Involve Convolutional-NN to Generate Item Latent Factor Consider Product Genre to Increase Robustness in Product Sparse Data for E-commerce Recommendation

Hanafi$^{1,2}$, N Suryana$^2$ and A S H Basari$^2$

$^1$Department of Computer Science, Universitas Amikom Yogyakarta, Jalan Ringroad Utara, Condongcatur, Depok, Sleman, Yogyakarta 55581, INDONESIA

$^2$Faculty of Information and Communication Technology University Teknikal Malaysia Malaka (UTeM), Jalan Hang Tuah Jaya, Durian Tunggal, Malaka 76100, MALAYSIA

Abstract. Online shopping also popular named e-commerce business need a computer machine to provide product information for customer or buyer candidate. Relevant information served by ecommerce system engine famous dubbed recommender system. It will impact seriously in increasing of marketing target achievement. The character of information in ecommerce have to be specific, personalized, relevant and fit according to customer profiling. There are four kind of recommender system to provide ecommerce recommender system, however only one model that most successful to applied in real ecommerce industry that namely collaborative filtering. These approaches rely on rating as basic calculation to generate product recommendation. However, just a little number of rating that given by customer reference to several convince datasets. The problem causes of sparse product rating, it will bring the impact to product recommendation accuracy. Sometime, in extreme condition impossible to generate product recommendation. Some work efforts have been developing to tackle lack of rating, one of them is considering to involving text sentences document such as text genre, product review, abstract, product description, synopsis, etc. All of material text sentences useful as raw component to predict the product rating. Reference previous work method aims extract text sentences document such as product review to become rating value based on bag of word and word order, sometimes they have got the mis in deeper understanding of text sentences description about product. Therefore, it influenced the result of predict the rating was inaccurate. In this research, author proposed novel method enhance variant of convolutional neural network dubbed dynamic convolutional neural network to improve scalability and increase deeper understanding to increase accuracy level. Based on our experiment, our model outperforms over existing state of the art in previous work in extracting text sentences product description based on evaluation approach uses RMSE.

1. Introduction

E-commerce was changing the mode in many companies in the word to do business transaction. E-commerce is not only has become alternative but an imperative. Many companies face with the most classical matter: what is the best way for developing and conducting business in the digital economy? Some companies are moving their businesses entirely to the Web (e.g.egghead.com). Some are establishing subsidiaries, then spinning them off as separate online business entities (e.g. barnesandnoble.com) [1]. Develop E-commerce business necessary a machine to serve relevant information about product then deliver through web sites portal or mobile phone, the system famous called recommender system. Recommender framework is a basic part in industry region. It is an essential equipment to share information about product and service for so many online e-commerce portal and mobile applications. For cases, 80 percent of films viewed on Netflix originated from suggestions [2], 60 percent of video clicks originated from web page preference in YouTube [3]. Following [4] found that business operators with recommendation by the NetPerceptions framework accomplished 60% higher classical approach and 50% higher achievement target than utilizing conventional approach dependent on investigations held at a U.K.- based retail and business gathering, GUS pls (www.gus.co.uk).
Recommender system divided into 4 types based on technical approach [5], 1). Content based: the model of recommendation considers of item distinction approach. It is tending information retrieval to produce item suggestion 2). Knowledge based: this method develops for specific necessary recommendation, the specific character is to provide product infrequently to purchase for personal requirement for instance building, leasing, assurance, wagon. 3). Demographic based: product recommendation in which established to provide product recommendation based demographic information. 4). Collaborative Filtering: product recommendation based on user behaviour in the past for example the term of behaviour is rating, comment, testimony, purchasing and etc. This method is the most successful approach that applied in many large e-commerce company, due they have ability to provide recommendation character as follow; provide product fit, serve relevant information, accurate, serendipity [6]. In widely use, collaborative filtering adopted rating as explicit feedback to calculate similarity user for product uses rating matrix to generate product recommendation. The big problem in collaborative filtering is just slight of user population who gave rating for product, totally about only less than 1 percent. The problem popular called sparse data also in extreme condition sparse data famous called cold start. When cold start happens, there is no recommendation possible generated by system.

An example works to solve the serious issue as mention on above by using concentrating information from items. Fig 1 on below display an example of product review on movie commerce given by customer. Extract content feature have done by several author, for instance author [7] develop music recommendation involve deep convolutional neural network (deep CNN) to make classification music based on type of sound (music). The approach as mention also to handling unpopular music (long tail). Different work by author [8] enhance extraction deep colour feature to create fashion recommendation uses colour classifier. This research aims to predict rating of the product by use product genre by textual analysis involving deep learning convolutional layer. Its important to do, due several approaches based on convolutional use text document are very high computation and having scalability issue point of view. Enhancing text to build a part of recommendation component have done many researcher, for example [9] involve traditional method use bag of word come from users to estimate rating grade. By using classical approach includes bag of word and word order still having shortcoming. There are misunderstanding to detect the meaning of text or sentences. It became a reason, the result of recommendation base on sentences extraction sometime have been inaccurate. For example, movie review portal from IMDB as a major popular movie review. There are several text resources aim to improve accuracy for example product description, user opinion, user comment, testimony. According fig 1, IMDB show a product description by producer.

In another challenge, movie portal facility where customer is free to give an opinion, in my best knowledge it is become a representation of product rating. However, sentences necessary to converted to be a rating star. An example of movie review shown on figure 2. According marketing representation study, mostly customer would not purchase a novelty item unless impression. Customer would make sure the quality of the product involving opinion or testimony for every customer that having expertise based on product has been consumptions.
The contribution in this study is develop novel approach by use convolutional neural network to catch the relationship item latent factor based on text sentence sentiment, so that it can interpretation the movie specification. Then produce a novel method to create a product recommendation movie genre involve text analysis point of view.

2. Previous Work

CF is the most favourite recommender system approach that implemented in online business due they have good characteristic of being able to serve important product information and giving revelation information to e-commerce customers in the form of product information. Even so, collaborative filtering has a very prominent weakness, namely rare data due lack of customers in giving product rating. Because collaborative filtering relies on rating to calculating a product recommendation. Some efforts have been developing to decrease rarely data so that the recommendations still accurate. The involving of side information is sure to improve the accuracy of a product recommendation. Some of the auxiliary information that have been explored to handle sparse data for instance audio features in music recommendations [7] [10], colour characteristic for fashion commerce transaction [8], documents recommendation for news online [11]. Utilize of sentences to improve accuracy has been done by several researchers. Extracting feature-based content feature, there are several types of sentence expressions for instance, items descriptions, synopsis, item review, comments, and testimonials. According authors [12] we recapitulate the precious information that could be concentrate from reviews and employed to improve as mentioned on above became normative recommendation model. Author [13] proposed a model of text sentence document representation involving bag of word (BOW) method to utilize user review movie to enhance predict rating value based on product review. Exploit bag of word can be work normally, even though the classical problem of this method is less ability to capture contextual meaning. It was failing to detect “not good”, “not bad”, “not very well” for example case. It’s the reason the product recommendation result will inaccurate. There is a challenge how to increase tier of accuracy that have applied involve CNN for instance in text sentence classification [14].

| No | Method | Description | Ref. |
|----|--------|-------------|-----|
| 1  | PMF    | Probabilistic Matrix Factorization (PMF) is a favourite model for rating prediction that just consideration rating as probabilistic matrix factorisation approach by users. | [11] |
| 2  | CTR    | Collaborative Topic Regression (CTR) is another genius approach recommender system method, which incorporating collaborative filtering (PMF) and topic modelling approach based on Latent Dirichlet Architecture (LDA) to use both ratings and text sentence documents. | [12] |
| 3  | NMF    | Recommender system based on collaborative filtering applied use non-negative matrix factorization (NMF) to generate recommendation | [13] |
| 4  | CDL    | Collaborative Deep Learning (CDL) is another model recommendation model in which involving with enhancing rating prediction to increasing the accuracy by analysing text sentences documents of product review by using deep learning approach. | [14] |
| 5  | SVD    | Recommender system based on collaborative filtering applied use Singular value decomposition (SVD) as low rank dimensional factorization. | [15] |

3. Proposed Method

This research aims to deal with the problems as explained on above, we proposed novel approach that consideration convolutional neural network a class of deep learning. Our model focus in generating item/product latent factor that involve product genre, where in our research we applied product genre as a major element to generate latent factor based on genre contextual model. Indeed, Several approaches aim to improve effectiveness item latent factor representation on previous work that adopted from
convolutional strategy to concentrate content characteristic according to text sentence document [15], the detail technical approach and enhancing method displayed on below involving several method as follow:

![Figure 3](image.png)

**Figure 3.** Novel approach movie genre based on sentiment analysis

### 3.1. Text Sentiment Analysis for product genre based on Convolutional NN

The CNN model developed that’s aiming to acquire text sentences document latent vector from genre of item. Where to combine the items latent factor with epsilon value. Figure 4 describe our CNN model that including 4-layer approach 1. Step input layer, 2). Step embedding layer, 3). Step convolutional layer, 4). Step pooling layer, 5). Step generate output layer.

**Input Layer task:** this step has major task to generate information resource based on feature information of the content such as product detail specification. There are several approaches to extract content feature for example music recommender system, online fashion shop and several approaches based on text document description of the product. The most common product detail description is the term in text document. This layer has important task to transport product description into text sentence imputation become preprocessing stage.

**Embedding word layer task:** the task of this stage is to convert of raw text sentences document into matrix array following the number of the words, that will be transform to convolution production in the text pattern. For example, when the number of words in a text sentences document is \( l \), then a vector of each word could be combined to the matrix referring with the sequences. The vector developed with a pre trained processing with glove in example as follow:
Convert word to numeric process where $p$ stands for the dimension of word embedding and $w_{[1:p,i]}$ represents raw word $i$ in the document.

Convolution layer task: This task is to transform contextual aware based on text sentences document representation. The text sentences document given by movie consumer. Convolution was design to throw away the documents. The particular formulation is described according in equation 1.

$$c_i^j = f(w_c^j * D(c_i : (i+ws-1)) + b_c^j)$$

(1)

Where contextual representation given by: $c_i^j \in \mathbb{R}$ is extracted by $j$th shared weight $w_c^j \in \mathbb{R}^{p \times ws}$ and windows size $ws$ acquire by the count of surrounding words shown formulation on equation 1. This convolution require activation function, we consider involve linear unit (ReLU) among another activation function approach such sigmoid and tanh. The next stage, to develop contextual feature vector denotes $c^j \in \mathbb{R}^{l ws+1}$ of document with $W_c^j$ was built by equation 2.

$$C^j = [c_1^j, c_2^j, c_3^j, c_4^j, c_5^j, ..., c_{l ws+1}^j]$$

(2)

Thus, the design for multiple contextual feature was given by equation 3.

$$n_c \text{ of } w_c \text{ (i.e., } w_c^j \text{ where } j=1, 2, ..., n_c)$$

(3)

Pooling Layer Task: The section has responsible not just concentrate representative characteristic from convolution section, but they also distribute with varying distance of text sentences documents trough pooling production that designed a regular distance characteristic vector. Since the convolution process has generated, a text sentences document has been deputized given $n_c$ contextual characteristic length vectors, where every contextual characteristic vector has varying distance (that is, $l - ws + 1$ contextual feature). Nevertheless, such like deputation compels dual issue: 1) there beyond number contextual $c$, in which highly contextual characteristic could not support improve the ability of the model; 2) the distance of the contextual characteristic vector varied, whose made this complicated to build the according the process. Hence, max-pooling was employed to make smaller the deputation of text sentences document into a variable $n_c$ permanent distance vector by extracting simply the optimum contextual characteristic come from every contextual characteristic vector as equation 4.

$$df = [\max(c_1^j), \max(c_2^j), ..., \max(c_{n_c})]$$

(4)
where \( c^j \) is a contextual feature vector of length \( l - ws + 1 \) extracted by \( j \)th shared weight \( w^c_j \).

**Output Layer task:** The task of output section, representation-layer characteristic acquires from the foregoing process that has been transform into a particular responsible. Thus, \( df \) was forecasted use \( k \)-dimensional plot of product latent factor and user latent factors for the suggestion work, to acquire a text sentences document latent factor, we consider non-linier projection by tanh, detail formulation given by equation 5.

\[
s = \tanh(W_{f1}\left\{ \tanh(W_{f1}d_f + b_{f1}) + b_{f2} \right\}) + b_{f2}
\]

(5)

where \( w_{f1} \in \mathbb{R}^{l \times s} \), \( w_{f2} \in \mathbb{R}^{s \times l} \) are projection matrices, and \( b_{f1} \in \mathbb{R}^l \), \( b_{f2} \in \mathbb{R}^s \) is a bias vector for \( w_{f1} \), \( w_{f2} \) with \( s \in \mathbb{R}^s \). After that CNN construction became purpose whome bring a crude document to usage and revert a latent vector for every document became output result as shown in equation 6.

\[
s_j = dcnn(W, Y_j)
\]

(6)

whereabout \( W \) symbolize of quality and bias variables to avoid clutter, \( Y_j \) symbolize crude document for item \( j \), and \( s_j \) symbolize a document’s latent vector for item \( j \).

### 3.2 Probabilistic Matrix Factorization

The idea inspired by Dong [15] and we modified aim to increase contextual also level of scalability, we consider to incorporate (PMF) method to incorporated our new in convolutional neural network. Expected, we belong \( N \) representation of users and \( M \) representation of items, after that acquire rating are suppose by \( R \in \mathbb{R}^{N \times M} \) matrix, Then, our purpose is to discover user and item latent approach \((U \in \mathbb{R}^{N \times M} \text{ and } V \in \mathbb{R}^{N \times M})\) which generate \((U^T V)\) reconstruct the rating matrix \( R \). Referring to PMF approach, the model distribution through predict rating is denotes by equation 7.
\begin{align}
\rho(R | U, V, \sigma^2) &= \sum_{i} \sum_{j} N(v_{ij} | u_i^T v_j, \sigma^2 I_g) \\
\end{align}

(7)

The explanation as equation 8, \( N(x | \mu, \sigma^2) \) is representation of probability solidity purpose of Gaussian normal distribution with mean \( \mu \) and variance \( \sigma^2 \), and \( I_g \) is an indicator function. To expect user latent model, we use traditional priori, a zero-mean spherical Gaussian prior on user latent model use variance \( \sigma^2_U \).

\begin{align}
\rho(U | \sigma^2_U) &= \sum_{i} N(u_i | 0, \sigma^2_U I) \\
\end{align}

(8)

It is very different way over traditional item latent factor base PMF, we consider three variable 1, inside quality \( W \) in our strategy, 2. \( X_j \) deputation the document of item \( j \), 3. Epsilon varying as Gaussian noise. Consider optimizing the product latent representation to predict the rating, after that the indicate as equation 9 to generate final item latent model:

\begin{align}
v_j &= cmn(W, X_j) + \sum_{i} \sum_{j} I_g N(0, \sigma^2_I) \\
\end{align}

(9)

Where about quality \( w_k \) in \( W \), we assume zero-mean spherical Gaussian prior, the really famous by given prior as shown in equation 10.

\begin{align}
\rho(W | \sigma^2_W) &= \sum_{k} N(w_k | 0, \sigma^2_W) \\
\end{align}

(10)

Where, the allocation item latent model indicates as equation 11.

\begin{align}
\rho(W, X, \sigma^2) &= \sum_{i} \sum_{j} N(v_{ij} | cmn(W, X), \sigma^2_I) \\
\end{align}

(11)

Explanation of product latent vector acquired of CNN strategy was employed the mean of Gaussian distribution noise of item as established as the variance of Gaussian distribution to incorporate between Convolution and Probabilistic to produced including rating and text sentences description.

3.3. Measure Metric RMSE

RMSE is a method to evaluation metric that frequent used to evaluate of the distinction among grade (sample and population grade) estimated by an approach or predictor and the grade really inspected. The RMSE reflected the sample standard deviation of the dissimilar among estimated grade and inspected grade. Root Mean Squared Error (RMSE) is might the most famous metric used in evaluating accuracy of predicting rating

3.4. Dataset used

To show the robustness of the approach of belonging of us, in form of rating estimation, we applied 3 truly industry datasets acquired come from MovieLens [16] and Amazon information video (AIV). Including datasets information is consumer’ explicit ratings for products on rating scale of 1 to 5. Amazon information video including opinion for products as item characteristic documents. Because of MovieLens dataset does not contain item characteristic documents, we consider to generate the documents use corresponding items from IMDB server.

Table 2. Dataset specification

| Datasets | Number of users | Number of movies | Number of rating | sparse | Additional information | Release |
|----------|----------------|------------------|-----------------|-------|------------------------|--------|
| ML100k   | 943            | 1.682            | 100,000         | 6.3%  | demographic            | 4/1998 |
| ML1M     | 6.040          | 3.900            | 1,000,209       | 4.2%  |                        | 2/2003 |
### 3.5. Library and Tools

Our experiment involves several tools include software and hardware. There are several tools and software that involve includes Python with some libraries are included, for instance tensor flow for CNN learning application, Nvidia GeForce GTX 1001 to implementing convolutional neural networks help by pentium Xeon 2.4 Ghz.

**Table 3. list tool and library**

| No. | Device/tools/library       | Specification                          |
|-----|---------------------------|----------------------------------------|
| 1   | Processor                 | Xeon Quad core, 2.4 Ghz                |
| 2   | Memory                    | 16 Gb                                  |
| 3   | GPU                       | Nvidia GeForce GTX 1001                |
| 4   | Tensor Flow               | Deep learning library                  |
| 5   | Keras                     | Deep learning library                  |
| 6   | Anaconda                  | Python web interface                   |
| 7   | Python                    | Core programming tools                 |
| 8   | Scikit-learn              | Support library                        |
| 9   | Surface                   | Recommender system data analytic and comparison method tools |

### 4. Result and Discussion

According our experiment result on tables 3 mean of evaluation metric approach applied by RMSE of rating prediction approach based on Probabilistic Matrix Factorization (PMF), a model that implemented Collaborative Topic Regression (CTR) model, state of the art in rating prediction based on Collaborative Deep Learning (CDL) and Convolutional Probabilistic Matrix Factorization (ConMF) and our approach model with different percentages of testing data on four datasets. First, we can observe from the result that ConMF is the best performance over all approach model in item side information in the term of text sentences document of product. The bear of the meaning for the effectiveness of hybridization auxiliary information. Our approach involving variant of Convolutional Neural network dubbed DCNN stay outperform over previous work based on traditional sentiment analysis approach uses bag of word (BOW). It proven when increase increasing understanding for text sentences description going to upgrading the level of accuracy. According on table 3, this is show that our approach can be work normally such another approach PMF, CDL, CTR, ConMF. The increasing of accuracy level grows up constantly according ratio training and testing composition ration. Comparison result over state of the art in recommendation, involving additional information include text sentences description about product still out perform over juts involving user and rating such PMF, Traditional approach to extract product review using bag of word and word order such as CDL and CTR are loses against text sentences product description that applied deep learning variant in this technique involve convolutional neural network (ConvMF)and dynamic convolutional neural network (DCNN). However, ConvMF by Dong [15] stay out perform compare over all of them method include our approach by dynamic convolutional neural network (DCNN). According our test use similar parameter, device utility, library that used by ConvMF. Indeed, our approach out perform in time consuming over ConvMF. According our comparison in time to compute, DCNN need 50 hours, 21 minutes, 48 second, ConvMF need a lot of time to compute exactly 168 hours, 50 minutes, 02 second.
Table 4. Experiment result and comparison

| Technical Approach | Ratio of Training Set (ML-1M) |
|--------------------|-------------------------------|
|                    | 20%  | 30%  | 40%  | 50%  | 60%  | 70%  | 80%  |
| PMF                | 1.0168 | 0.9711 | 0.9497 | 0.9354 | 0.9197 | 0.9083 | 0.8971 |
| CTR                | 1.0124 | 0.9685 | 0.9481 | 0.9337 | 0.9194 | 0.9089 | 0.8969 |
| CDL                | 1.0044 | 0.9639 | 0.9377 | 0.9211 | 0.9068 | 0.8970 | 0.8879 |
| CNN-G              | 0.9885 | 0.9490 | 0.9223 | 0.9052 | 0.8891 | 0.8825 | 0.8698 |
| ConvMF             | 0.9745 | 0.9330 | 0.9063 | 0.8897 | 0.8726 | 0.8676 | 0.8531 |

5. Conclusion

Based on our experiment, involving side information in the term of product description on movie field commerce business. Our proposed model can handle to increase accuracy level involving product description. Applied of product description have proven could be increase the level of accuracy to predict the rating compared to several models that already existed before. The involvement of convolutional neural networks to extract product description is proven to increase a deeper understanding of the product description, so that the results we can see are able to increase the accuracy of rating prediction. In this research we involve a contextual product description using dynamic convolutional neural network and involves a standard probabilistic matrix factorization. Our method possibly to integrated with another standard matrix factorization such as based on SVD, NMF, SVD++. The hybridization approach may have possibly to increase performance in rating prediction.

According our experiment result, we expect there is the potential to further research to improve scalability and performance of convolutional approach by involving derivatives from convolutional, such as dilated convolutional neural network, in which is likely to produce accuracy and decrease time savings for computation. Our proposed also suitable for model that needed accuracy but have no hardware resource to implemented ConvMF such no need high hardware computer, memory also GPU with high specification. That mean limitation hardware specification is still compromised with our model.

6. References

[1] G. R and G. J, “Get the right mix of bricks and clicks,” Harv. Bus. Rev., 2000.
[2] C. A. Gomez-uribe and N. Hunt, “The Netflix Recommender System: Algorithms, Business Value,” vol. 6, no. 4, 2015.
[3] J. Davidson and B. Liebald, “The YouTube video recommendation system The YouTube Video Recommendation System,” no. August, 2014.
[4] J. Ben Schafer, J. Konstan, and J. Riedl, “E-commerce recommendation applications,” Appl. Data Min. to Electron. ..., pp. 115–153, 2001.
[5] Hanafi, N. Suryana, and A. Sammad, “AN UNDERSTANDING AND APPROACH SOLUTION FOR COLD START PROBLEM ASSOCIATED WITH RECOMMENDER SYSTEM: A LITERATURE REVIEW,” J. Theor. Appl. Inf. Technol., vol. 96, no. 09, 2018.
[6] F. Ricci, L. Rokach, and B. Shapira, Recommender Systems Handbook, vol. 54. 2015.
[7] A. van den Oord, S. Dieleman, and B. Schrauwen, “Deep Content-Based Music Recommendation,” pp. 2643–2651, 2013.
[8] S. Jaradat, “Deep Cross-Domain Fashion Recommendation,” Proc. Elev. ACM Conf. Recomm. Syst. - RecSys ’17, pp. 407–410, 2017.
[9] J. Wang, “Predicting Yelp Star Ratings Based on Text,” Standfor Univ. Univ., pp. 1–5.
[10] X. Wang and Y. Wang, “Improving Content-based and Hybrid Music Recommendation Using Deep Learning,” Proc. 22Nd ACM Int. Conf. Multimed., pp. 627–636, 2014.
[11] K. Park, J. Lee, and J. Choi, “Deep Neural Networks for News Recommendations,” CIKM’17, Novemb. 6-10, 2017, Singapore, 2017.
[12] L. Chen, G. Chen, and F. Wang, “Recommender systems based on user reviews: the state of the art Recommender Systems Based on User Reviews: The State of the Art,” no. April, 2015.

[13] L. Qu, G. Ifrim, and G. Weikum, “The Bag-of-Opinions Method for Review Rating Prediction from Sparse Text Patterns,” no. August, pp. 913–921, 2010.

[14] A. Severyn and A. Moschitti, “Twitter Sentiment Analysis with Deep Convolutional Neural Networks,” ACM, pp. 959–962, 2015.

[15] D. Kim, C. Park, J. Oh, S. Lee, and H. Yu, “Convolutional Matrix Factorization for Document Context-Aware Recommendation,” Proc. 10th ACM Conf. Recomm. Syst. - RecSys ’16, pp. 233–240, 2016.

[16] F. M. Harper and J. A. Konstan, “The MovieLens Datasets: History and Context,” ACM Trans. Interact. Intell. Syst., vol. 5, no. 4, p. 19:1–19:19, 2015.