# Trust Based Recommender System in E - Commerce using Sentiment and Entropy Method

Hla Sann Sint
*University of Computer Studies, Yangon*
Yangon, Myanmar
hlasannsint@ucsy.edu.mm

Khine Khine Oo
*University of Computer Studies, Yangon*
Yangon, Myanmar
khinekhineoo@ucsy.edu.mm

*Abstract*— ***The developing for e - commerce, understanding the techniques for trust is basic. Moreover, trust issues are important to a few taking care of corporate obligation; web-based business, and the social communication. The current period pattern is to look into reviews, expert opinions and mining on Web. The user can make an informed trust decision-making. Sentiment analysis, likewise called opinion mining, which is the computational study of opinions, sentiments and users emotions expressed in natural language processing and text analysis. In this framework, assessment of clients on the administrations gave by online marketing website is considered. The assessment or opinion of individuals is inferred by comments from Facebook. This framework put together ascertain with respect to opinion, client similitude is based from the writing's comments by the clients for the trust with text-mining methods utilizing Sentiment Analysis to assist clients with choosing on the web items utilizing on the web reviews. This paper proposes a method using the combination of Sentiment Analysis and Shannon's Entropy in online social commerce. Computing trust is ranking based on the comments given by "trustable users", weighted by their trust score without any direct ratings from customers. Trust decision making output is calculated by Shannon's Entropy. This system based on the individual's accurate information, can get more trusted over trust.***

**Keywords— Online Reviews, Sentiment Analysis, stochastic learning, Data Mining, Entropy, Trust.**

## I. INTRODUCTION

In today's technology, communication has very grown and very easy and cheap to communicate and connect with each other. Communication techniques have improved from wired devices to wireless. The internet has also fast changes in the forms of communications that are available and currently being used. It has led the rise of value of social networks. Trust can be truly important in social associations. In the public arena, trust requirements for individuals, yet additionally establishments and frameworks. There is an overall agreement among contemporary social framework that social trust is significant, for both social and political reasons. Users inclinations and feelings are influenced by other users‟ reviews on the web or writes or over long-range informal communication stages. Sentiment analysis is generally utilizing and applied to feeling of the customers materials, for example, remarks, surveys and evaluations for applications that range from advertising to customer service to purchase the item productively [5][16].

In this system, Sentiment analysis and recommended product regarding the services provided by online marketing websites are considered. Customers emotions are expressed unequivocally with a vote or certainly with comments for item. This framework comes from consolidating the fields of sentiment analysis and suggestion utilizing entropy to produce a unique and functioning recommender framework. To ease online marketing website regarding recommended services for users. It is intended to provide product recommendation system and then to support e-commerce applications.

The rest of this paper is coordinated as follows: describe trust issue in online social marketing in section II and related works in section III. Then discuss evaluation for user trust prediction in section IV and then empirical result is section V. At last, section VI describe conclusion.

## II. TRUST ISSUE IN E- COMMERCE

When looking at social media and relationships, there are plenty of people who get upset that their significant other doesn't post about them often enough. There are also those who feel social media is a downfall to their relationship, as it causes distrust between the couple. "Trust is one of the most important synthetic forces within society". Trust is additionally supposed to be at the focal point of a group of different ideas that are as significant in sociology hypothesis as in down to earth day by day life, including life fulfillment and happiness, idealism, prosperity, well-being, financial success, instructive achievement, government assistance, interest, network, common society, and majority rules system. For instance, there is proof that believing individuals are more advantageous and more joyful and live more than doubting individuals do. Furthermore, obviously, social trust is a center part of social capital, and is typically utilized as a critical pointer of it, in some cases as the best or just single marker. In the event that trust is to be sure as significant as this, at that point it should be very intriguing to find out about the roots of social trust [4][16][17].

## III. RELATED WORKS

R. Lydia Priyadharsini, M. Lovelin Ponn Felciah [13] proposed recommendation system based on hybrid recommendation algorithm for E-commerce web sites and the customer can easily search the products ratings, reviews are analysed by sentiments with the help of review based MAC filtering that can be avoided fake reviews. The benefit of this framework is the visual association of the information dependent on the underlying structure, and a huge decrease in the size of the search space per result output.

Au Yau Leung, Rex[2] investigated trust level online-shopping in Hong Kong. This system result based on primary data analysis and descriptive statistics, reliability analysis and regression analysis. Data collected from the scholars and students were based on by using paper-form questionnaire. The result showed factors that contribute to trust, purchase intentions, and other influencing variables must be measured.

Dimah Alahmadi and XaioJun Zeng[7] proposed a structure of understood social trust and estimation (ISTS) in view of suggestion frameworks from friends‟ short posts in microblogging as miniature reviews by utilizing using machine learning methods including Naive Bayes, Logistic Regression and Decision Trees. The results show that the these short and inconsistent posts can empower the users preferences data in particular when no preferences of history were available.

Sonja Grabner-Kräuter [14] built up a hypothetical system, which encourages a staggered and multidimensional investigation of exploration issues identified with trust in OSNs. This paper talked about various kinds and wellsprings of trust from the trust writing and their significance for trust-related choices and practices in OSNs. The examination uncovered social capital can be seen both as a result picked up by people in an OSN and as an instrument for encouraging the administration of such spaces and furthermore it was demonstrated the connection between social capital and trust isn't unidirectional yet corresponding.

Amel Ziani, Nabiha Azizi, Didier Schwab, Monther Aldwairi, Nassira Chekkai, Djamel Zenakhra, Soraya Cheriguene[1] proposed a multilingual recommender framework dependent on sentiment analysis using online product reviews. This paper was utilized semi regulated SVM for the assessments characterization assignment to evade the absence of explained information issue. The test results recommended high exactness and a review of 100%. This outcomes investigation assessment gives fascinating discoveries on the effect of coordinating feeling examination into a proposal technique based on collaborative filtering.

Some problem drawbacks have not been solved by above mentioned methods in online marketing. Most of research approaches require rating from users to calculate trust evaluations to work properly. In this paper, trust values evaluation with user rating that can be calculated user similarity using sentiment analysis. These are calculated based on reviews by the user.

## IV. CALCULATION OF TRUST PREDICTION

In this section, calculation described the procedure of evaluation trust prediction with sentiment analysis and Shannon entropy. The comparison is done with the single profile in the database that matches the current user best. This method involves gathering data on user preferences and behaviors and then using that data to algorithmically produce recommendations for new users. This system will display a list of recommended products to the end user.

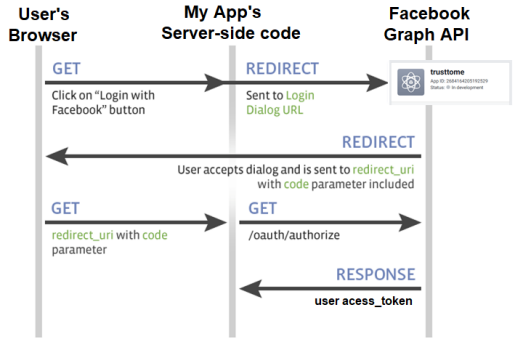


Fig1. User Access token Flow Diagram

In system trust calculated using with Facebook data (such as post, comments). the system admin must be accessed token flow to get the Facebook data.

In flow diagram, firstly user login to the Facebook and find my Facebook id in the find my Facebook id.com and then enter Facebook graph API and get the user access token.

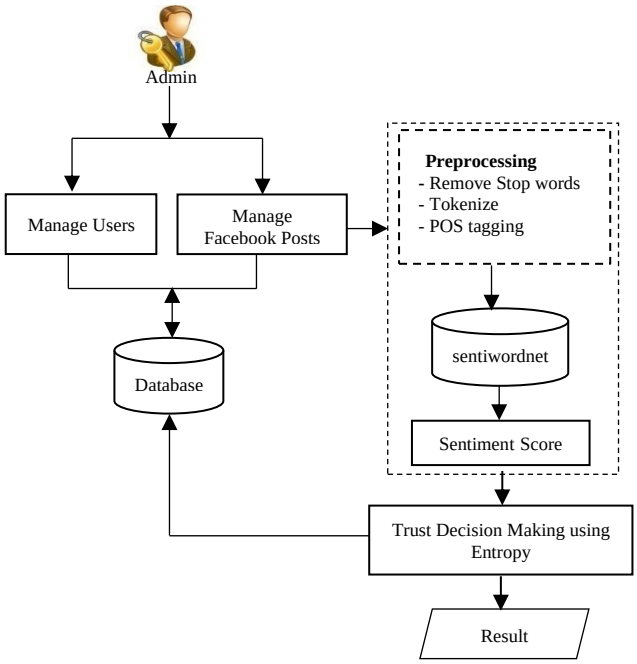


Fig. 2 Trust System Design for Admin

In system design for admin include admin can manage user and manage Facebook post. Manage user is the user enter or registration to system application and also manage Facebook post is to watch trust calculation of Facebook share posts.

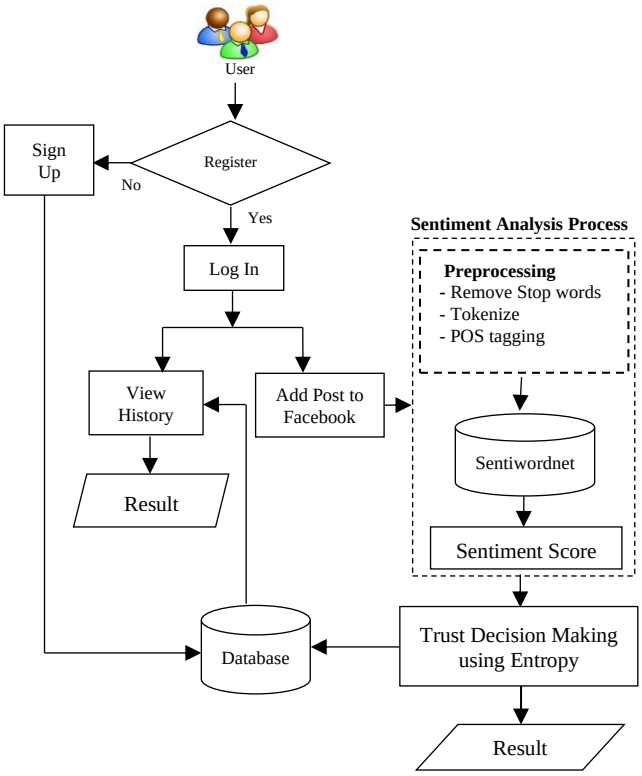


Fig.3. Trust System Design for user

In figure 3, initially, user will login to the framework by utilizing the username and secret key and will survey different items and can give comments about the items. Administrator can examine where the item is positive or negative utilizing sentiment analysis.

As sentiment analysis explains before, the data is collected from by users, and the texts annotated as reviews. The information collected from each of these sources form a text file which would be mined to extract similarity.

## V. Evaluation Methods

### A. *Sentiment Analysis Process*

The first step of sentiment analysis consists of collecting data from user generated comment in Facebook post.

**Removing Stop Words:** A stop word is a usually utilized word (such "a" and "the") that a search engine has been customized to disregard, both when ordering passages for looking and while retrieving them as the result of a search query.

The rundown of words that are not to be added is known as a stop list. Stop words are deemed irrelevant for searching purposes since they happen as often as possible in the language for which the indexing engine has been tuned.

**Replace Words:** A replace words is a believer to single word from given content. The words are changed over as a solitary word as they are the for the most part comparative.

For example, “not too good” sentence replaces with “bad”, “not bad “sentence replaces with “good”, “not too bad” sentence replaces with “nice”.

**Post Tagging:** The process of assigning a tag to a word in a corpus is Tagging. Parts of speech features (POS), word classes, morpho-logical classes, or lexical labels give data dependent on the both its definition and its unique situation. POS Tagging is relationship with adjoining and related words in an expression, sentence, or section.

There are eight basic POS have been distinguished: Noun, verb, pronoun, preposition, adverb, conjunction, adjective, and article

**Analyze sense:** Sentiment analysis is the automated process of understanding an opinion mining for a given subject from written or spoken language [6][7][8][9].

Algorithm 1. Facebook Comments preprocessing algorithm

Output: corpus of normalized comments
I=0, n= number of comments in dataset

1. While (i<n) do
2. input review(i)
3. keep English characters [A-Z, a-z] in comment (i)
4. convert comment (i) to lower case
5. for all words in review(i)
   - Remove stop word in comment(i)
   - Replace word in comment(i)
   - PosTagging(word) in comment(i)
6. add comment to list corpus []
7. i = i + 1
8. end

The preprocessing algorithm process the comments text from posts, which decrease the efficiency of machine learning for sentiment analysis operation such as remove stop words, finding the PosTagging of words in each comment text, replace text and then saving the result refined text in corpus list database [1][10].

In system application database in figure 4 includes 6 entities name UserAcc, FB_group, FB_post, Result, Senti and Post_temp. UserAcc has 6 attributes and that is many to many relationships to FB_group and then FB_group is one to many relationships to FB_post and others entities are also one to many relationships.

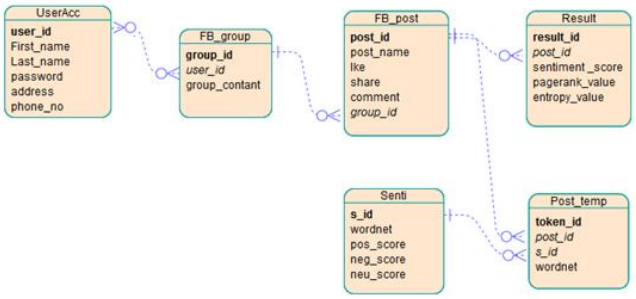


Fig 4. Database Design

## VI. Empirical Results

The following test results are calculated using sentiment analysis with some comment from share Facebook post for online marketing.

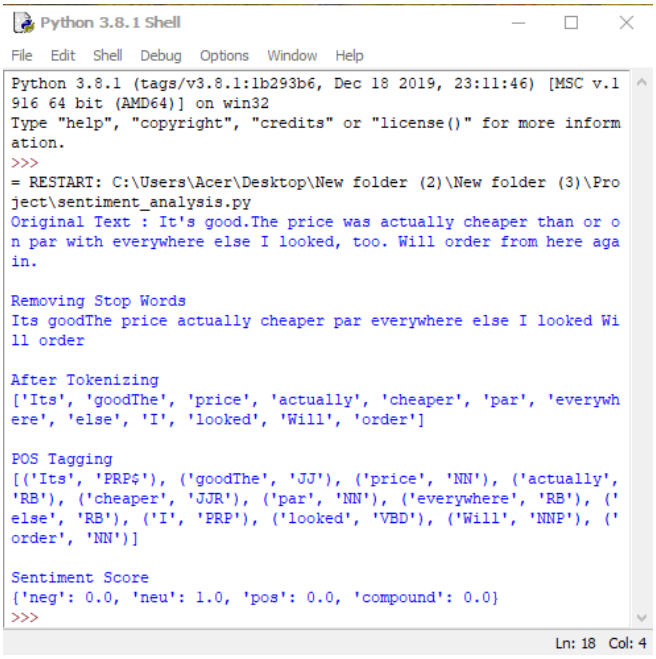


Fig 5. Sample result of comment from Facebook

In figure 5 show the output of sentiment analysis test result for comment from Facebook posts. In testing, the result of removing stop words, tokenizing, POS tagging and sentiment score.

In POS tagging using Penn Treebank Tagset and PRP means Possessive pronoun, JJ is Adjective, NN for Noun, RB is Adverb, JJR is Adjective, comparative, VBD is Verb, past tense NNP is Proper noun, singular.

In Sentiment score, neg is negative value, neu is neutral value and pos is positive value.

TABLE 1. SENTIMENT SCORE FOR FACEBOOK POST'S COMMENTS

| Share Posts from Facebook | Total comments | Negative Score | Neutral Score | Positive Score |
|---|---|---|---|---|
| share1 | 300 | 0.23 | 0.02 | 0.57 |
| Share2 | 250 | 0.2 | 0.4 | 0.4 |
| share3 | 540 | 0.15 | 0.6 | 0.25 |
| share4 | 700 | 0.1 | 0.75 | 0.15 |
| share5 | 159 | 0.5 | 0.8 | 0.85 |

After the sentiment analysis, table 1 show output result for sample five share posts' comments from Facebook.

### B. *Information Entropy*

As a proportion of confusion or vulnerability in the information, was presented by Shannon.

Shannon entropy is portrayed by a level of vulnerability related with the event of the outcome. The Shannon entropy of a random event is subsequently the expected, which thusly can be perceived as the self-data or "surprisal" related with the $i^{th}$ result. Shannon entropy is portrayed by a level of vulnerability related with the event of the outcome. This framework gives the two of sentiment score and Entropy where measure this data by the utilization of the entropy of dynamical organizations on realistic models of social connections [8].

The Shannon entropy can also calculate off-line and achieve more accurate performance.

Shannon Entropy

Entropy H (X) = $-\sum_{i=1}^{n} p(x_i)\, log_2 p(x_i)$

H = entropy

$p(x_i)$ = Pr $(X = x_i)$ is the probability of the $i^{th}$ outcome of X.

This system calculation using the base of Shannon entropy formula.

H(X) = - [(sumpossscore*$log_2$ sumpossscore)
+ (sumnegscore* $log_2$ sumnegscore)
+ (sumneuscore* $log_2$ sumneuscore)]

In trust calculation based on the Shannon entropy,

H(X) = Entropy
Sumposscore= sum positive score
Sumnegscore = sum negative score
Sumneuscore = sum neutral score

$H(X)=-[(2.82log_2 2.82)+(1.18log_2 1.18) + (2.15log_2 2.15)]$
= 1.50

Finally, calculation of trust using Entropy method with above score results output the trust decision making for user. This above value becomes the "Trust Decision Making" that will be shown in percentage.

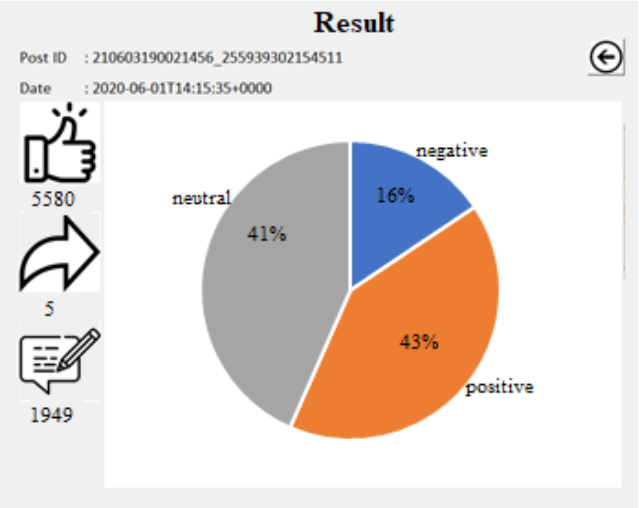


Fig. 6. Empirical Trust Result

In chart fig.6. show the positive, neutral and negative percentage result of sentiment analysis for five share post from Facebook. The user can easily be seen how many trusts on online marketing by using this system.

## VII. CONCLUSION

In this paper is proposed inferring trust values from existing customers in online social commerce. The benefit of this system proposes improving recommender system based on sentiment analysis and Shannon entropy to help customers decide on products using online marketing Facebook comments. Accordingly, it can be used in any online marketing and social networks where some contextual information about the customer is available. Moreover, this framework proposed to execute recommending applied on information from e-commerce and how the consequences of sentiment analysis of textual comments and trust dynamic use with entropy. This assessment can be utilized for customer sentiment regarding with respect to web based promoting and it can likewise be applied to any important and adequately huge informational collection containing notion information. Likewise, Also, Shannon Entropy success in predicting trust for customer based on comments in post shares and also to find the most trusted path in as minimum time more possible than other and at the time of trust calculation.

## REFERENCES


[1] Amel Ziani, Nabiha Azizi, Didier Schwab, Monther Aldwairi, Nassira Chekkai,Djamel Zenakhra, Soraya Cheriguene, "Recommender System Through Sentiment Analysis", 13 Jan 2018.

[2] Au Yau Leung,Rex, "A StudyoftheTrust in On-line Shopping in HongKong: Differences between Scholars and Students", April 2005.

[3] Amany A. Naem, Neveen I. Ghali," Optimizing community detection in social networks using antlion and K-median", v8i4.1196, *IAES journal*, Vol 8, No.4, December 2019, p 1433-1440

[4] Bo Yang, Yu Lei, Jiming Liu, Fellow, "Social Collaborative Filtering by Trust", Transactions on Pattern Analysis and Machine Intelligence , ( Volume: 39 , Issue: 8 , Aug. 1 )2017 IEEE.

[5] Darshana Karna, Ilsang Ko, "We-Intention, Moral Trust and Self-Motivation on Accelerating Knowledge Sharing in Social Collaboration", 48th Hawaii International Conference on System Sciences 2015 IEEE.

[6] Dimah Alahmadi and Xaio-Jun Zeng, "Improving Recommendation Using Trust and Sentiment Inference from OSNs", journal, Vol. 1, No. 1, June 2015.

[7] Edi Sutoyo, Ahmad Almaarif, *"*Twitter sentiment analysis of the relocation of Indonesia's capital city*", IAES journal*, Vol 9, No 4, August 2020, p 1620-1630.

[8] Farhad Hosseinzadeh Lotfi and Reza Fallahnejad," Imprecise Shannon's Entropy and Multi Attribute Decision Making", Entropy 2010, 12, 53-62; doi:10.3390/e12010053, ISSN 1099-4300.

[9] Hla Sann Sint, Khine Khine Oo. "Consumer Trust Recommendation in Online Social Commerce", 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), 2019. ISBN:981-1-7281-3575-5/19/$31.00 2019IEEE, p 445 – 447.

[10] Heru Agus Santoso, Eko Hari Rachmawanto, Adhitya Nugraha, Akbar Aji Nugroho, De Rosal Ignatius Moses Setiadi, Ruri Suko Basuki", Hoax classification and sentiment analysis of Indonesian news using Naive Bayes optimization", TELKOMNIKA Telecommunication, Computing, Electronics and Control Vol. 18, No. 2, April 2020, pp. 799~806.

[11] K. Arun, A. Srinagesh," Multilingual twitter sentiment analysis using machine learning*",* International Journal of Electrical and Computer Engineering (IJECE), Vol 10, No 6, December 2020.

[12] Lu, S., Li, X., Bai, S., Wang, S., (2000) "An improved approach to weighting terms in text", Journal of Chinese Information Processing, Vol. 14(6), pp. 8-13.

[13] Mimi Aminah binti Wan Nordin, Dmitry vedenyapin, Muhammad Fhreza Alghifaari, Teddy Surya Gunawan, "The disruptometer: an artificial intelligence algorithm for market insights", IAES journal, Vol 8, No 2, June 2019, p 727-734.

[14] R. Lydia Priyadharsini, M. Lovelin Ponn Felciah, "Recommendation System in E-Commerce using Sentiment Analysis", journal, Volume 49 Number 7 July 2017.

[15] Sonja Grabner-Krauter and Sofie Bitter, "Trust in online social network :A multifaced perspective", 2015.

[16] Munmun Bhattacharya, Kolkata Nashreen Nesa, "An Algorithm for Predicting Local Trust based on Trust Propagation in Online Social Networks", International Journal of Computer Applications (0975 – 8887) Volume 156 – No 7, December 2016.

[17] Manasa S. M., Manjula S. H., Venugopal K. R., "Trust Aware System for Social Networks: A Comprehensive Survey", International Journal of Computer Applications (0975 – 8887) Volume 162 – No 5, March 2017

[18] Peter Hakansson, Hope Witmer, "Social Media and Trust – A Systematic Literature Review", Journal of business and economics, ISSN 2155-7950, USA, March 2015.

[19] Raed A. Hasan, Royida A. Ibrahem Alhayali, Nashwan Dheyaa Zaki, Ahmed Hussien Ali," An adaptive clustering and classification algorithm for Twitter data streaming in Apache Spark", TELKOMNIKA, Vol.17, No.6, December 2019, pp.3086~3099.

[20] Reed, J., Jiao, Y., Potok, T., Klump, B., Elmore, M., and A. Hurson, (2006), "TF-ICF: A New Term Weighting Scheme for Clustering Dynamic Data Streams", International Conference on Machine Learning and Applications (ICMLA'06) Orlando, Florida, USA, 14-16 December 2006.

.