downsides of spreading the emotions through social network content, a personalization model is created to filter the unwanted content being received by a user as illustrated in Figure 1.

To demonstrate effectiveness of the method, Twitter contents were taken into consideration and personalization of Twitter page of the user is endeavored.

The process is approached by implementing three subtasks namely, extraction of tweets through Twitter API, classification of tweets according to the expressed emotion, and finally the content filtering process to filter the unpleasant content.

3.1. Twitter API and Tweets Extraction

The main purpose of API is to define a set of functionalities that are independent of their respective implementation. Usually, an API is delivered in the form of a library, including specifications for routines, data structures, object classes, and variables. In some cases, notably for Simple Object Access Protocol (SOAP) and REpresentational State Transfer (REST) services, an API comes as a specification of remote calls exposed to the API users. As far as Twitter is concerned, it is built up on REST architecture, referring to a collection of network design principles that define resources and methods to address and access data.

The following tasks applied to extract the tweets involves,

- Construction of URL
- API Key Access
- Rate Limitation

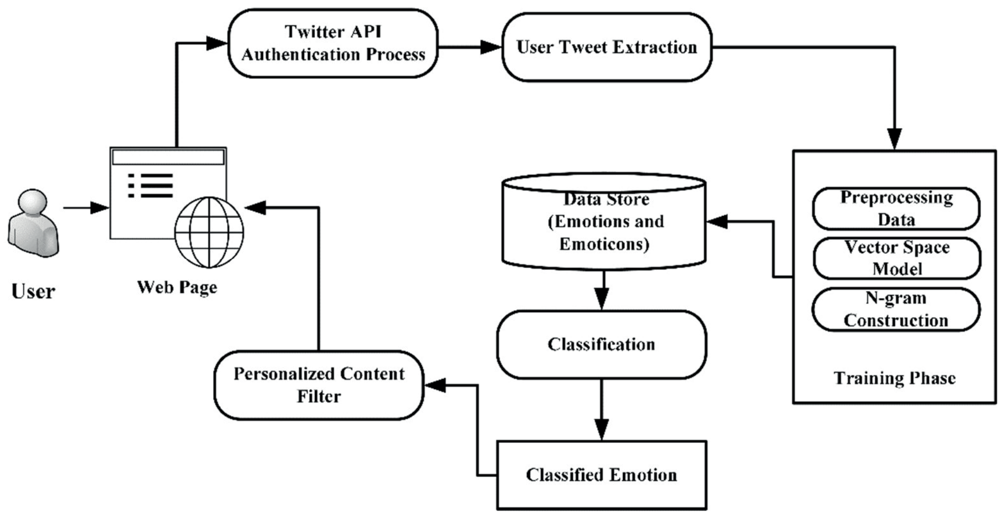

Figure 1. Personalization model

3.1.1. Construction of URLs

The usual access into Tweets happened as an unauthorized user which could extract only minimum tweets for analysis. Therefore, the URL needed to be reconstructed using permitted key to access the Tweets as an authorized user.

3.1.2. API Key Access

In order to access the tweets with proper authorization, the API key provided by Twitter for the registered user needed to be obtained. The API key was further applied to reconstruct the URL and hence authorized access to tweets was established.

3.1.3. Rate Limitation

Due to the heavy usage of Twitter API, the access to twitter content was limited. For unauthorized users, it was permitted to access 150 tweets per day. On the other hand, the API accepts 15 requests per minute for authorized users.

Thus, using the Twitter API, Tweets are extracted to carry forward further analysis. To continue with the process on the shared contents, a supervised learning approach was considered for which a data store was required to be built. Hence, around 5000 Tweets were extracted and processed to form the data store.

3.2. Data Store Formation

Initially, the training set was constructed by extracting approximately 5,000 labeled tweets. To generate a training set, the social network is represented as a set of building blocks carrying out activities such as identity creation, sharing, conversation, relationship establishment, presence indication, group participation and reputation. These activities tend to share content in form of text, images and video. The content being shared among the users reflect the different forms of emotions that the user possesses at the time of information sharing. According to the design, initially a mapping is performed between the seven social network activities and nine emotions to categorize the type of activities followed for expressing different emotions. This mapping process provided an idea about the nine different emotions which is later applied in segregating the emoticons extracted from the standard library. Such segregated emoticons are stored in data store. In addition, from the collected labeled tweets, meaningful words were extracted manually and stored as a collection of words in a data store. Along with the meaning of full words, emoticons, if present, were extracted and stored in the data store Prior to the creation of the collection of words, the manually extracted meaning of full words were subjected to pre-processing that involved the following subtasks:

- Cleaning tweets
- Removal of stop words
- Tagging of parts of speech
- Vector space model construction
- N-gram model construction

3.2.1. Cleaning of tweets

The data cleaning process of tweets was intended to remove meaningless words and involved the following subtasks:

- URL removal
- Decoding of Unicode
- Apostrophe lookup
- Slang replacement
- Removal of special characters

3.2.2. Removal of stop words

This process was implemented to remove irrelevant words from the tweets. It was performed by language stemming, which is an information retrieval process to reduce inflected words to their roots or derive the roots of words to return a complete word.

3.2.3. Tagging of parts of speech

The tagging of parts of speech involves extracting a part-of-speech-based Penn Treebank corpus, and, hence, adjectives, adverbs, nouns, and verbs are segregated.

3.2.4. Vector space model construction

A vector space model or vector model is usually created to represent any text document as vectors containing identifiers such as key terms or index terms. In this case, a set of words from tweets form a feature set. It represents each tweet as a vector of features $f = (f_1,\ f_2, f_3, \ldots, f_n)\quad where\ f_i : (1 <\ i < n)$ and is a nonnegative value denoting the single or multiple occurrences of term i in tweet d. Thus, each unique term in the tweet collection corresponds to a dimension in the space.

To further improve the accuracy of text-based emotion detection, the constructed vector model was enhanced to form a collection of words containing text expressing emotions and associated semantically similar words based on WordNet [36]. The collection of words was strengthened by adding semantically similar words from the WordNet lexicon database. The process was performed by finding the similarity between text expressing emotion and the words present in the lexicon database. A cosine similarity measure, as indicated below, was applied to extract all words expressing the same type of emotions [37]. The cosine of two vectors $(w_i,\ w_j)$ can be derived by using the Euclidean dot product formula:

$$w_i, w_j = \|w_i\|\ \|w_j\|\ \cos\theta$$

Given two vectors of attributes, w_i and w_j the cosine similarity, $\cos\theta$ is represented using a dot product and magnitude as

$$similarity(w_i,\ w_j) = \cos\theta = \frac{(w_i,\ w_j)}{\|w_i\|\ \|w_j\|} = \frac{\sum_{i=1}^{n} w_i * w_j}{\sqrt{\sum_{i=1}^{n}(w_i)^2} * \sqrt{\sum_{j=1}^{n}(w_j)^2}}$$

Thus, a semantic container was constructed using a vector space model, which was then subjected to n-gram model construction.

3.2.5. N-gram model

An n-gram model is a type of probabilistic model used to predict the next item in a sequence in the form of an $(n-1)$ order [38]. It uses the previous $N-1$ words in the sequence to predict the next word and forms a sequence of n items that can be syllables, letters, words, or base pairs. The model uses the Markov assumption to compute Maximum Likelihood Estimates (MLE) for the last word in the sequence by considering the previous word (2-gram), the previous two words (3-gram), or a maximum of the previous three words (4-gram).

For a given word sequence of n words, $W = \{w_1, w_2, \dots, w_n\}$, the language model probability is rewritten as:

$$p(W) = p(w_1, w_2, \dots, w_n) = \prod_{i=1}^{n} p(w_i | w_0, \dots, w_{(n-1)})$$

where w_0 is chosen appropriately to handle the initial condition.

In this attempt, the tweets that were extracted and stored as a collection of words were tokenized into 2-gram, 3-gram, and 4-gram tokens to build an n-gram model. The initial set of MLEs was calculated for the n-grams and later optimized via interpolation of the dataset. The interpolated MLEs were assigned to all n-grams. The procedure developed for the same is detailed in Algorithm 1.

```
Data: Tweets
Result: freature_labels
slang_dict = {.......};
remove_tags = {.......};
labelled_tweets = [....];
emotion_words = [......];
for tweet in labelled_tweets do
        tokens = tweet.split(' ');
        if word.beginswith(remove_tags) then
                tokens.remove(word)
        end
        if word.beginswith(remove tags) then
                word=slang_dic[word]
        end
        tagged = POS_TAG(tokens)
end
for tag in tagged do
        tag=similarity(tag,emotion_words);
        tweet_list.add(ngram(tag,3),label);
        ngramlist.add(ngram(tag,3));
        ngramlist = set(ngramlist) ;
end
for tweet in tweet_list do
        feature={};
        for f in ngramlist do
                if f in tweet then
                        feature[f] = 1
                end
                feature_labels.add(feature,label)
        end
end
```

Algorithm 1. Generating n-gram model and feature labels

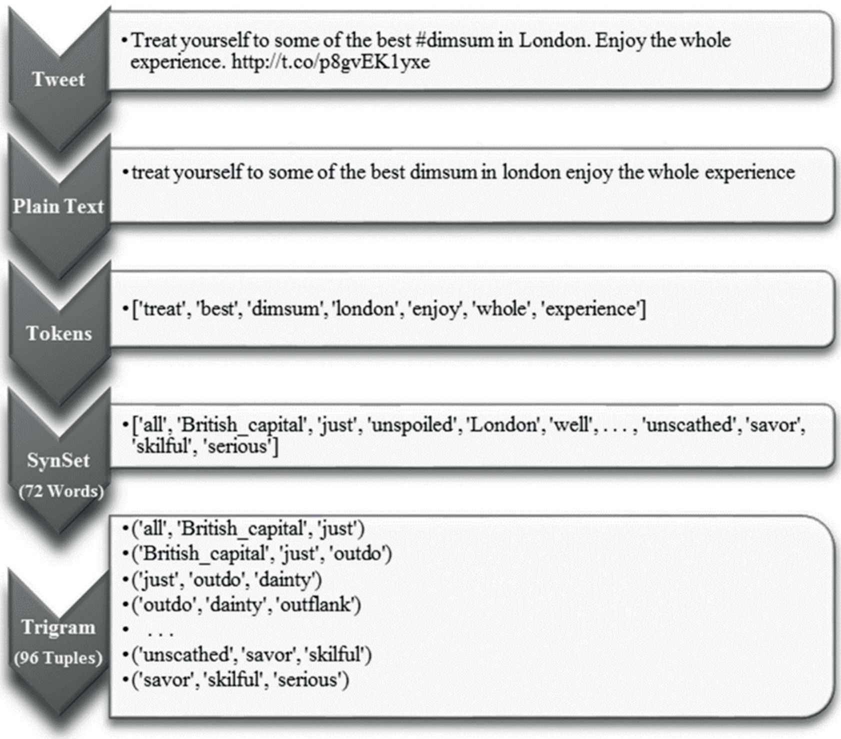

Figure 2. Construction of n-gram model for classifier

The constructed n-grams were also stored in the data store for supporting supervised learning. Thus, the data store was developed to contain classified emotion expressing words for all nine emotions along with their emoticons and all possible n-grams from the tweets. The entire data set was considered the training data for the further process.

Along with the n-gram vector of words, emoticons that are extracted from standard Unicode consortium are also stored as reference data for further processing.

4. User Content Classification

Classification is a process of dividing the data into different classes according to the probability of constraints occurrence. The process may be approached either as a supervised or unsupervised method. The supervised approach requires a training data set with which the classification is guided while unsupervised approach deals with unlabeled classification.

To select an appropriate method of social network user content classification based on emotions, a study of existing popular classification methods is performed. Five standard procedures of classification: Multi-nominal Naive Bayes, SVM, Decision Tree, Regression, and Perceptron are considered for the study. The effectiveness of the chosen method is tested on Tweets and is evaluated using the measures precision, recall, and F1-Score.

By comparing the performance of all the classifier it is found that the classifiers decision tree and SVM performs the best for the selected social network contents.

	Emotions	Love	Happiness	Pity	Furiousness	Heroism	Fearfulness	Disgust	Wonder	Peace	Average
Naive Bayes	P	0.81	0.50	1.00	0.70	1.00	1.00	1.00	1.00	1.00	**0.80**
	R	0.57	0.99	0.73	0.80	0.38	0.67	0.38	0.09	0.72	**0.69**
	FM	0.67	0.66	0.85	0.75	0.56	0.80	0.55	0.17	0.85	**0.69**
SVM	P	0.89	0.88	0.88	0.72	1.00	1.00	1.00	1.00	1.00	**0.90**
	R	0.68	0.91	0.93	0.94	0.85	0.94	0.94	0.91	0.93	**0.89**
	FM	0.77	0.90	0.90	0.82	0.92	0.97	0.97	0.95	0.96	**0.89**
Decision Trees	P	0.69	0.94	1.00	0.96	1.00	1.00	0.94	1.00	1.00	**0.94**
	R	0.95	0.94	0.93	0.88	0.85	0.88	1.00	0.91	0.93	**0.92**
	FM	0.80	0.94	0.97	0.92	0.92	0.92	0.97	0.95	0.96	**0.93**
Regressi-on	P	0.86	0.57	1.00	0.84	1.00	1.00	1.00	1.00	1.00	**0.84**
	R	0.65	1.00	0.73	0.82	0.54	0.78	0.62	0.36	0.72	**0.78**
	FM	0.74	0.72	0.85	0.83	0.70	0.88	0.77	0.53	0.84	**0.78**
Percept-ron	P	1.00	0.93	1.00	0.65	1.00	0.88	1.00	0.86	0.59	**0.78**
	R	0.32	0.40	0.60	0.52	0.77	0.78	0.62	0.55	0.66	**0.58**
	FM	0.49	0.56	0.75	0.58	0.87	0.82	0.77	0.67	0.62	**0.60**

Table 2. Performance of classifiers (P –Precision, R-Recall, FM-F1 Measure)

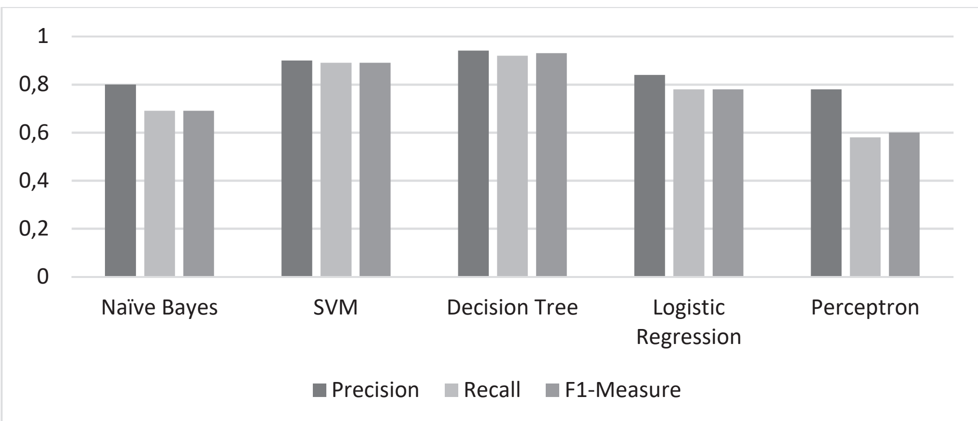

Figure 3. Average of precision and recall for all Classifiers

The performance of classifiers through average of precision and recall based on human emotions in expressed in social networks are portrayed in Table 2 and Figure 3.

The results obtained from decision tree approaches is found to be most suitable for tweet classification based on the users' emotions.

5. Personalization Process

The ultimate objective to develop the unified framework is to perform personalization based on the user's dominant emotion that creates an expectation of desirable content from the social network. Generally, the personalized system is a type of information filtering application which is eventually built by representing the user's information needs and preferences based on the knowledge inferred from their log, ratings, or any other interactions performed by the users. One of the popular method applied for such filtering process is user based collaborative filtering, which is elaborated in the next sub section.

5.1. User Based Collaborative Filtering

Collaborative filtering referred to as social filtering is used to filter information for any recommender system. It processes certain characteristics with which the users have agreed upon and are likely to continue the same in the future. The method spins around three major factors such as: preferences, interests, and interests stability [39], [40]. Two different approaches are possible for user based collaborative filtering namely, user based nearest neighbor and item based nearest neighbor. In the user based nearest neighbor, a prediction on similar users based on certain characteristics is performed while in item based approach similarities between the characteristics of an individual are predicted. Hence, in the process of personalization of social network page, the item based approach is extended to synchronize with the nine emotions expressed by an individual through social network content over a period of time. Considering an individual's tweets over a month, an analysis is performed to find an expressed emotion through the contents in Twitter. The tweet count for a month, per week for each emotion is mined from the tweets posted by the individual as shown in the graph in Figure 4.

From the obtained data, it is understood that the user frequently updates contents related to happiness and heroism than content reflecting any other emotion. Thus, the analysis discovered dominant emotion of the user to be happiness and heroism for that period. Similarly, a correlation between the expressed emotion and the period is drawn to obtain the status of mood change in an individual during the mentioned period. This correlation is necessary to determine the frequency of personalization process to be carried out on social network page of an individual. On analyzing the user's dominant emotion, for the chosen user in the selected period, it was found that the happiness and love appeared to be dominant. Thus, the Twitter page is personalized according to love and happiness.

5.2. Personalization of User's Social Network Page

Personalization is the process of delivering content tailored to a particular user or segment of users. Computer systems that perform personalization on the user needs require recognizable patterns in the user's behavior. One of the method by which user behavior was ascertained is by means of the emotions expressed by them. In social networking, the contents that are shared among the users carry the emotion. Hence, the content is personalized based on the dominant emotion of users.

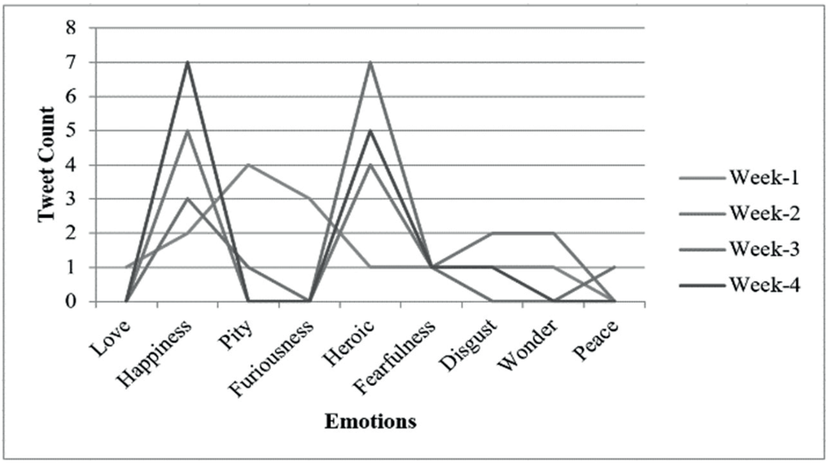

Figure 4. Individual's Emotion Expressed through Tweets

User (U)	TR	TF	NT	TM	PC	PN	PM
U1	64	56	4	4	87.5	6.3	6.3
U2	205	187	12	6	91.2	5.9	2.9
U3	175	150	18	9	85.7	10.3	5.1
U4	316	298	12	6	94.3	3.8	1.9
U5	108	85	11	12	78.7	10.2	11.1
U6	97	84	4	3	86.6	4.1	3.1
U7	68	57	6	3	83.8	8.8	4.4
U8	76	68	4	4	89.5	5.3	5.3
U9	123	94	12	17	76.4	9.8	13.8
U10	84	78	4	2	92.9	4.8	2.4

Table 3. Statistics of the filtered Tweets

(TR - Number of Tweets Received, TF- Number of Tweets Filtered, NT-Number of Neutral Tweets, TM-Number of Tweets Misclassified, PC-Percentage of Correctness, PN-Percentage of Neutral, PM-Percentage of Misclassification)

Finally, the developed model is evaluated using Twitter page of few users. On analyzing the user's dominant emotion for about 2 weeks, of the chosen 10 users, it was found that the happiness and heroism appeared to be dominant for 6 users, peace and love dominated in 2 users while 2 other users were dominant with heroism and furiousness. By applying the personalization model, the statistics of data filtered are tabulated in Table 3.

According to the obtained results, for the considered 10 chosen users, the model could provide an average of 86.7% of correctness in filtering process. While average of neutral tweets counted to 6.9% and the average wrongly classified tweets showed up to 5.6%. Thus, the impact of developed model in the personalization process appeared to be satisfactory and acceptable.

6. Conclusion

The primary contribution of this paper is the development of a novel model and system to support emotion-based personalization of content displayed in Twitter. The developed model aids in the process of personalization of social network page of a user. The developed model is evaluated by applying filtering process on the Twitter page of selected users. The model provided a satisfactory outcome of around 87%. However, the model is also yielded a notable amount of neutral and misclassification rate. The percentage of correctness and the misclassification rate might be positive and negative inflation, as the dominant emotions of the users are determined periodically. Also, the classification depends upon the dictionary created. So, the increase in knowledge size would be reflected in performance of the model and subsequently the results may be varied. To further improve the performance, the model may be extended by grouping similar user's emotions to decide the dominant emotions expressed with respect to time stamp. This factor might help in designing a common template for each emotion personalization and hence may be articulated for the entire social network users. Consequently, the process would considerably increase the harmony of social network usage by reducing the unpleasant content from each individual's perspective. The personalization approach demonstrated in this paper focuses on text information (in particular Tweets). Hence, it is expected that the method can be generalized to be used in other social networks that contain text information.

References

[1] C. E. Izard, *Human emotions*. Springer Science & Business Media, 2013.

[2] A. Adelman, "RASA: Serenity, Valor, Love," University Of North Texas, 2002.

[3] D. M. Boyd and N. B. Ellison, "Social network sites: definition, history, and scholarship," *J. Comput. Commun.*, vol. 13, no. 1, pp. 210–230, 2007.

[4] C. E. Izard, *The psychology of emotions*. Springer Science & Business Media, 1991.

[5] P. Ekman, "An argument for basic emotions," *Cogn. Emot.*, vol. 6, no. 3–4, pp. 169–200, 1992.

[6] V. Johnston, “The origin and function of pleasure,” *Cogn. Emot.*, vol. 17, no. 2, pp. 167–179, 2003.

[7] P. Marchand, “The yoga of the nine emotions,” *Tantric Pract. Rasa Sadhana based teachings Harish Johari*, 2006.

[8] M. De Choudhury, S. Counts, and M. Gamon, “Not All Moods are Created Equal! Exploring Human Emotional States in Social Media.,” in *ICWSM*, 2012.

[9] A. L. Hill, D. G. Rand, M. a Nowak, and N. a Christakis, “Emotions as infectious diseases in a large social network: the SISa model.,” *Proc. Biol. Sci.*, vol. 277, no. July, pp. 3827–3835, 2010.

[10] J. Arafat, M. A. Habib, and R. Hossain, “Analyzing Public Emotion and Predicting Stock Market Using Social Media,” no. 9, pp. 265–275, 2013.

[11] J. Berger and K. L. Milkman, “Social Transmission , Emotion , and the Virality of Online Content,” no. 10, 2010.

[12] K. J. P. Espinosa and A. M. J. Bernales, “Characterizing Influence Factors Affecting Emotion Diffusion in Facebook,” vol. II, pp. 22–24, 2014.

[13] J. Tang, J. Sun, C. Wang, and Z. Yang, “Social influence analysis in large-scale networks,” in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 807–816.

[14] H. Jalonen, “Social Media – An Arena for Venting Negative Emotions,” no. April, pp. 224–230, 2014.

[15] M. D. C. Scott and M. Gamon, “Not All Moods are Created Equal! Exploring Human Emotional States in Social Media,” in *The 6th International AAAI Conference on Weblogs and Social Media*, 2012, pp. 66–73.

[16] R. Fan, J. Zhao, Y. Chen, and K. Xu, “Anger is more influential than joy: sentiment correlation in Weibo,” *PLoS One*, vol. 9, no. 10, p. e11084, Oct. 2014.

[17] F. Kivran-Swaine and M. Naaman, “Network properties and social sharing of emotions in social awareness streams,” in *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 2011, pp. 379–382.

[18] H. Lin and L. Qiu, “Sharing emotion on Facebook: network size, density, and individual motivation,” in *CHI’12 Extended Abstracts on Human Factors in Computing Systems*, 2012, pp. 2573–2578.

[19] S. Kim, J. Bak, and A. H. Oh, “Do You Feel What I Feel? Social Aspects of Emotions in Twitter Conversations,” in *Sixth International AAAI Conference on Weblogs and Social Media*, 2012, pp. 495–498.

[20] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, “User-level

sentiment analysis incorporating social networks," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2011, pp. 1397–1405.

[21] S. Tchokni, D. O. Séaghdha, and D. Quercia, "Emoticons and Phrases: Status Symbols in Social Media," in *Eighth International AAAI Conference on Weblogs and Social Media*, 2014, pp. 1–10.

[22] D. G. Becerra, "Modeling collective emotions in online communities," 2012.

[23] A. Anagnostopoulos, R. Kumar, and M. Mahdian, "Influence and correlation in social networks," in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2008, pp. 7–15.

[24] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining.," in *LREC*, 2010, vol. 10, pp. 1320–1326.

[25] J. B. Walther and K. P. DAddario, "The impacts of emoticons on message interpretation in computer-mediated communication," *Soc. Sci. Comput. Rev.*, vol. 19, no. 3, pp. 324–347, 2001.

[26] E. Cambria, M. Grassi, A. Hussain, and C. Havasi, "Sentic Computing for social media marketing," *Multimed. Tools Appl.*, vol. 59, no. 2, pp. 557–577, May 2011.

[27] P. S. Hudson, M. S. Roth, and T. J. Madden, "Customer Communications Management in the New Digital Era," *Darla Moore Sch. Business, Univ. South Carolina*, Jan. 2012.

[28] Q. Yan, L. Wu, and L. Zheng, "Social network based microblog user behavior analysis," *Phys. A Stat. Mech. Its Appl.*, vol. 392, no. 7, pp. 1712–1723, 2013.

[29] C. D. Manning, P. Raghavan, H. Schütze, and others, *Introduction to information retrieval*, vol. 1. Cambridge university press Cambridge, 2008.

[30] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp. 144–152.

[31] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. CRC press, 1984.

[32] T. Hastie, R. Tibshirani, and J. Friedman, *Unsupervised learning*. Springer, 2009.

[33] H.-F. Yu, F.-L. Huang, and C.-J. Lin, "Dual coordinate descent methods for logistic regression and maximum entropy models," *Mach. Learn.*, vol. 85, no. 1–2, pp. 41–75, 2011.

[34] K. Nigam, J. Lafferty, and A. McCallum, "Using maximum entropy for text classification," in *IJCAI-99 workshop on machine learning for information filtering*, 1999, vol. 1, pp. 61–67.

[35] Y. Freund and R. E. Schapire, "Large margin classification using the perceptron algorithm," *Mach. Learn.*, vol. 37, no. 3, pp. 277–296, 1999.

[36] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to wordnet: An on-line lexical database*," *Int. J. Lexicogr.*, vol. 3, no. 4, pp. 235–244, 1990.

[37] P. Resnik and others, "Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language," *J. Artif. Intell. Res.(JAIR)*, vol. 11, pp. 95–130, 1999.

[38] P. F. Brown, P. V Desouza, R. L. Mercer, V. J. Della Pietra, and J. C. Lai, "Class-based n-gram models of natural language," *Comput. Linguist.*, vol. 18, no. 4, pp. 467–479, 1992.

[39] G. Linden, B. Smith, and J. York, "Amazon. com recommendations: Item-to-item collaborative filtering," *Internet Comput. IEEE*, vol. 7, no. 1, pp. 76–80, 2003.

[40] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, 2001, pp. 285–295.