

# Product Review Analysis with Filtering Vulgarity & Ranking System Based On Transaction Id and OTP

Mrs. S. Vidya

Assistant Professor

Department Of Computer Science and Engineering

St. Joseph's Institute Of Technology

India

## ABSTRACT

Recommender systems represent technologies that assist users in finding a set of interesting or relevant items, typically by predicting the “rating” (i.e., an indicator of preference) that users would give to an item they had not yet considered. Recommender systems play an important role in a variety of settings, most prominently ecommerce. Many companies, including Amazon, Netflix, and Pandora, use recommender systems to suggest alternative or cross-selling products to their customers. For example, Netflix has reported that roughly 75 percent of what their subscribers watch (including both DVDs by mail and videos streamed online) has been recommended to customers by its recommender system. Therefore, offering good recommendations to customers is critical in order to retain users and can also contribute to the enhancement of product Reviews, product ranking, opinion mining, POS, xml documents

**Keywords:-** mHLR, OTP

## I. INTRODUCTION

Recommender systems employ users' feedback on consumed items to generate a list of recommended items.

The input used in the recommender systems includes explicitly provided feedback in the form of ratings or tags, as well as feedback that can be implicitly inferred by monitoring users' behavior such as browsing, linking, or buying patterns. The most common approach to modeling users' preferences for items is via numeric ratings. For example, Netflix, an online movie rental company, asks users to rate the movies they have watched on a 5-star numeric scale (with 1 being “Hated it”, and 5—“Loved it”). The Netflix recommender system then analyzes patterns of users' past ratings and predicts users' preferences for unseen movies.

Once ratings for the unseen movies are estimated, the movie(s) with the highest estimated rating(s) can be recommended to the user. In the recommender systems literature, evaluating performance of recommendation algorithms has always been a key issue, and recommendation accuracy has been the major focus in developing evaluation metrics.

As a result, much of the research in the recommender systems area has focused on proposing new techniques to enhance the accuracy of recommendation algorithms in predicting what users will like, as exemplified by the \$1M Netflix prize competition. Prediction accuracy metrics typically

compare the rating values estimated by a recommendation algorithm against the actual rating values and reflect the closeness of the system's predictions to users' true ratings. In addition to accuracy, researchers have proposed a number of alternative types of measures, including recommendation coverage, diversity, novelty, serendipity, and several others, to evaluate the performance of recommender systems. Of special interest to us is the recently introduced measure of recommendation stability.

According to the definition, stability is the consistent agreement of predictions made on the same items by the same algorithm, when any new incoming ratings are in complete agreement to system's prior estimations. Thus, stability is designed to capture the level of internal consistency among predictions made by a given recommendation algorithm. Consider an example where the system makes rating predictions for two items,  $i_1$  and  $i_2$ , for user  $u$ , and that both predictions are of 5 stars. Let's assume that prediction for  $i_1$  is precisely accurate (i.e., user  $u$  completely agrees with the system) and, after user  $u$  consumes item  $i_1$ , this 5-star rating gets added to the system as part of the known ratings. The rating values estimated by a recommendation algorithm against the actual rating values and reflect Prediction accuracy metrics typically compare the closeness of the system's predictions to users' true ratings.

Additionally, the process of iterative smoothing involves multiple iterations to adjust the predictions of unknown ratings. One of the goals in this study is to find whether predictions converge during the process of iterative smoothing and, if so, when. This is directly related to a key decision in iterative smoothing: how many iterations should be used to compute the final predictions?

The recommendation algorithm then re-computes all other predictions in light of the newly added data. A product aspect ranking framework to automatically identify the important aspects of products from numerous consumer reviews. A probabilistic aspect ranking algorithm to infer the importance of various aspects by simultaneously exploiting aspect frequency and the influence of consumers' opinions given to each aspect over their overall opinions on the product.

The potential of aspect ranking in real-world applications. Significant performance improvements are obtained on the applications of document-level sentiment classification and extractive review summarization by making use of aspect ranking. Moreover, the proposed framework and its components are domain-independent and generally applicable in other domains, such as hotel, hawk center, and clothes etc. Firms adopt these systems to provide increase in benefits and their popularity can be explained in the online world. If a customer adopts an RS and purchases a product and finds it not satisfying his expectation, then it is expected that the system would fail to meet the customer requirement.

Support Vector Machines (SVMs) are the newest supervised machine learning technique. SVM uses the notion of a "margin"- a hyper plane that divide two data classes.. An upper bound on the expected generalization error can be reduced by maximizing the margin and thereby largest possible distance between separating hyper plane instances on either side of it. In the case of linearly separable data, once the optimum separating hyper plane is found, data points that lie on its margin are known as support vector points whose solution is represented as a linear combination of only these points . Other data points are ignored.

Therefore, the model complexity of an SVM is unaffected by the number of features encountered in the training data. For this reason, SVMs are well suited for learning tasks where the number of features is large with respect to the number of training instances.

## **II. PROBLEM DEFINITION**

The existing system suffers mainly from high time consumption, less data transmission rate and less secure. There is no Proper stability in the product recommender system. This makes it a bit unreliable in terms of recommending a product.

The recommendations are based on both the anonymous and authenticated user. The proposed system tries to avoid the recommendation based on the anonymous and non-authenticated user. So the recommender system will have higher stability in terms of recommending a product.

Providing stable and consistent recommendations is important in many contexts, prior research has demonstrated that some of the most popular algorithms used in real-world applications (e.g., the widely used item- and user-based nearest-neighbor collaborative filtering approaches) suffer from high degree of instability. Stability of a recommender system measures the consistency of its predictions. It is an important property of recommendation algorithms, because unstable or inconsistent recommendations could lead to user confusion and reduce trust in recommender systems, which in turn can have negative impact on users' acceptance and harm the success of the system.

The iterative smoothing approach uses multiple iterations to repeatedly and explicitly adjust predictions of a recommendation algorithm based on its other predictions in order to make them more consistent with each other.

## **III. LITERATURE SURVEY**

Rutuja Tikait, Ranjana Badre, Mayura Kinikar et.al proposed a product that may have hundred of aspects[1]. Some of the product aspects are more important than the others and have strong influence on the eventual consumer's decision making as well as firm's product development strategies. Identification of important product aspects become necessary as both consumers and firms are benefited by this. Consumers can easily make purchasing decision by paying attention to the important aspects as well as firms can focus on improving the quality of these aspects and thus enhance product reputation efficiently. This paper provides the description of various techniques for product aspect identification and classification. This in turn will help the customer to buy the product based on the reviews[1]. However there may be some demerits in this technique where a brand product

cannot be determined based only on the review.

Ming Ha, Christian Rohrdantz et.al proposed Twitter currently receives over 190 million tweets (small text-based Web posts) and manufacturing companies receive over 10 thousand web product surveys a day, in which people share their thoughts regarding a wide range of products and their features. A large number of tweets and customer surveys include opinions about products and services[2]. However, with Twitter being a relatively new phenomenon, these tweets are underutilized as a source for determining customer sentiments. To explore high-volume customer feedback streams, this paper integrate three time series-based visual analysis techniques: [1]feature-based sentiment analysis that extracts, measures, and maps customer feedback; [2] a novel idea of term associations that identify attributes, verbs, and adjectives frequently occurring together; and new pixel cell-based sentiment calendars, geo-temporal map visualizations and self-organizing maps to identify co-occurring and influential opinions. This paper combines the techniques into a well-fitted solution for an effective analysis of large customer feedback streams such as for movie reviews (e.g., Kung-Fu Panda) or web surveys (buyers).

Güneş Erkan, Dragomir R. Radev et.al proposed the paper that introduces a stochastic graph-based method for computing relative importance of textual units or Natural Language Processing. This paper tests the technique on the problem of Text Summarization (TS). Extractive TS relies on the concept of sentence salience to identify the most important sentences in a document or set of documents. Salience is typically defined in terms of the presence of particular important words or in terms of similarity to a centroid pseudo-sentence. This paper creates an approach [3] LexRank, for computing sentence importance based on the concept of eigenvector centrality in a graph representation of sentences. In this model, a connectivity matrix based on intra-sentence cosine similarity is used as the adjacency matrix of the graph representation of sentences. This system, based on LexRank ranked in first place in computing more than one task in the recent DUC 2004 evaluation. In this paper, a detailed analysis approach is applied to a larger data set including data from earlier DUC evaluations. [3]This paper discusses several methods to compute centrality using the similarity graph. The results show that degree-based methods (including LexRank) outperform both centroid-based methods and other systems participating in DUC in most of the cases. Furthermore, the LexRank with threshold method

outperforms the other degree-based techniques including continuous LexRank. This paper describes that the approach is quite insensitive to the noise in the data that may result from an imperfect topical clustering of document.

W. Liu, H. Liu, D. Tao\*, Y. Wang, Ke Lu et.al proposed that, with the rapid development of social media sharing, people often need to manage the growing volume of multimedia data such as large scale video classification and annotation, especially to organize those videos containing human activities. Recently, manifold regularized semi-supervised learning (SSL), which explores the intrinsic data probability distribution and then improves the generalization ability with only a small number of labeled data, has emerged as a promising paradigm for semiautomatic video classification. In addition, human action videos often have multi-modal content and different representations. To tackle the above problems, this paper proposes multiview Hessian regularized logistic regression (mHLR) for human action recognition. Compared with existing work, the advantages of mHLR lie in three folds: [1]mHLR combines multiple Hessian regularization, each of which obtained from a particular representation of instance, to leverage the exploring of local geometry; [2]mHLR naturally handle multi-view instances with multiple representations; [3]mHLR employs a smooth loss function and then can be effectively optimized. This paper carefully conducts extensive experiments on the unstructured social activity attribute (USAA) dataset and the experimental results demonstrate the effectiveness of the proposed multiview Hessian regularized logistic regression for human action recognition.

Yu Zhang et.al proposed that all the existing multi-task local learning methods are defined on homogeneous neighborhood which consists of all data points from only one task. In this paper, different from existing methods, it proposes local learning methods for multitask classification and regression problems based on heterogeneous neighborhood which is defined on data points from all tasks. Specifically, this paper extends the k- nearest-neighbor classifier by formulating the decision function for each data point as this paperrighted voting among the neighbors from all tasks where this paper rights are task-specific. By defining a regularizer to enforce the task-specific thing, thispaper uses the right matrix to approach a symmetric one, a regularized objective function is proposed and an efficient coordinate descent method is developed to solve it. For regression problems, this paper extends the kernel regression to multi-task

setting in a similar way to the classification case. Experiments on some toy data and real-world datasets demonstrate the effectiveness of our proposed methods.

Raphael Hunger, Michael Joham, and Wolfgang Utschick et.al proposed that a novel MSE balancing paradigm is presented which is based on the sum-MSE minimization with individual. This paper is constraint in conjunction with an adaptive allocation. By means of a counter-example, this paper shows that even the two-user MSE region may be non-convex in case of multi-antenna user which is in contrast to the predominant opinion. Any algorithm utilizing the sum-MSE minimization hence cannot achieve the complete MSE region for any channel realization. [6] This resulting gap due to the nonconvex dent is closed by the presented approach. For the sum-MSE minimization with individual constraints which represents the core of the balancing algorithm, this paper comes up with an extremely fast converging alternating optimization outperforming all hitherto existing approaches.

MingjieQian, ChengxiangZhai et.al proposed a new unsupervised feature selection method, i.e., Robust Unsupervised Feature Selection (RUFFS), is proposed. Unlike traditional unsupervised feature selection methods, pseudo cluster labels are learned via local learning regularized robust nonnegative matrix factorization. During the label learning process, feature selection is performed simultaneously by robust joint  $l_{2,1}$  norms minimization. Since RUFFS utilizes  $l_{2,1}$  norm minimization on processes of both label learning and feature learning, outliers and noise could be effectively handled and redundant or noisy features could be effectively reduced. This project adopts the advantages of robust nonnegative matrix factorization, local learning, and robust feature learning. In order to make RUFFS be scalable,[7] this paper designs a (projected) limited- memory BFGS based iterative algorithm to efficiently solve the optimization problem of RUFFS in terms of both memory consumption and computation complexity. Experimental results on different benchmark real world datasets show the promising performance of RUFFS over the state-of-the-arts.

Chih-Chung Chang and Chih-Jen Lin et.al proposed [8] LIBSVM is a library for Support Vector Machines (SVMs). This paper has been actively developing this package since the year 2000. The goal is to help users to easily apply SVM to their applications. LIBSVM has gained wide popularity in machine learning and many other areas. In this article, describes all implementation details of

LIBSVM. Issues such as solving SVM optimization problems, theoretical convergence, multi-class classification, probability estimates, and parameter selection are discussed in detail.

Jianxing Yu, Zheng-Jun Zha, MengWang, Tat-Seng Chua et.al proposed [9] this paper, aims to automatically identify important product aspects from online consumer reviews. The important aspects are identified according to two observations: (a) the important aspects of a product are

usually commented by a large number of consumers; and (b) consumers opinions on the important aspects greatly influence their overall opinions on the product. In particular, given consumer reviews of a product, [9] this paper first identifies the product aspects by a shallow dependency parser and determine consumer's opinions on these aspects via a sentiment classifier. This paper then develops an aspect ranking algorithm to identify the important aspects by simultaneously considering the aspect frequency and the influence of consumer's opinion given to each aspect on their overall opinions. The experimental results on 11 popular products in four domains demonstrate the effectiveness of our approach. This paper further applies the aspect ranking results to the application of documentlevel sentiment classification, and improve the performance significantly.

Jianxing Yu, Zheng-Jun Zha, Meng Wang, Tat-Seng Chua et.al proposed [10] This paper dedicates to the topic of aspect ranking, which aims to automatically identify important product aspects from online consumer reviews. this paper first identifies the product aspects by a shallow dependency parser and determine consumer's opinions on these aspects via a sentiment classifier. [10]This paper then develops an aspect ranking algorithm to identify the important aspects by simultaneously considering the aspect frequency and the influence of consumer's opinion given to each aspect on their overall opinions.

#### **IV.EXISTING SYSTEM**

In the existing system, Content based, Collaborative Filtering & Hybrid algorithm has been used. But the major problem is we could not come up proper stability with this Recommendation process. In the existing system, the product ranking will be based only on the user perception. The existing system comprises of the feedbacks of the product, however the feedback that has been entered cannot be identified whether it is from the authenticated user who has actually purchased and used the product. The ranking of a product is purely based on analyzing the feedbacks of the user. The technique has major

drawback of distinguishing an authenticated user and a non-authenticated user. In the previous technique they use iterative smoothing algorithm to analyze the feedback of the user. The ratings and the rankings of the product are purely based on the user perception.

#### 4.1 Disadvantages Of Existing System

- High time consumption
- Less data transmission rate
- Less effective
- Instability
- Less security

## V. PROPOSED SYSTEM

In the proposed system, Iterative Smoothing along with SVM algorithm is used which is aimed at stability improvement. The key idea is predictions computed during the current session are the feedback into the data to predict other instances in subsequent iterations. In the proposed system, the product ranking will be based on the frequently commented in consumer reviews and the consumer's opinions on these aspects which greatly influence their overall opinions on the product. According to the product aspect identification and sentiment classification the product will be ranked. The modification in the project will generate the Transaction ID for Every Product Purchase. User will give the Feedback about the Product by Keying the Transaction ID. Once the Transaction ID is matched, an OTP will be generated to the User's Mobile Number. Only after Authentication Feedback is Accepted and Published in the Website. User is not allowed to give Feedback for Non Purchased Products & Area.

#### 5.1 Advantages Of Proposed System

- Less time consumption
- High data transmission rate
- More effective
- Stability
- High security

## VI. SYSTEM DESIGN

People search for products eagerly to buy good product. This is due to enormous production of large number of products in the world. Their decision

to choose a product highly depends on word-of-mouth. Customers greatly observe the views of different people to make decisions. For this, new system emerged called Recommender systems (RS). They help people to get products of their interest. Many people perform more search operation to choose right products. Many people don't know the right way to get products of their interest. Recommender Systems helps consumer to choose the product among so many options.

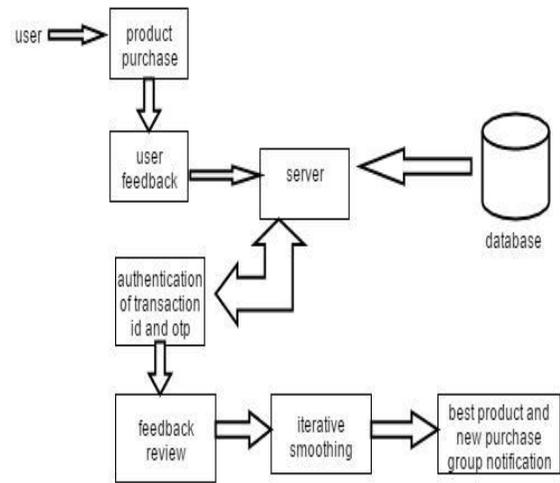


Fig.1. Control flow of system and working procedure

## VII. MODULES

### 7.1 User Registration

In client side user can enter all details. Then user can login using particular username and password. All the inserted and updated items are added into the product list. Then user selects the wanted items and add all items into cart products with count of the each item. A warning message will display in the dialogue box when the customer type the quantity above the constraint value mentioned in the database. All selected items are displayed in the cart product list. Then purchase the required items.

### 7.2 Purchase Portal

Consumer's behavior is the sum of a consumer's attitudes, preferences, intentions and choices relating to the consumer's behavior in the marketplace when purchasing a product or service. The study of consumer behavior draws upon social science disciplines of sociology, and economics. At this stage, the consumer will make a purchasing decision. The ultimate decision may be based on

factors such as price or availability. For instance, our client has determined to get a specific model of car because its price was the most effective, they could negotiate and the car was accessible rightaway.

### 7.3 SERVER

Also the Server needs to establish the connection to communicate with the Users. The Server can update the each User's activities in its database. The Server will authenticate each user before they access the Application. So, that the Server can stop the Unauthorized User from accessing the Application.

### 7.4 Feedback

Process in which the effect or output of an action is 'returned' (fed-back) to change the consequent action. Feedback is important to the working and survival of all regulatory mechanisms found throughout living and non-living nature, and in man-made systems such as education system, online shopping system and economy. As a two-way flow, feedback is inherent to all interactions, whether human-to-human, human-to-machine, or machine-to-machine.

In an organizational context, feedback is the information sent to an entity(individual or group) regarding its previous behavior so that the entity may adjust its current and future behavior to achieve the required result. Feedback occurs when an environment reacts to an action or behavior.For example customer's feedback is the buyers reaction to a firm products and policies and operational feedback is the internally generated information on a firm's performance. Response to a stimuli is considered as a feedback only if it brings a modification in the recipient's behavior.

### 7.5 Otp Generation And Verification

A one-time password (OTP) is a password that is valid for only one login session or transaction. OTPs avoid a number of shortcomings that are associated with traditional (static) passwords. The most important shortcoming that is addressed by OTPs is that, in contrast to static passwords, they are not vulnerable to replay attacks. This means that a potential intruder who manages to record an OTP that was already used to log into a service or to conduct a transaction will not be able to abuse it, since it will be no longer valid. On the downside, OTPs are difficult for human beings to memorize. Therefore they require additional technology to work and verification code will be sent to the e-mail after that only feedback is accepted Based on the feedback value, the admin will rate the promising items and find out the promising items. Candidate item sets can

### 7.6 PRODUCT RANKING

The server will monitor the entire user's information in their database and verify them if needed.Also the server will store the entire user's information in their database.

be generated efficiently with only two scans of database. Mining high utility item sets from database refers to the discovery of item sets with high utility like profit. So the user can the feedback base product to purchase. This will be useful for the new user to buy the product

## VIII. SYSTEM ARCHITECTURE

A system architecture or Systems architecture is shown in Fig.2.

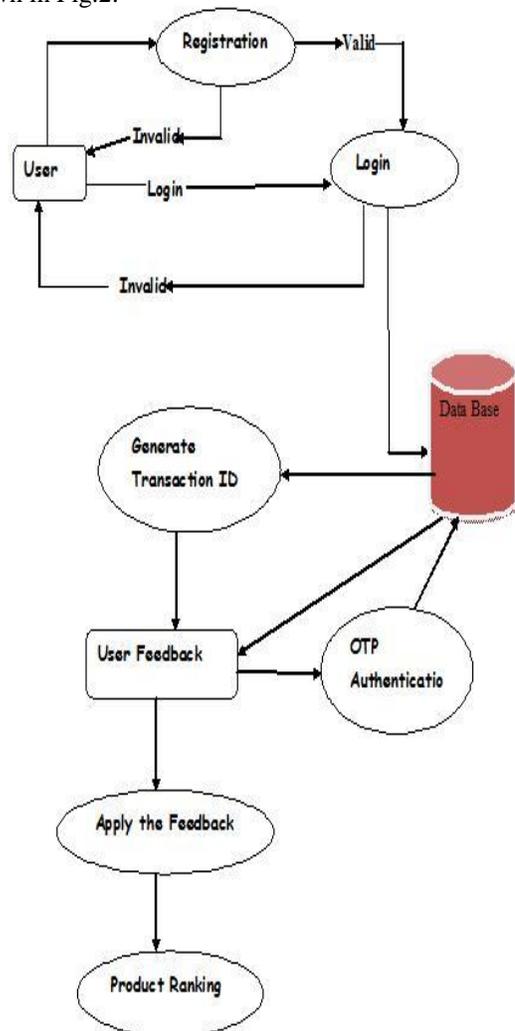


Fig.2.System architecture for product review

It is the conceptual model that defines the structure, behavior and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. The goal of a Recommender System is to generate meaningful recommendations to a collection of users for items or products that might interest them. Recommender Systems helps consumer to choose the product among so many options.

### 8.1 Transaction Id And Otp Generation

Transaction ID is used for authorization for user. Once the id is checked OTP is also checked for conforming the user whether they are human are not. Checking the id and OTP is mainly for the authorization purpose so that authorized person are only allowed to post their feedback. SO two type of password is checked. OTP generation algorithms typically make use of randomness, making prediction of successor OTPs by an attacker difficult, and also hash functions, which can be used to derive a value but are hard to reverse and therefore difficult for an attacker to obtain the data that was used for the hash. This is necessary because otherwise it would be easy to predict future OTPs by observing previous ones.

2nd Factor Authentication

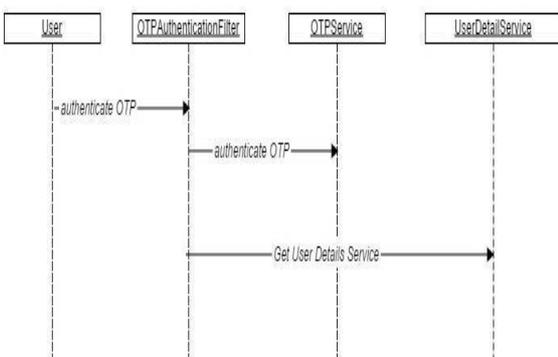


Fig.3.OTP generation process

In Fig. 3, it describes the working procedure of OTP generation algorithm which in turn makes the review posted by the user an effective and secure one. The OTP generation plays a vital role in the whole process as it enhances the system by securing and differentiating authenticated user from anonymous user.

### 8.2 Streaming Algorithm For Removing Vulgar Words:

In user updates there may be many contamination of vulgar words. So to avoid those vulgar word stream clustering is used to avoid those words. All vulgar words are stored as keyword in the database. When the stream cluster start to analyze it match the words with the keyword stored in the database and

block all the vulgar words. User generated data is posts/comments given by people on any news or story of real world.

Posts/comments are formed by sequence of words, keywords or term. So to perform this task there is need of user data stream. Also, the removal of slang words from user data is another approach covered here. For showing results one website is created. Data mining technique clustering is used and one hybrid algorithm is designed for current event detection.

### 8.3 Analysis User And Review Using Svm

SVM is used to analyze the user who post the reviews. After analyzing the post are separated on two basis, positive review and negative review. After separating the review, ranking is based on the reviews. This process is done to identify the best product. Support Vector Machines (SVM) covers emerging techniques which have proven successful in many traditionally neural network (NN)-dominated applications. An interesting property of this approach is that it is an approximate implementation of the Structural Risk Minimisation (SRM) induction principle that aims at minimising a bound on the generalisation error of a model, rather than minimising the mean square error over the data set. In this paper, the basic ideas underlying SVM are reviewed and the potential of this method for feature classification and multiple regression (modelling) problems is demonstrated using digital remote sensing data and data on the horizontal force exerted by dynamic waves on a vertical structure respectively.

SVMs can be used to solve various real world problems. SVMs are helpful in text and hypertext categorization as their application can significantly reduce the need for labeled training instances in both the standard inductive and transductive settings. Classification of images can also be performed using SVMs. Experimental results show that SVMs achieve significantly higher search accuracy than traditional query refinement schemes after just three to four rounds of relevance feedback.

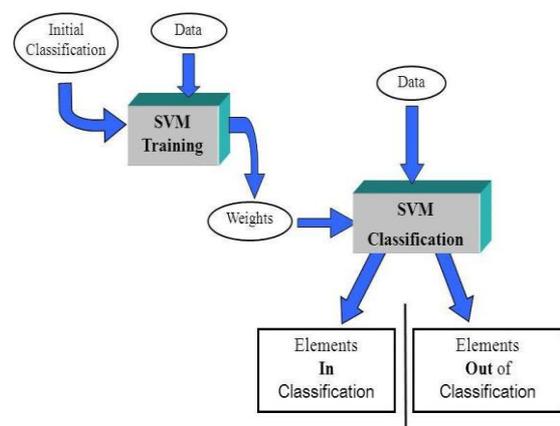


Fig.4.SVM Process Overview

#### **8.4 Recommendation:**

The user who have same product are grouped in same group but no one know each other that they are their group meat. But one person from the group is interested to buy other product and searched for new product then automatically information is sent to the group member that your group member is interested on so and so product so are you interested to buy that product. This kind of recommendation is given to the user who are all in same group. Retailers can make personalized, data-driven product recommendations by looking at a customer's past purchase history. Working Person's Store offers personalized customer experiences with the help of targeted product recommendation campaigns. Working Person's Store triggers this personalized message to customers who purchase specific products as a way to motivate a repeat purchase. The email includes recommendations for products that accessorize with or are similar to the originally purchased item(s), with colorful images and links to specific product pages. The light-hearted and conversational tone of the message also works to further engage the customer.

#### **IX. CONCLUSION**

Stability of a recommender system measures the consistency of its predictions. It is an important property of recommendation algorithms, because unstable or inconsistent recommendations could lead to user confusion and reduce trust in recommender systems, which in turn can have negative impact on users' acceptance and harm the success of the system. This paper explores two general-purpose and practical meta-algorithmic approaches—based on the traditional bagging technique and the proposed iterative smoothing technique—that can be used to improve stability of a wide variety of state-of-the-art recommendation algorithms. The bagging approach extracts multiple training sub-samples from the original dataset and combine predictions made based on these samples to form an aggregate final prediction for each unknown rating. The iterative smoothing approach uses multiple iterations to repeatedly and explicitly adjust predictions of a recommendation algorithm based on its other predictions in order to make them more consistent with each other.

#### **X. FUTURE ENHANCEMENT**

Recommender system can improve the stability along with the consistency by implementing the algorithm that works well with the system. In this we try to improve the stability of the system but the consistency was not considered. So the future work might try to be user friendly with limited complexity and higher consistency. Recommender system can enhance the user by interacting with them and finding out the interest of user by their frequent search of item/product in the social media. It can also engage the user friendly option

where it makes the recommendation a bit easier and sorting out the unwanted product/item listed for the user.

#### **REFERENCES**

- [1] Abd. Samad Hasan Basari, Burairah Hussin, I. Gede Pramudya Ananta and Junta Zeniarja, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization", Published by Elsevier Ltd 2013.
- [2] Yin-Fu Huang and Heng Lin, "Web Product Ranking Using Opinion Mining" Conference on Computational Intelligence and Data Mining (CIDM), 2013 IEEE Symposium on, pages.184 – 190.
- [3] Tanvir Ahmad and Mohammad Najmud Doja, "Opinion Mining using Frequent Pattern Growth Method from Unstructured Text", Computational and Business Intelligence (ISCBI), 2013 International Symposium on IEEE Conference Publications, pages. 92 – 95.
- [4] Po-Wei Liang and Bi-Ru Dai, "Opinion Mining on Social Media Data", 14th International Conference on Mobile Data Management, 2013 IEEE , pages 91-96.
- [5] Jalel Akaichi, Zeineb Dhouioui and Maria José López- Huertas Pérez, "Text Mining Facebook Status Updates for Sentiment Classification" System Theory, Control and Computing (ICSTCC), 2013 17th International Conference, pages. 640 - 645
- [6] Akaichi, J., "Social Networks' Facebook' Statutes Updates Mining for Sentiment Classification", Social Computing (SocialCom), 2013 International Conference on, pages. 886 – 891.
- [7] Na Chen and Viktor K. Prasanna, "Rankbox: An Adaptive Ranking System for Mining Complex Semantic Relationships Using User Feedback", Information Reuse and Integration (IRI), 2012 IEEE 13th International Conference on, pages. 77 - 84.
- [8] Darena, F., Zizka, J., Burda, K., "Grouping of Customer Opinions Written in Natural Language Using Unsupervised Machine Learning", Symbolic and Numeric

- Algorithms for Scientific Computing (SYNASC), 2012 14th International Symposium on, pages. 265 – 270.
- [9] Ningzhong, Yuefeng Li, and Sheng-Tang Wu, “Effective Pattern Discovery for Text Mining”, IEEE transactions on knowledge and data engineering, vol. 24, no. 1, January 2012, pages. 30 – 44.
- [10] Dominique Ziegelmayr and Rainer Schrader, “Sentiment polarity classification using statistical data compression models”, 2012 IEEE 12th International Conference on Data Mining Workshops, pages. 731 – 738.
- [11] Balakrishnangokulakrishnan, Pavalanathan Priyanthan , thiruchittampalamragavan , Nadarajah Prasath and ashehan Perera, “Opinion Mining and Sentiment Analysis on a Twitter Data Stream”, The International Conference on Advances in ICT for Emerging Regions – ictcr 2012, pages. 182 – 188.
- [12] Krzysztof Jędrzejewski, Mikołaj Morzy proposed “Opinion Mining and Social Networks: A Promising Match”, 2011 International Conference on Advances in Social Networks Analysis and Mining, pages. 599 – 604.
- [13] Jianwei Wu, Bing Xu and Sheng Li “An Unsupervised Approach to Rank Product Reviews”, 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pages. 1769 – 1772.
- [14] Peng Jiang, Chunxia Zhang, Hongping Fu, Zhendong Niu, Qing Yang, “An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews”, 2010 IEEE International Conference on Data Mining, pages. 256 - 265
- [15] Weishu Hu, Zhiguo Gong and Jingzhi Guo, “Mining Product Features from Online Reviews”, 2010 IEEE International Conference on E-Business Engineering, pages. 24 – 29.
- [16] P. Tian, Y. Liu, M. Liu, and S. Zhu, "Research of product ranking technology based on opinion mining" Proc. 2009, The 2nd International Conference on Intelligent Computation Technology and Automation, pages. 239 - 243.
- [17] Juling Ding, Zhongjian Le, Ping Zhou, Gensheng Wang, Wei Shu, “An Opinion-Tree based Flexible Opinion Mining Model”, 2009 International Conference on Web Information Systems and Mining, IEEE, pages. 149 – 152.
- [18] Jung-Yeon Yang, Jaeseok Myung and Sang-goo Lee, “The Method for a Summarization of Product Reviews Using the User’s Opinion”, 2009 International Conference on Information, Process, and Knowledge Management, pages. 84 – 89.
- [19] Alexandra, Balahur and Andres Montoyo, “A Feature Dependent Method for Opinion Mining and Classification”, Natural Language Processing and Knowledge Engineering, 2008, NLP-KE '08, IEEE
- [20] Jian Liu , Gengfeng Wu and Jianxin Yao, “Opinion Searching in Multi-product Reviews”, Proceedings of The Sixth IEEE International Conference on Computer and Information Technology (CIT'06) 2006, pages. 25