ACCELERATED - GENERIC GRADIENT DESCENT FOR E-COMMERCE RECOMMENDER SYSTEMS

A Suresh

Research Scholar, Department of Computer Science and Engineering,
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology Avadi,
Chennai, Tamil Nadu. 600062,India.
vtd592@veltech.edu.in

Dr. M J Carmel Mary Belinda
Professor, Department of Computer Science and Engineering,
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology Avadi,
Chennai, Tamil Nadu. 600062,India.
carmelbelinda@veltech.edu.in

Abstract

The online product reviews play an important role in the field of e-commerce as rating of the product infers the preference for customers and they rely on them while making purchases. The main aim of providing an efficient recommender system is to gain useful information for user based on quality, worth and quantity of the product. The reviews given by the customers are used to extract the features of products that evaluate the similarity content reviewed by another customer. Therefore, the reviews written by the distinct persons would be similar, but with different words of expression shows difficulty during review differentiation. The developed model performed mining customer review for a set of manufactured goods and their characteristics are used as product features. The proposed Gated Recurrent Unit with Accelerated - Generic Gradient Descent (GRU- AGGD) technique extracts the relevant and sufficient features from the review to overcome the existing optimization problems. These relevant features are the summary of the product used to predict the rank using the users-attributes and users-items matrix and these two matrixes give rise to Non-Negative Matrix Factorization model to provide recommendations to user for the product. The experimental results of product recommender system are evaluated using the proposed GRU-AGGD that shows the improvement in terms of performance measures. The average MSE and RMSE values obtained for the proposed GRU-AGGD technique are 0.84 and 0.91 respectively, which are higher when compared to the Bayesian recommender models.

Keywords: Customer review, E-commerce, Non-Negative Matrix Factorization, Online Product, Ratings, Recommender Systems, Users-Attributes, Users-Items Matrix.

1. Introduction

Nowadays, the e-commerce has emerged a great demand by helping the consumers to receive desire product on their place and have improved the information load for products. In India, 329.1 million people are dependent upon buying the goods with help of online services. In online shopping websites, the reviews given by the previous users for a particular product play a major role in the decision making of the new users of the same product. However, providing item recommendations would not work due to a large number of product attributes. These product attributes have their own importance in terms of time because as the days pass the perspectives on the product will also be changed by the consumer.

The selection of desired product will be in customer hand and once the order is placed, the customers are going to get the products on to their own place wherever they desire to receive. As the time required by the customers to select a product is high due to their ambiguous state in making a decision. The recommender systems are developed to ease the selection of the desired product and helps automatic recommendation process for customers [1].

The recommender systems usually filter the information to predict the preference of a user for an item that provides suggestions for customer individually [2]. The reviews and stars obtained for each product will be

A Suresh et al. / Indian Journal of Computer Science and Engineering (IJCSE)

analyzed by the customers before purchasing the products from the online platform, which contributes a high impact on the product sales [3]. However, the irrelevant or repeated information in the system often creates problems for the users, so the automatic recommendation systems are used to overcome the excess information problems [4]. The limited availability of content leads to many problems, including overspecialization [3]. The traditional recommendation model was analyzed the rankings based on their ratings that are similar to infer user preferences [5].

The improved hybrid recommendation models combine the rating matrix, product matrix and user-review matrix for product ranking analysis. The developed hybrid recommendation models facilitated one-to-one correspondence between the latent rating dimensions and the latent topics which are extracted from reviews [6]. The researchers suggested numerous methodologies that evaluated reviews thereby improved the ratings of recommendations for the desired product. The various methodologies were used for analyzing the reviews based on sentimental analysis that recommended the products for customers based on their reviews [7,8].

A piece of text will predict whether the review about the product is positive or negative, opposed or supported etc. In addition, the semantic analysis was employed to know the service quality provided by E-commerce [9]. The developed hybrid recommendation model performed customer review mining that extracted characteristics for a set of products and analyzed the opinions mentioned in the form of reviews. The items are represented based on the subjective attributes and item is selected based on their subjective attributes. Features that represent user preferences in a better way are not taken into account. Deep Learning systems are now increasingly being applied to structured datasets and form an area of active research [10]. The objectives of the model are to identify the user's point of interest, recommending products/services based on the user's latent interests [11]. In this situation, recommender systems have emerged as an effective mechanism to provide personalized recommendation services, which can effectively alleviate the information overload problem [12, 13].

The gated recurrent unit methods are used to extract the characteristics to summarize them into an abstract using self-attention layer. The set of attributes is constructed by the users-attributes matrix, but is very sparse as it has less attributes in the review. The problem can be severe in context-aware recommenders as they use multidimensional vectors, where it becomes very difficult to provide accurate recommendations for very few rated items. Thus, Sparsity is one of the challenging faces during developing recommender systems. The sparse rating in recommendation systems determines the accurate predictions about items. The proposed GRU-AGGD technique is used to extract the summary for recommender system and to overcome the optimization problem. In addition to this, users-items matrix and users-attributes product gives rise to non-negative matrix factorization model that predicts ratings to provide recommendations.

The structure of the research paper is arranged as follows. Section 2 describes the literature review of the existing methodologies. Section 3 explained about the proposed GRU-AGGD technique in detail and Section 4 describes the results of the proposed GRU-AGGD technique in terms of quantitative analysis and comparative analysis. The conclusion and future work of this research is given Section 5.

2. Literature Review

Various methodologies were used to analyze the reviews based on sentimental analysis that recommends the products for customers based on their reviews. This section shows the advantage and limitation of the existing models used in recommendation system.

Paradarami, et al. 2017 [14] presented hybrid recommended system using Artificial Neural network (ANN) framework that utilized the business reviews and content based features. The generated features used for the prediction model combines the user and business for ranking products. The developed model used Yelp Academic Dataset by performing the computational experiments. The collaborative and the set of content features were given as the input for the neural network model to reduce the misclassification errors and to improve the optimization using a stochastic gradient descent algorithm. The advantage of the developed model was that the parameters required was saved in the memory, thus lowering the memory loss and improving scalability for the products. However, the developed model considered only star rating to predict the results and the review texts were not considered for ranking the products.

Wu *et al.* 2017 [15] developed a DNN with Dual-Regularized Matrix Factorization (DRMF) for Recommender systems. The developed model used Apps for Android (AA), the Kindle Store (KS) and Amazon Instant Video (AIV) datasets to conduct the research. The developed model used the DRMF with DNN that adopted multilayered Network model for stacking. The Recurrent Neural Network (RNN) generated an independent distribution systems for the contents that consisted of users, items attributes by generating matrix factorization. The DRMF adopt a multilayered Neural Network model that stacked CNN and gated RNN generated content to represent the items and user matrixes. The experimental results obtained from the

developed model showed that the dual-way regularization strategy improved the matrix factorization by achieving the best accuracy. However, the developed model needs to examine for training large dataset to obtain a complete framework of DRMF.

Sohail *et al.* 2018 [16] developed a comprehensive approach for recommended systems based on the feedback obtained from the online shopping. The developed model used local dataset for evaluating the rank aggregation algorithm. This Rank aggregation algorithm was used to obtain a final ranking when compared with the base of ranking evaluation. The developed model evaluated the fake or biased feedbacks to state that the recommender system was genuine. The comparative analysis among the existing approach shows a great superiority in aforementioned approach with respect to various parameters. However, the model failed to provide a reliable online shopping scheme for assessing the recommender system.

Nilashi *et al.* 2019 [17] analyzed Travelers' Online reviews on Social networking sites using Fuzzy Logic Approach. The developed model used Trip Advisor dataset to evaluate the recommendation system for travelers in the domain of tourism. The developed model used crisp based multi-criteria algorithms to determinine the similarities in the ratings provided by the travelers. The recommendation model used fuzzy logic approach for clustering and prediction of product rate. The developed method significantly improved the quality of recommendations in the domain of tourism. However, the developed model needs to be improved in the scalability issue of multi-criteria based algorithms in the domain of tourism.

Lee *et al.* 2019 [18] developed a hierarchical deep neural network model to determine the product ranking for the online products based on the customer reviews. The developed model used as a large dataset called Amazon Product Data to evaluate the ranking of the products for specific categories. The system included hierarchical attention network from where the additional features were used for ranking the product and recommending the users. The developed model effectively captured the underlying features of the online products in a hierarchical manner that were represented by attention weights. The advantage of the developed model was the retailers expected the future sales of the product based on the online reviews. However, the developed model focused only on overall user preference of particular product apart from interest of people.

Mingming Jiang et al. [19] developed a Bayesian Recommender Model for reviewing profiles and user rating exploited for Amazon dataset. The existing recommender systems took rating scores that granted and discarded the wealth of information accompanied reviews. The developed model served as a "recommender" task with very sparse datasets. However, the reviews of each user or item were not fully exploited for the recommendation. In addition, due to the hardware limitation up to 50,000 users and up to 5000 items for very large dataset, the system performance was reduced.

Liu et al. [20] developed an Attention-Based Adaptive Memory Network (AAMN) for Recommendation with Review and Rating for leverage historical reviews and ratings systemically for amazon dataset. The developed AAMN learned the static features from distinct historical records by modeled the adaptive features obtained from the items and users. The static features were developed adaptively that extracts the semantic information from reviews which was not characterized based on the global view distinguished an importance of distinct reviews. However, the developed model had explored moderate efficiency during the factorization for each machine layer model thereby impacted poor on the system performance.

From the literature survey of existing methods, it was analysed that generally the existing researches limited the availability of content that lead to problems including overspecialization. The overspecialization problem created the complexity during recommendation, as the subjective attributes were important for selecting an item. Features that represent user preferences in a better way are not taken into account. The reviews given by the customers were used to extract the features of products that evaluates the similarity content reviewed by another customer. Therefore, the reviews written from distinct persons would be similar, but with different words of expression showed difficulty during review differentiation. The sparse rating in systems makes difficult to make accurate predictions about items. The scalability becomes difficult for a typical recommender to process such large-scale data. The proposed GRU-AGGD technique is used to extract the summary for recommender system and to overcome the optimization problem. In addition, users-items matrix and users-attributes product gives rise to non-negative matrix factorization model that predicts ratings to provide recommendations.

3. Proposed methodology

The flow chart of the proposed GRU- AGGD methodology is shown in the figure 1. The block diagram consists of Amazon data's customer review from where the text forms of the main ideas are extracted by the process of Abstractive summarization. The summary will be extracted and thereby followed to find out the polarity in the text. By Constructing the two matrices such as users attribute and users-items matrix the polarities are

calculated. The proposed NNMF finally to predict the recommendations for items based on the amazon reviews that overcomes the problem of sparsity and scalability.
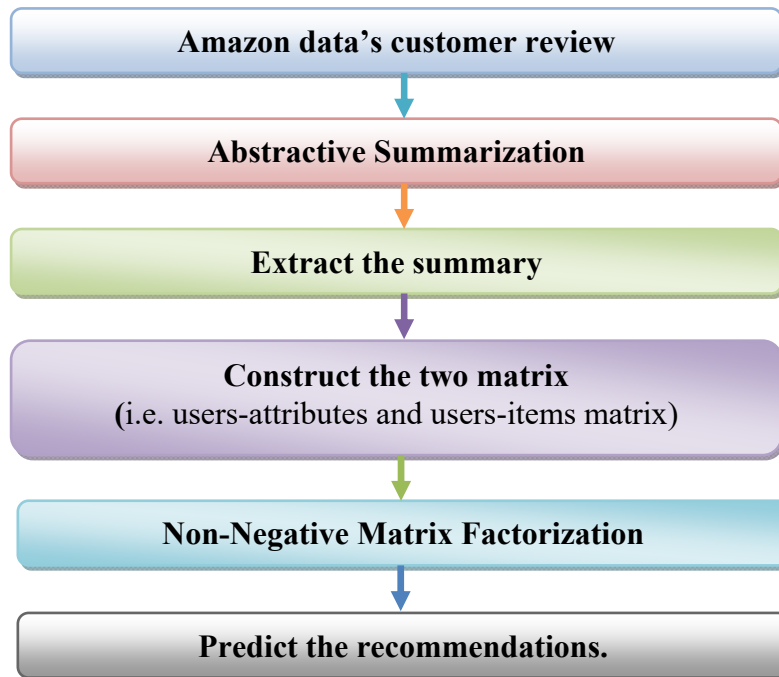
```
┌─────────────────────────────────────────┐
│      Amazon data's customer review        │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│         Abstractive Summarization         │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│            Extract the summary            │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│         Construct the two matrix          │
│ (i.e. users-attributes and users-items matrix) │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│      Non-Negative Matrix Factorization    │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│        Predict the recommendations.       │
└─────────────────────────────────────────┘
```

Figure 1: Flow chart of the proposed methodology

3.1. Data Collection

3.1.1.   Amazon Review dataset:

The proposed method uses Amazon Review dataset for experimentationand it contains 142.8 million reviews on product between the months of May 1996 and July 2014 in addition with item metadata and users profile. This metadata consisted of reviewer ID, Product ID, reviewer's name, ratings for the product, reviewer's text to the product, helpfulness rating, review summary and review time. The dataset consists of the product name, price for the product, the products related to the selected one, information about sales rank, name of the brand, categories for the list of products. The customers have shared their experience of using the product by expressing in the form of comments. The Amazon customers have contributed over a hundred million reviews for expressing their opinions and have shared their experiences about the products they purchased. From these reviews, about hundred million in number described the experiences based on the products on the website and is involved in studies in the field of Natural Language Processing (NLP), Machine Learning and Information Retrieval. Accordingly, the Amazon data are used in further research with respect to multiple disciplines relative to the understanding of product experiences from the customer. This dataset was specifically constructed for representing customer evaluations, variations in the perception, opinions across geographical regions. This research analyzed the product recommendation, Amazon Instant Video, Cell Phones & Accessories, Kindle store reviews for evaluating the results.

3.2. Abstractive Summarization

Initially, the reviews from the datasets are considered to perform abstractive summarization that extracts new phrases using words or possibly rephrasing that do not belong to the original text. Abstractive methods choose words based on semantic understanding where the chosen words are not originated from the source documents. As the extraction of natural abstractive summary perfectly is harder, a model is required to understand the document for expressing in short possibilities using the new phrases or words. The abstractive summarization performs three functions such as generalization, paraphrasing and integrating the knowledge that present in the real-world consequences. In order to extract the abstractive summary, an open source neural network library called Keras is used.

3.2.1. Keras model for summary extraction

Keras is a high level neural network model that has ability to run complex modules using Python. In Keras model, GRU in addition with Accelerated Generic Adaptive Method is used as the proposed model. The Gated Recurrent Unit (GRU) model is used to develop the abstractive summarization and also performs gating mechanism in recurrent neural networks. The GRU is like an LSTM with forget gate, but the performance of GRUs exhibit even better performance on certain smaller datasets than LSTM. The GRU primarily controls the gating network signals based on the previous memory. The signal obtained from the previous memory is used as an input that updated with an activation state result to produce the current state. The present gating state has the set of weights that are adapted and updated to learning phase. The LSTM unit includes an input gate, cell, and an output gate and forget gate and the forget gate is used for facilitating remembering values in the cell based on the arbitrary time. The gates that facilitate the flow of information in and out of the cell control the flow of information. The block diagram of the LSTM model is shown in the figure 2. The output generated by the LSTMs has the ability to support the automated feature learning by extracting only the relevant features for the classification.
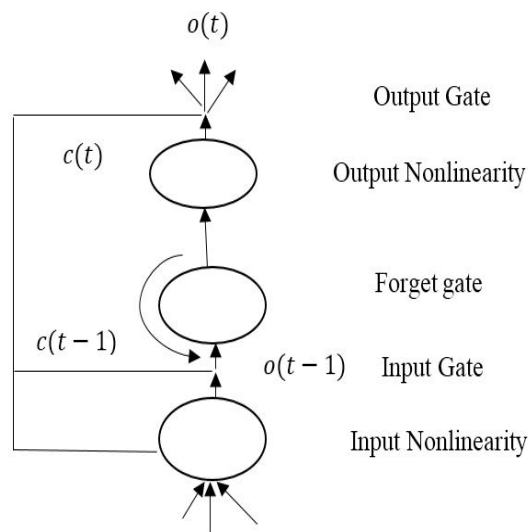

Figure 2: LSTM model

There are several variations on the full gated unit as the data uses the previous hidden state and performs various bias combinations. The combinations generated are simplified into a form called minimal gated unit. The hidden state of an LSTM cell can be calculated using the equation 5. The layers of the LSTM are calculated using the equations from 1 to 4.

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \qquad (1)$$

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \qquad (2)$$

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \qquad (3)$$

$$C_{t=} \sigma(f_t \times C_{t-1} + i_t \times \hat{C}_t) \qquad (4)$$

$$h_t = \tan h(C_t) \times o_t \qquad (4)$$

Here,

i $\rightarrow$ input gate

f $\rightarrow$ forget gate

o $\rightarrow$ output gate

W$\rightarrow$recurrent connection at the previous hidden layer and current hidden layer,

U$\rightarrow$ weight matrix connecting the inputs to the current hidden layer.

The Keras implementation of LSTM updatesinvolved parameter information thereby pertains or retains the state of the overall network. Thus, all information regarding the current input and the previous hidden states is reflected in the latest state variable. Moreover, all the adaptive parameters update the involved components of the internal state of the system. In addition to this, the neural network process is also used to extract the summary based on the two computational procedures namely forward process and back propagation. In the

forward process, all input signals are processed in the forward direction in the activated network layers i.e.., input to output. The error value is calculated to estimate the connection weight and bias term. For the better understanding of the Artificial Neural Network, the three-layered feed-forwarded network is discussed in this section. The input signals are initially produced with a training pattern that is presented to the network and then multiplied by next layer's connection weight values. Each bias term is added with corresponding weighted signal and then imported to a suitable activation function to develop the outputs of hidden layer. The outputs are weighted based on the error value and then summed up in the output neuron. The input and output layers are explained in the following section based on the aforementioned process.

Input Unit:

The input unit is mathematically expressed in Equation (5)

$$o_i^1 = y \tag{5}$$

Hidden Units:

The hidden unit equation is expressed in Equation (6) and (7)

$$o_i^2 = f(net_i), i = 1, \dots, I \tag{6}$$

$$net_i = y.w_i^1 + b_i, \tag{7}$$

Where,

f is the sigmoid activation function.

Output Unit:

The Output unit equation is expressed in Equation (8)

$$N(y) = \sum_{i=1}^{I}(w_i^2.o_i^2) = \sum_{i=1}^{I}(w_i^2.f(w_i^1.y + b_i)) \tag{8}$$

The designed neural architecture is a pro-type model and some minor changes are needed to overcome the fractional problem. The modified Adam optimization algorithm is given as follows,

Algorithm 1: Adam optimization algorithm

Input: $x_1 \epsilon F$, step size $(\alpha_t > 0)_{t=1}^{T}$, sequence of functions $(\varphi_t, \psi_t)_{t=1}^{T}$

$\qquad t \leftarrow 1$

$\qquad S \leftarrow threshold$

Repeat

$\qquad g_t = \nabla f_t(x_t)$

$\qquad V_t = \psi_t(g_1, \dots, g_t)$

if $g_t.m_{t-1} > 0$ and $|m_{t-1} - g_t| > S$ **then**

$\qquad gm_t = g_t + m_t$

$\qquad m_t = \beta_1.m_{t-1} + (1-\beta_1)_{gm_t}$

else

$\qquad m_t = \beta_1.m_{t-1} + (1-\beta_1)_{g_t}$

$\qquad \hat{x}_t + 1 = x_t - \alpha_t \, m_t V_t^{-\frac{1}{2}}$

$\qquad \hat{x}_t + 1 = \Pi_{F,\sqrt{V_t}}(\hat{x}_t + 1)$

$\qquad t \leftarrow t + 1$

until t > T

The Adam optimization algorithm is modified by adding the gradient values, thereby generate an improvised Adam optimization algorithm known as Accelerated-Generic Gradient Descent Method. The change in the Accelerated Generic Gradient Descent optimizer does not alter its convergence since the quantity essentially measures the change in the inverse of learning rate with respect to time. Thus, the integrated GRU and AGGD algorithm results with an improved performance.

Algorithm 2: Accelerated-Generic Gradient Descent Method

Input: $x_1 \in F$, step size $(\alpha > 0)$

$$t \leftarrow 1$$
$$S \leftarrow threshold$$

Repeat

$$g_t = \nabla f_t(x_t)$$
$$new_{-g} = g_t$$

If $g_t \cdot g_{t-1} > 0$ and $|g_{t-1} - g_t| > S$

then

$$new_{-g} = g_t + g_{t-1}$$
$$x_{t+1} = x_t - \alpha_t . new_{-g}$$
$$t \leftarrow t + 1$$

until $t > T$

The summary from the abstract is obtained and processed to extract the product and user attributes. The proposed GRU –AGGD overcomes the optimization problem as the related features are only extracted from the research.

3.3 Construction of users-attributes and users-items matrix

The user-attributes matrix and user-items matrix are used to extract the attributes of user and items that recommends the user to buy the product based on the obtained ranking. The user-attributes matrix and the item matrix are constructed to extract the attributes from user pointed reviews expressed in the form of binary format. One of the challenges with the matrix are sparsity, where such an issue is overcome by finalizing the set of product characteristics with proper analysis of customer review during mining. The set of attributes is constructed for user-attribute matrix. However, developing recommender system in the sparsity is challenging and creates sparsity problem. The user-items and user attributes matrix are used with Non-Negative Matrix Factorization (NNMF) to predict rating that provides recommendations. The ratings are used to predict the quality of the product using the aspects of item's and a user's information. The user-review model is constructed in the proposed GRU –AGGD technique to provide recommendations based on the ratings provided by the customers. The cross-domain task is accomplished in the model that transfers embedded word. The ratings are predicted in the user-items matrix to generate recommendations with the help of NNMF which acts an efficient filtering technique.

The NNMF approximation is a group of algorithms that performs multivariate analysis and linear algebra. The term V is the factorized matrix that is expressed into two matrices W and H having the property of non-negative elements. The non-negativity makes the results easier by defining two matrices W and H, where W is the user matrix and H is the item matrix. Let matrix V be the product of the matrices W and H. The equation for the NNMF is expressed as shown in equation (9)

$$V = WH \qquad (9)$$

The matrix multiplication is implemented by computing the column vectors V and the combinations of linear column vectors W are used as coefficients that forms the column vectors H. Each column of V is expressed in the equation 10

$$v_t = W h_t \qquad (10)$$

Where v_t is the ith column vector of Matrix V and h_t is the ith column vector of H matrix.

An advantage of using the proposed GRU –AGGD technique is that the learning process used will speed up when the predictive model deals with large datasets. The developed model reduces the deficiencies during sparsity for the attributes. Thus, the neural network enhances the characteristics obtained for the list of user to improve the quality of the recommendation system.

3.4 Predicting the recommendation ratings

The last step in recommendations is that predicts the rating using NNMF model. The rating estimation for each item with respect to the user underwent NNMF approach where the product of respective user performed item embedding. The dot product of user and item embedding is merged together to obtain the

product rating. As both of the users-deep features and users-items are used for employing NNMF, it is easy to perform ranking predictions for the items in generating the records of the recommendation as per the size of the list.

4. Results

The experimental setup and the system requirements used for the proposed GRU –AGGD method is shown in the table 1. An operating system of Windows 10 (64 bit) with i7-configured processor is used in the research work. The memory space of 16 GB RAM and the Anaconda Navigator tool in python 3.7 language is utilized for code writing.

SYSTEM SETUP	CONFIGURATION
Operating system	Windows 10 (64 bit)
Processor	intel core i7
RAM	16GB
Software	Anaconda Navigator 3.5.2.0 (64-bit) Python 3.7

Table 1: The experimental setup for the proposed GRU-AGGD method

4.1 Evaluation Metrics

- **Mean Absolute Error**

Mean Absolute Error (MAE) is the difference between prediction and actual observation where all individual differences have equal weight. The MAE is expressed in the equation 11.

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j| \qquad (11)$$

- **Mean Square Error**

Mean Square Error (MSE) is the measure of the average square errors obtained when the average squared difference between the actual and the estimated value is evaluated. The MSE is expressed as shown in the equation 12.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2 \qquad (12)$$

- **Root-Mean-Square Error**

The Root-Mean-Square Error (RMSE) measures the average magnitude error that follows the quadratic scoring rule. The RMSE is defined as the square root of the average squared differences between the actual and prediction observations The RMSE is expressed as shown in the equation 13

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2} \qquad (13)$$

Where,

y_j is the actual value

$\hat{y}_j$ is the predicted value

n is the number of observations

4.2 Quantitative analysis

The experimental result performs the training and testing process using Amazon review dataset. The experimentation used 80% of the dataset for training, 20% for testing. The amazon review dataset considered categories such as Amazon Instant Video, Cell Phones & amp; Accessories and Kindle stores for experimentation. This collected review from the datasets undergoes abstract summarization. Finally, each user's review is extracted based on matching set of words that gives rise to construct the user-attributes matrix and item-attribute matrix. The obtained feature parameters are applied for the NNMF and Neural Network. The results obtained for the proposed GRU –AGGD is tabulated as shown in the Table 2

Amazon Review dataset	Non-Negative Matrix Factorization (NNMF)		
	Performance measure		
	MAE	MSE	RMSE
Amazon Instant Video	0.65	0.85	0.92
Electronics	1.02	1.12	1.06
Kindle stores	0.52	0.56	0.75

Table 2: Performance measures obtained for the proposed NNMF

The quantitative analysis is performed for three categories from the amazon review dataset that includes Amazon instant video, Cell Phones and Accessories, and Kindle store. The least error value for MSE, MAE and RMSE are obtained for the Amazon Instant Video Category as 0.75, 1.55, and 1.25 respectively, whereas the cell phones category obtained the highest error values of 1.22, 2.67, and 1.63 respectively. The kindle store obtained moderated error values such as 0.83, 1.96 and 1.39 when compared to the Amazon Instant Video Kindle store categories. The comparison graph for the proposed NNMF for all the three categories is shown in the figure 3.
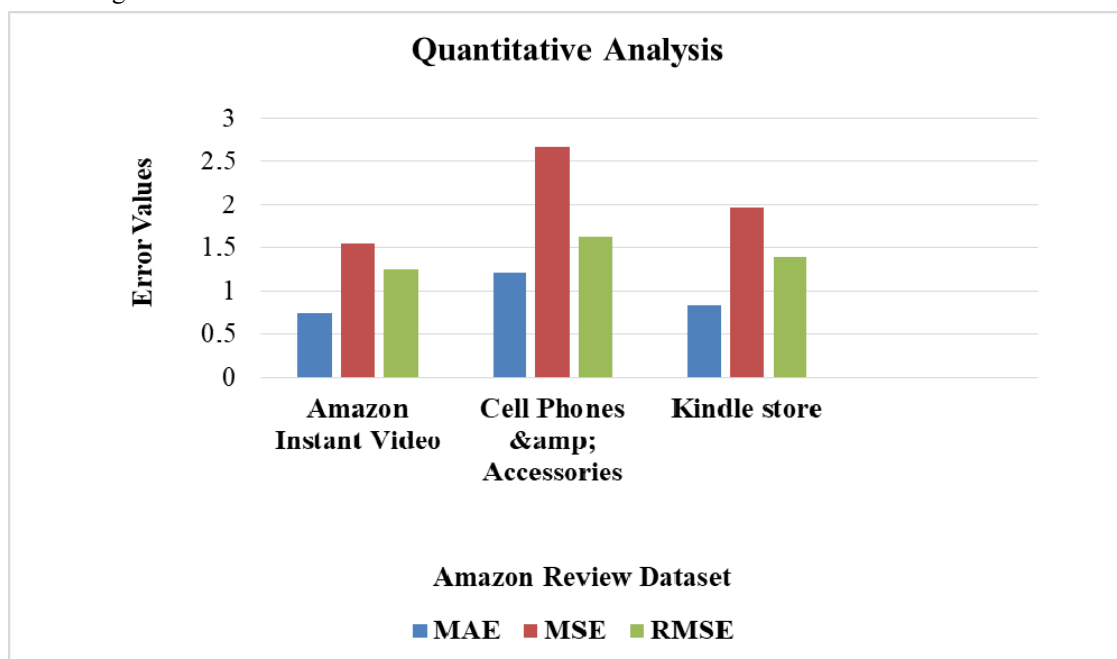

Figure 3: Quantitative representation of the proposed GRU –AGGD

4.3 Comparative Analysis

The comparative analysis is performed among the existing Matrix Factorization and KNN algorithm. The significant improvement in MAE is obtained in the proposed GRU –AGGD model better than the KNN algorithm. The KNN algorithm obtained least error value when compared to the existing Matrix Factorization method. Moreover, the improvement in the performance shows higher to a cell phone accessory categories and kindle stores, where the customers are discussed more about the product attributes.

	Matrix Factorization			KNN Algorithm			Non-Negative Matrix Factorization		
Performance measure	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE
Amazon Instant Video	2.31	8.97	2.99	1.17	3.11	1.76	0.65	0.85	0.92
Electronics	3.46	14.39	3.79	1.66	4.77	2.18	1.02	1.12	1.06
Kindle stores	3.75	16.39	4.05	1.11	2.89	1.69	0.52	0.56	0.75

Table 3: Comparison table obtained for the performance measures for the existing and proposed methodologies

The comparative analysis of various performance measures obtained from the existing and the proposed method is shown in table 3. The proposed method performs better for all of the product categories such as Amazon Instant Video, Kindle stores, Cell Phones and accessories. Compared to the Matrix Factorization model, the proposed NNMF method is capable of predicting the average improvement from 13.25 to 0.84 in terms of MSE averagely. Among the three cases, NNMF obtained better performance for the categories such as Cell Phones and accessories and Kindle stores. The proposed GRU–AGGD is compared with the existing KNN, which showed 1% of improvement in the performance with respect to the Cell Phone Accessories and Kindle stores categories. The same scenario is repeated among the HFT model and the proposed GRU –AGGD showed better improvement in error values. The proposed GRU –AGGD showed great improvement in the predictions up to 3.14% of these 2 cases and also showed better performance in terms of MAE. The proposed GRU –AGGD obtained MSE values of 0.85, 1.12 and 0.56 for the Amazon Instant Video, Cell Phones & Accessories, and Kindle stores categories. The RMSE values of 0.92, 1.06, and 0.75 for the Amazon Instant Video, Cell Phones & Accessories, and Kindle stores categories. Similarly, MAE values of 0.65, 1.02, and 0.523 for the Amazon Instant Video, Cell Phones & Accessories, and Kindle stores categories. The comparative analysis of the existing and the proposed GRU –AGGD is shown in the Figure 4.
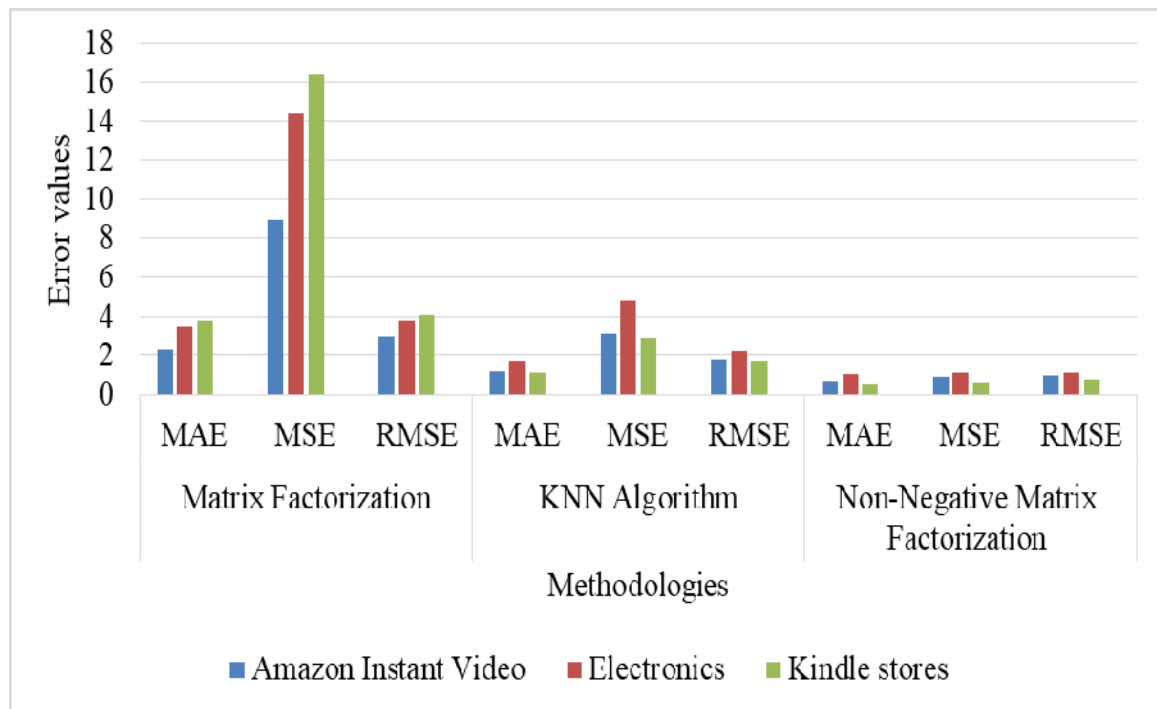
Figure 4: Graphical representation of the comparative Analysis for the proposed GRU –AGGD methodology and the existing methodologies.

The results obtained from the proposed model extracted the words based on the summarization process that made user-attributes more informative by providing the information. The relation between the user and the items are recommended based on the review ratings written by customers. Thus, the proposed GRU-AGGD model works better in the field of e-commerce thereby plays a crucial role in real systems to satisfy customer satisfaction. The comparative analysis of the proposed NNMF with the existing Matrix Factorization and Bayesian Recommender Model is shown in the table 4. The existing method

Categories	Matrix Factorization [2] (2019)	Bayesian Recommender Model [19] (2015)	Cross-Domain Deep Neural Network (CD-DNN) [20] (2020)	NNMF
Amazon Instant Video	1.437	1.028	0.948	0.85
Cell Phones and Accessories	2.337	2.217	1.209	1.12
Kindle store	1.66	1.6436	0.624	0.56

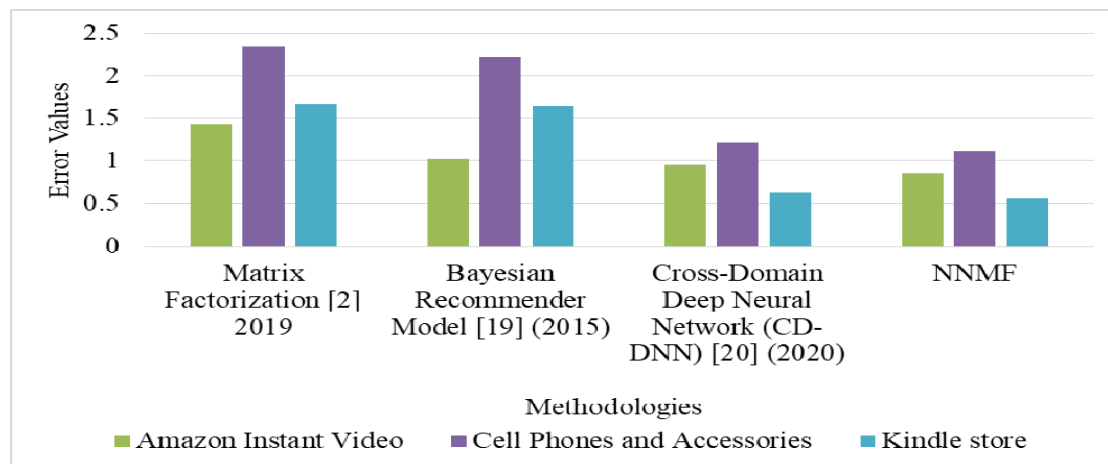Table 4: Comparative analysis of MSE of the existing and the proposed model

Figure 5: Comparison graph for MSE values for the proposed and the existing methods

The experimental results obtained from the existing models such as Matrix Factorization [2], Bayesian Recommender Model [19] and CD-DNN [20] are compared with the proposed NNMF model. The results comparison is performed for three main categories such as Amazon Instant Video, Cell phone &Accessories, and Kindle store taken from the Amazon Review Dataset. The proposed NNMF model shows lesser MSE error values when compared to the existing models. The proposed NNMF model showed improvement of MSE value approximately up to 0.97% when compared to the existing models. The comparison graph of MSE values for the proposed and the existing methods is graphically represented as shown in figure 5. The Amazon Instant Video category showed better MSE values when compared to other two categories. However, the proposed NNMF achieved slightly more MSE when compared with the existing Bayesian model.

5. Conclusion

The proposed GRU-AGGD is an efficient model for summarizing the customer reviews that provides recommender systems for E - commerce site. Initially, the reviews from the customers were collected from the Amazon Review dataset for the experimentation and these reviews were processed for abstractive summarization using Keras to extract summary of reviews. The proposed GRU-AGGD model was used to predict ratings to generate recommendations to the user. For estimating the rating of the items, the proposed GRU –AGGD approach used the product of the respective user and item embedding. To obtain the rating, the dot product of user and item embedding are merged together which will find the best product to recommend the user. The proposed GRU –AGGD model performed better when compared with the existing Bayesian Recommender Model and MF methods. The NNMF outperforms MSE values up to 0.97 error values better for GRU-AGGD by using Amazon dataset. In the future, proposed GRU-AGGD method can be extended to other real time datasets such as Trip Advisor, Better Business Bureau, Facebook etc., using the improvised Adam optimizer recommender systems for customers.

REFERENCES

[1]   Jiang, L., Cheng, Y., Yang, L., Li, J., Yan, H. and Wang, X., 2019. A trust-based collaborative filtering algorithm for E-commerce recommendation system. *Journal of Ambient Intelligence and Humanized Computing*, *10*(8), pp.3023-3034.
[2]   Shoja, B.M. and Tabrizi, N., 2019. Customer Reviews Analysis with Deep Neural Networks for E-Commerce Recommender Systems. *IEEE Access*, *7*, pp.119121-119130.
[3]   Bai, T., Zhao, W.X., He, Y., Nie, J.Y. and Wen, J.R., 2018. Characterizing and predicting early reviewers for effective product marketing on e-commerce websites. *IEEE Transactions on Knowledge and Data Engineering*, *30*(12), pp.2271-2284.
[4]   Jamshidi, S., Rejaie, R. and Li, J., 2019. Characterizing the dynamics and evolution of incentivized online reviews on Amazon. *Social Network Analysis and Mining*, *9*(1), p.22.
[5]   Sun, L., Guo, J. and Zhu, Y., 2019. Applying uncertainty theory into the restaurant recommender system based on sentiment analysis of online Chinese reviews. *World Wide Web*, *22*(1), pp.83-100.
[6]   Sun, L., Guo, J. and Zhu, Y., 2018. A multi-aspect user-interest model based on sentiment analysis and uncertainty theory for recommender systems. *Electronic Commerce Research*, pp.1-26.
[7]   Bagheri, A., Saraee, M. and De Jong, F., 2013. Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews. *Knowledge-Based Systems*, *52*, pp.201-213.
[8]   Sun, Q., Niu, J., Yao, Z. and Yan, H., 2019. Exploring eWOM in online customer reviews: Sentiment analysis at a fine-grained level. *Engineering Applications of Artificial Intelligence*, *81*, pp.68-78.
[9]   Qiu, L., Gao, S., Cheng, W. and Guo, J., 2016. Aspect-based latent factor model by integrating ratings and reviews for recommender system. *Knowledge-Based Systems*, *110*, pp.233-243.
[10] Kiran, R., Kumar, P. and Bhasker, B., 2020. DNNRec: A novel deep learning based hybrid recommender system. *Expert Systems with Applications*, 144, pp. 113054.

[11] Sivaramakrishnan, N., Subramaniyaswamy, V., Viloria, A., Vijayakumar, V. and Senthilselvan, N., 2020. A deep learning-based hybrid model for recommendation generation and ranking. Neural *Computing & Applications.* https://doi.org/10.1007/s00521-020-04844-4

[12] Liu, D., Li, J., Du, B., Chang, J., Gao, R. and Wu, Y., 2020. A hybrid neural network approach to combine textual information and rating information for item recommendation. *Knowledge and Information Systems*, pp.1-26. https://doi.org/10.1007/s10115-020-01528-2

[13] Hong, W., Zheng, N., Xiong, Z. and Hu, Z., 2020. A Parallel Deep Neural Network Using Reviews and Item Metadata for Cross-Domain Recommendation. I*EEE Access*, 8, pp. 41774-41783.

[14] Liu, W., Lin, Z., Zhu, H., Wang, J. and Sangaiah, A.K., 2020. Attention-based adaptive memory network for recommendation with review and rating. *IEEE Access,* 8, pp.113953-113966.

[15] Paradarami, T.K., Bastian, N.D. and Wightman, J.L., 2017. A hybrid recommender system using artificial neural networks. *Expert Systems with Applications* pp.300-313.

[16] Wu, H., Zhang, Z., Yue, K., Zhang, B., He, J. and Sun, L., 2018. Dual-regularized matrix factorization with deep neural networks for recommender systems. *Knowledge-Based Systems*, *145*, pp.46-58.

[17] Sohail, S.S., Siddiqui, J. and Ali, R., 2019. A comprehensive approach for the evaluation of recommender systems using implicit feedback. *International Journal of Information Technology*, *11*(3), pp.549-567.

[18] Nilashi, M., Yadegaridehkordi, E., Ibrahim, O., Samad, S., Ahani, A. and Sanzogni, L., 2019. Analysis of Travellers' Online Reviews in Social Networking Sites Using Fuzzy Logic Approach. *International Journal of Fuzzy Systems*, *21*(5), pp.1367-1378.

[19] Lee, H.C., Rim, H.C. and Lee, D.G., 2019. Learning to rank products based on online product reviews using a hierarchical deep neural network. *Electronic Commerce Research and Applications*, *36*, p.100874.

[20] Jiang, M., Song, D., Liao, L. and Zhu, F., 2015. A Bayesian recommender model for user rating and review profiling. Tsinghua Science and Technology, 20(6), pp.634-643.

Author 1

Mr. A SURESH has 6+ years of Academic experience and 3+ years of research experience. He has done his B.E. in CSE from GRT Institute of Engineering and Technology affiliated to Anna University and M.Tech in CSE from Seshachala Institute of Engineering and Technology affiliated to JNTUA Currently he is pursuing his Ph.D. in Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi Chennai. His interested research area includes Big Data Analytics, Data Mining, and Recommendation Systems.

Author 2

Dr. M. J. Carmel Mary Belinda, currently working as Professor at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology. She completed her Ph.D. in the area Wireless Sensor Networks. Having 20+ years of experience in the teaching field. Published more than 20 research papers in her research area