USER RATING CONFIDENCE PREDICTION FOR PRODUCTS USING SOCIAL USER’S TEXTUAL AND TEMPORAL INFORMATION

1V Swedha, 2A Pravi
1Master of Engineering in Software Engineering Department of Computer Science Anna University, Chennai-600025
2Assistant Professor Department of Information Technology Sri Muthukumaran Institute of Technology, Chennai-600069

ABSTRACT:
Opinion Mining or sentiment analysis is a computation study and ongoing research field which aims to analyse the people’s opinion, sentiment and emotions. E-commerce websites allows the people to share their feedback/opinion as textual reviews regarding the products/services. In traditional, many researches has focused on personalized recommendation, which uses the historical data for recommendation. The historical data contains review rating and reviews for prediction and recommendation. Recommendation mainly focuses on User-item ratings. The problem arises when there is drift in user preferences/interest in purchasing the product. In the proposed work, the concept of User rating’s confidence is used to denote the trustworthiness of user ratings. To predict the user confidence rating and analyse the best qualified product, entropy is calculated. In addition to find accuracy in the user rating confidence, constraints like sentiment and temporal features are used. Finally, all the three factors are merged to find each User item rating and overall item rating. The overall item rating suggests the quality of product. This system results in accurate performance accuracy and better prediction of ratings.

Keywords—Data mining, recommender system, sentiment analysis, review ratings, social networks.

1. INTRODUCTION
In recent years, enhancement of internet access helps the social users to upgrade the knowledge in various domains. E-commerce websites have been used to purchase the products share their opinions and gain much more information regarding the web services. Recommendation system helps to seek the user preference/interest over the rating of an item. Mostly recommendation system based on user–item ratings. There exists some difference while rating and reviewing the products posted by social users. To avoid the difference in rating the prediction, user confidence ratings is enhanced for trustworthiness of user-rating. To represent the inter user similarity between the social users, entropy value is calculated for user item ratings using the old user ratings. To constraint the user ratings confidence, sentiment and temporal features are used. Then the entropy values and constraints are fused to predict the exact user confidence ratings for their rated items. Thus the model uses the pre-processing in which the user confidence rating (entropy values) and low level constraints(sentiment and temporal features extraction using 4th degree Gaussian method) calculation steps are done to find the user confidence rating and product quality. During the low-level constraints calculation, the sentiment score is calculated for predicting the user confidence rating using sentiment features and curve fitting values are calculated using temporal features. Finally all three factors are fused, each user-item confidence rating is calculated and overall item rating is calculated. Thus the overall item rating suggests the quality of the product.

To address the problems in text analysis, the product reviews containing unwanted contents, special symbols problem must be handled. The constraint in the user confidence rating must be handled at regular interval. Since, the user’s interests drifts over change in time. The overall rating of each product must be accurately calculated. The sparse matrix problem must be handled. The proposed work includes the user’s confidence rating calculation and low level constraints calculation to handle the text analysis and overall rating of each product.

The major contributions of the proposed system are: i) filters the unwanted symbols, Uniform Resource Locator (URL)s, stemming the product description. ii)User Rating Confidence values are calculated with entropy value of the user and product rates iii) Low – level constraints values are calculated using the sentiment and temporal features. Curve fittings values are used for calculation. If the UserConfidence ratings are calculated with user confidence values and the low–level constraints values. Finally all three factors are fused, each user-item confidence rating is calculated and overall item rating is calculated. The experimental results shows better performance in rating prediction and in suggesting the quality of the product.

The remainder of this paper is organized as follows. In Section 2, we present the related work. In section 3, we describe the proposed approach. In section 4, the dataset introduction were provided. In section 5, Experiment results
were discussed. In section 6, Conclusions were made. In section 7, References are drawn.

2. RELATED WORK

The opinion mining is used to collect and categorize the opinions about the products and also find the product quality through rating and reviews as stated by the social user’s. It generally analyze the user’s preferences on the product. The purchasing of the products dependson the historical data about the user ratings and the product features. The existing system only calculates the user’s rating according to the reviews and its rating given. The problem occurs when there is deviation in the user’s confidence about the products from time to time. To overcome user’s confidence rating with constraint (sentiment and temporal) and to predict the exact user- rating on the product and to check the overall product quality the proposed system is developed.

A. Entropy Calculation (User Confidence Rating)

User preferences and the overall rating for the products differ over time. Entropy is a measure of uncertainty. The entropy value represents the confidence value of user ratings. So the lower the entropy value is, more stable the product quality is, and the more confident the user’s rating is.

Bell and Volinsky [1] developed a personalized product recommendation which involves Matrix factorization method for analyzing the latent factors and nearest-neighbour techniques. In addition, conventional Singular Value Decomposition (SVD) is used for overcoming sparse matrix. However this system lacks in finding the sentiment score and spatial–temporal features. The dataset used was Amazon and Netflix.

Zhao & Quain [11] developed a framework to predict User Servicerating. This framework fuses four factors namely, user personal interest (related to user and the item’s topics), interpersonal interest similarity (related to user interest), interpersonal rating behavior similarity (related to users’ rating behavior habits), and interpersonal rating behavior diffusion (related to users’ behavior diffusions), into a unified matrix-factorized framework. But however, no temporal features prediction rating were involved to find exactness of user’s interest. The dataset used is Yelp and Douban Movie dataset for prediction.

Zhang & Ding [9] developed collaborative recommender system and used the user-item rating matrix along with sentiment classification to overcome the sparseness of matrix. The system lacked in finding the spatio-temporal features which in-turn lacked in finding user’s preference. Youko (Chinese dataset) was used for experiment.

B. User confidence rating (sentiment features)

Mostly, in e-commerce websites, the users are allowed to rate the product, share their opinions and attitudes by reviewing. The textual reviews and user rating may differ in rare circumstances. This system lacks in the exactness of user rating. It is necessary to analyze the relevance between user confidence and textural review sentiment. So the sentiment score is calculated using the reviews and sentiment dictionary and finally the relevance user confidence sentiment value is predicted.

Chelaru et al. [2] developed a query based recommendation system which involved sentiment analysis using the SentiWordNet thesaurus and a lexical based analysis for resource containing sentiment annotations. Query recommendation employed sentiment detection and sentiment classification to discover the topics. Sentiment Analysis was done with the available reviews and queries.

Tan et al. [6] proposed a Latent Dirichlet Allocation (LDA) using Foreground and Background LDA (FB-LDA) concept. This model was used to distill foreground topics and filter out longstanding background topics between two sets of documents. Twitter Dataset was used for the experiment.

Lei et al. [5] developed a sentiment-based rating prediction to improve prediction accuracy in recommender systems. The user’s sentiment scores and interpersonal sentiment score was calculated along with product reputation. The experiment was done using real-world dataset collected from Yelp. Even though, the system lacked in linguistic rules (and, but) while during sentiment analysis.

C. User Confidence Rating (Temporal Features)

User preference about the product may drift over change time. For predicting the changes, first, the averaged difference between user ratings and overall rating of product is calculated within a period. Secondly, the curve fitting values i.e., peak values, and centroid and peak widths (centroid to other values) are calculated of products within certain period (months) using 4th degree Gaussian method. Finally, the user confidence ratings are calculated using these curve fitting values.

Yang et al. [7] developed a cross-domain feature learning (CDFL) algorithm based on stacked denoising auto-encoder to handle the text, image and video of the context simultaneously. In this model, sentiment classification, spam filtering, and event classification was involved. The experiment used Amazon dataset. This system lacked in spatial features prediction.

Yang & Zhang [7] developed a users’ spatial temporal activity preference(STAP) model where the problem complexity of user’s social activities. This model separately considered the spatial and temporal features of user activities by introducing the notion of spatial specificity and temporal correlation. Finally, a context-aware fusion framework was designed and fused both features in activity preference inference. Three datasets where collected from two LBSNs, i.e., Foursquare and Gowalla was used for experiment.
D. Product Quality Rating Prediction

Product Quality Prediction is done to find exact user confidence rating and overall rating of each product. First the user confidence rating about the product is calculated using the old product rating, entropy values and the low level constraints value (sentiment features and temporal features). Then the overall rating of the product is calculated. This overall rating is used to suggest product quality and the user’s tastes.

Zhao & Qian [4] proposed a model to solve service objective evaluation by deep understanding social users since users’ tastes and habits are drifting over time. This model focused on user ratings confidence (trustworthiness of user ratings) in service objective evaluation. This model utilized entropy to calculate user ratings confidence and spatial-temporal features of user ratings to constrain confidence level of the user ratings. The experiment was conducted on Yelp datasets. The system lacked in finding the sentiments of social users’ reviews.

Lei X. & Qian X. [4] developed a social user’s reviews sentiment measurement system for calculating each user’s sentiment score on items/services, service reputation and accurate rating prediction based on probabilistic matrix factorization. The experiment was conducted on Yelp dataset. The system lacked in calculating the depth of social users’ sentiment.

3. DESIGN AND IMPLEMENTATION

A. Proposed work

The proposed system framework predicts the new user confidence ratings for the given rating after the inclusion of low level constraint. This system starts with the preprocessing stage which filters the unwanted information from the reviews dataset; perform the stop words and stemming process to obtain preprocessed content. The preprocessed content are used to calculate the entropy values and the overall user confidence rating of the products without the constraint information. Later the low level constraints, such as sentiment user confidence rating and the temporal user confidence ratings are calculated. Then the overall user confidence ratings for each user-item pairs are calculated with the entropy and the low level constraints values and finally the overall confidence rating of items are calculated, to find the product quality and likes among the users.

Figure 1 describes the overall system architecture of the proposed user confidence rating prediction for products using social user’s textual and temporal information.

B. Module Description

3.1 Data Pre-processing (Step I)

Input: Product Review Dataset, Stop words list
Output: Pre-processed data

Steps:
1. Loop through the products reviews.
2. Removal of URL, Special character and digits.
3. Tokenization is done for each product reviews.
4. Remove stop-words from tokenized content.
5. Stem each word of the content.
6. Stemmed words for each sentence are recombined.
7. End the loop

Fig. 1 Overall architecture of the proposed framework

3.2 User Rating Confidence Prediction (Step II)

Input: Preprocessed data
Output: Overall user confidence rating for each product.

Steps:
1. The user-item matrix is generated.
2. For the user-item match, particular user rating is populated else 0 is populated. Since entropy is a measure of uncertainty. The information entropy value used to represent the confidence value of user ratings.
3. The entropy value is calculated and then normalized for each user is calculated by cumulating each user rating using Equation (i).

\[ E_u = -\frac{1}{|d_i|} \sum_{d_i} p(d_i) \log p(d_i) \]

Where,

\[ E_u \] is the entropy value of each user, \[ p(d_i) \]

is the probability of item rating.

The lower entropy value is the more stable the system is, the more confident the user’s rating is.

4. Overall user confidence rating for each item is calculated using the entropy values and normalized
values using Equation (ii)

\[ Eu^u = \sum_{i} Eu^u(i) \]  

(ii)

Thus, the overall user confidence ratings for all items without constraints are calculated using Equation (iii) and obtained as output

\[ f_{ui} = \sum_{w} E_{ui} f_{ui} \]  

(iii)

3.3 Sentiment Score Calculation and User Confidence Rating (SentimentFeatures)

Input: Preprocessed Data, Sentiment Dictionary terms
Output: Sentiment score and user confidence rating with sentiment features.

Steps:
1) Loop through the pre-processed contents.
2) Extract all the positive sentiment, negativesentiment, conjugative words and words degree levels.
3) Calculate the sentiment score using Equation (iv)

\[ S(r) = \frac{1}{\sum_i \sum_w Q \omega^{Dw} R_w} \]  

(iv)

where,
- \( S(r) \) - Sentiment Score,
- \( N_c \) – Number of clauses
- \( Q \) – Negation Check Coefficient
- \( Dw \) – Degree level words (most, more …).
- \( Rw \) – Initial Score of sentiment word.

4) Normalize sentiment score

\[ E_{ui} = (10 / (1 + e^{-S[r]})) - 3 \]  

(v)

5) Calculate the user confidence rating with sentiment features.

\[ f_{ui} = \frac{1}{\sum_{w} \omega^{E_{ui}}(E_{ui})} \]  

(vi)

where,
- \( Su,i \) - user confidence rating of user (u) to item(i)
- \( aj \) - curve fitting value (high peak value of item(months))
- \( cj \) - curve fitting value (peak width (months)), \( R_{ui} \)
- Rating given by user (u) to item(i)

3.4 User Confidence Rating (Temporal Features)

Input: Pre-processed Data
Output: User Confidence rating with temporal features

Steps:
1) Loop through the pre-processed contents.
2) Calculate the average rating (user rating and overall rating).
3) Get the peak height (aj), centroid (bj) and peak width (cj) value as curve fitting values using Gaussian method for each month and item.

4) Calculate the user confidence rating with temporal using

\[ T_{ui} = \frac{1}{\sum_j \omega^{E_{ui}(j)}} \]  

(vii)

where,
- \( Tu,i \) - user confidence rating of user (u) to item(i)
- \( aj \) - curve fitting value (high peak value of item(months)), \( \omega^{E_{ui}(j)} \)
- \( bj \) - curve fitting value (centroid(months)), \( \omega^{E_{ui}(j)} \)
- \( cj \) - curve fitting value (peak width (months)), \( R_{ui} \)
- Rating given by user (u) to item(i)

3.5 Product Quality Prediction

Input: \( Tu,i \), \( Su,i \), \( Eu \), R
Output: Overall User Confidence Rating for Items and New user-item rating

Steps:
1) Loop through each item rating given by each user to calculate the overall user confidencerating.

\[ C_{ui} = A_{ui} T_{ui} + B_{ui} \omega^{Dw} R_w + C_{ui} \]  

(viii)

where,
- \( C_{ui} \) - User Confidence rating of user (u) to item(i),
- \( A_{ui} \) - User Confidence rating using temporal features of user (u) to item(i) at time (t),
- \( B_{ui} \) - User Confidence rating using sentiment features of user (u) to item(i) at time (t),
- \( C_{ui} \) - Entropy value with low – level – constraints user (u) to item(i) at time(t).

2) Calculate of overall rating of each item using

\[ r_i = \sum_{t=0}^{T} \left( \omega^{E_{ui}(j)}(E_{ui}(j)) \right) \]  

(ix)

where, \( ri \) - overall user confidence rating, \( Su,i \) - User Confidence rating of user (u) to item(i), \( r(u,i) \) - Rating of user(u) to item(i).

4. EXPERIMENT

A. Dataset Description

The dataset used for the experiment is Amazon. It contains the product reviews according to user’s interest or preferences. The product rating and reviews depict the overall interest of the consumers of Amazon products. List of features extracted from reviews and ratings information allows the system to handle the cold-start (new user and new item) problem. The table 1 contains the characteristics of the dataset which were utilised for the experiment. The important attributes used for the experiment are product id, reviewer id, Overall rating, year.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PID</td>
<td>ID of the product</td>
</tr>
<tr>
<td>ReviewerID</td>
<td>ID of the reviewer</td>
</tr>
<tr>
<td>Reviewer Name</td>
<td>Name of the reviewer</td>
</tr>
<tr>
<td>Year</td>
<td>Year in which the ratings and reviews are calculated.</td>
</tr>
<tr>
<td>Overall rating</td>
<td>Overall Star rating given by the social user’s.</td>
</tr>
</tbody>
</table>
B. Evaluation Measures

This section briefly describes about the evaluation measures.

1) Precision

Precision is known as positive predictive value. It is defined as the number of correct result divided by the number of the retrieved result. Precision returns the exactness (accurate) of the result. If the threshold value changes the precision may increase or decrease accordingly. The precision value is calculated after finding the new rating, to check whether the user had given the positive or negative ratings correctly.

\[
\text{Precision} = \frac{\sum TP}{\sum TP + \sum FP}
\]

where,
TP = True positive, total number of true positive conditions ranked
FP = False positive, total number of false conditions incorrectly ranked.

2) Recall:

Recall is known as true positive value or sensitivity. It is defined as the number of correct result divided by the number of the relevant result. Recall returns the completeness of the result but doesn’t depend on any threshold value. The recall value is calculated after finding the new rating, to check whether the user had given the positive or negative ratings correctly.

\[
\text{Recall} = \frac{\sum TP}{\sum P}
\]

where,
TP = True positive, total number of true positive conditions ranked.
P = Total number of correctly ranked.

Table 2: Calculated Precision and Recall values for various products

<table>
<thead>
<tr>
<th>system</th>
<th>seeds</th>
<th>fertilizer</th>
<th>crops</th>
<th>cereals</th>
</tr>
</thead>
<tbody>
<tr>
<td>entropy + SF</td>
<td>0.9225</td>
<td>0.8267</td>
<td>0.9667</td>
<td>0.9715</td>
</tr>
<tr>
<td>(precision)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entropy + SF</td>
<td>0.9174</td>
<td>0.9667</td>
<td>0.9176</td>
<td>0.9064</td>
</tr>
<tr>
<td>(recall)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entropy + SF + TF</td>
<td>0.9776</td>
<td>0.9729</td>
<td>0.9669</td>
<td>0.9765</td>
</tr>
<tr>
<td>(precision)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entropy + SF + TF</td>
<td>0.9325</td>
<td>0.9668</td>
<td>0.9177</td>
<td>0.9121</td>
</tr>
<tr>
<td>(recall)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3) Accuracy:

The Accuracy or recognition rate measures the proportion of true results to the total number of the instances in the dataset. It specifies how many instances correctly predicted. This affects the system performance. Misclassification measures the error rate (ie) the proportion of false results to the total number of the instances for the given dataset. It specifies how many instances falsely predicted. This affects the system performance.

\[
\text{Accuracy} = \frac{(\sum TP + \sum TN)}{(\sum P + \sum N)}
\]

where,
TP= True positive, total number of true positive reviews is classified as positive.
TN= True negative, total number of true negative reviews is classified as negative.
P = Total number of correctly ranked.

Table 3: Calculated Accuracy and Misclassification percentages for various products

<table>
<thead>
<tr>
<th>system</th>
<th>seeds</th>
<th>fertilizer</th>
<th>crops</th>
<th>cereals</th>
</tr>
</thead>
<tbody>
<tr>
<td>entropy + SF</td>
<td>88.54</td>
<td>83.6</td>
<td>84.2</td>
<td>85</td>
</tr>
<tr>
<td>(Acc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entropy + SF</td>
<td>11.46</td>
<td>16.4</td>
<td>15.8</td>
<td>15</td>
</tr>
<tr>
<td>(Mis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entropy + SF + TF</td>
<td>93.5</td>
<td>88.5</td>
<td>91</td>
<td>87.6</td>
</tr>
<tr>
<td>(Acc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entropy + SF + TF</td>
<td>5.5</td>
<td>11.5</td>
<td>9</td>
<td>12.4</td>
</tr>
<tr>
<td>(Mis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4) **Root Mean Squared Error (RMSE)**

The Root Mean Squared Error (RMSE) measures the average error rate calculated using the actual value and the predicted value. If the measure is low, the system performance is good else not good. If the measure is Zero. It proves that the system is best. The RMSE values is calculated after finding the new ratings for each review. Finally the difference is calculated between old and new ratings to check the error in the system.

$$RMSE=\sqrt{\frac{1}{m}\sum_{i=1}^{m}(t_i - y_i)^2}$$

Where,

- $m$ = total number of reviews accessed,
- $t_i$ = actual value predicted for the reviews.

4) **Mean Absolute Error (MAE)**

The Mean Absolute Error (MAE) measures the average deviation (error) between the predicted overall rating confidence and overall rating. This measure the forecasting error and misleading in the output predicted. If the MAE is 0 it is considered to best prediction result. The MAE value is calculated after finding the new ratings for each review.

$$\text{Misclassification}=1 - \text{accuracy}.$$  

<table>
<thead>
<tr>
<th>system</th>
<th>seeds</th>
<th>fertilizer</th>
<th>crops</th>
<th>cereals</th>
</tr>
</thead>
<tbody>
<tr>
<td>entropy + SF(mae)</td>
<td>0.4798</td>
<td>0.588</td>
<td>0.783</td>
<td>0.635</td>
</tr>
<tr>
<td>entropy + SF(rmse)</td>
<td>0.814</td>
<td>0.795</td>
<td>0.864</td>
<td>0.851</td>
</tr>
<tr>
<td>entropy + SF + TF (mae)</td>
<td>0.2346</td>
<td>0.3346</td>
<td>0.645</td>
<td>0.425</td>
</tr>
<tr>
<td>entropy + SF + TF(rmse)</td>
<td>0.7533</td>
<td>0.7663</td>
<td>0.6953</td>
<td>0.641</td>
</tr>
</tbody>
</table>

Table 4: RMSE and MAE value for various agricultural products.
REFERENCES


