International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 11 (2019) pp. 028-038

Received: 6 August, 2018; Accepted: 5 Jan 2019; Publish: 20 February 2019

A Recommender System Based on Group Method of Data Handling Neural Network

Meysam Shamshiri[1,*], Goh Ong Sing[2] and Yogan Jaya Kumar[2]
[1] Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Malaysia
[2]Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Malaysia
**e-mail: meysam@utem.edu.my*

***Abstract*: Nowadays, the transition of the customers' desire to move forward from traditional to intelligent systems has caused rising trends of Ecommerce businesses. This can bring lots of opportunities and challenges for internet businesses to absorb the customer desire in a competitive manner. In this regard, recommender systems help users to find and select their desired items. These systems cannot recommend without having enough information about users and their desired items such as film, music, book. One of the main goals in these systems is to collect various information about user interest and available items of the system. Most of these systems operate based on collaborative filtering method in which similarity measures are used to select similar neighbors for a user and then the recommendation is offered based on the evaluation of their comments. In this paper, a recommender system using GMDH Neural Network algorithm is proposed to recommend films to users. The proposed model is based on exploring implicit trust from active users' rates. Implicit trust networks among users are used to reduce prediction error of the improved user-oriented collaborative filtering algorithm. GMDH Neural Network offers a high learning speed even when there are few numbers of training data due to using the evolutionary genetic algorithm for the optimal design of the network structure. The proposed model is implemented using GevoM on MovieLens datasets and the results are compared with other algorithms in terms of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Results show that the proposed model outperforms other algorithms like MLP, Naïve Bayesian, J48, Bagging, SMO, RBF network, Logistic, and Random Forest with precision of 76% and absolute mean error of 0.273.**

***Keywords*:** Recommender System, Film, Neural Network, Group Method of Data Handling.

I. Introduction

Flourishment of Electronic Commerce has changed commercial behaviors and people tend to shop from online markets [1]. Considering the attraction of world wide web, some novel businesses have emerged which have their own customers. Moreover, due to the extent and continuity of the Internet, many sellers can offer their products and services without temporal and spatial constraints. As these customers are added to internet users, data and information of this network has started progressive growth. This growth has become one of the problems in using a high volume of information on the Internet. However, the progressive increase of information generated by Internet organizations and their users has created a problem known as information overhead which reduces satisfaction and fidelity of customers [2].

An approach to overcome these problems is to use tools known as Recommender Systems (RS). RS is defined as smart techniques based on the computer for facilitating high information overhead transactions [1]. These systems can benefit both sides of the transaction effectively. RS helps the customers to satisfy their needs by guiding them towards their interests. In addition, as more customers refer to these systems and considering their satisfaction, provider organizations also achieve their goals [1, 2].

RSs are in fact software techniques and tools which recommend a series of items (products and services) to users. These recommendations help the users decide. In general, these decisions might be watching a film, buying a product, visiting a web page, listening to a music and etc. In designing an RS, operational environment plays an important role. That is, the RS which is designed for film recommendation is different from the RS designed to recommend a book or grocery [2][3].

RS mainly operate in the two following forms [4][5]:

1) Predicting whether a user is interested in a specific product or not.

2) Predicting a class of products which might be selected by a user.

In fact, RSs are beneficial for both sides of an interaction and provide advantages for both of them. For instance, in a commercial interaction, customers can use RS to explore large volume information faster; sellers can also use these systems to increase customer satisfaction and increase their sale.

One of the obvious applications of RS is Cinema and recommending films to users based on their interests. Considering the large volume of cinema products, selecting one or more films is difficult for users. In such condition, a good RS which can recommend proper films based on previous users' interests to a new user in a short time might be very useful.

The purpose of this study is to present an efficient RS for recommending films to users such that proper recommendations are offered to the new user without finding similar users and using comments of other users and the trust in them. Since neural networks are one of the most common methods in exploring relationships and trust and Genetic

Algorithm (GA) has unique features in finding optimal values and exploring unpredictable spaces, using Group Method of Data Handling (GMDH) Neural Network [6][7] might be a suitable option for this RS as GMDH employs GA to design form of the Neural Network and determine its coefficients.

The rest of this paper is organized as follows: section II reviews the prior research art and literature review, section III introduces the user rates prediction and GMDH network with utilized datasets, section IV proposes a model for establishing the case study and describes the data preparing and processing, section V evaluates the performance and validation of the results, and finally section VI concludes this paper.

II. Literature Review

Social Recommender Systems (SRS) aim to solve information overhead of social media users through offering the most attractive and most related contents. Social recommenders also aim to increase adaptation, employment and contribution of new users and current social media sites. Content recommenders (weblogs and wikis) [8], tags [9], people [10] and communities [10] mainly use personalization methods matched with requirements and interests of a single user or a set of users.

RSs in Electronic Commerce area, remote education and learning improvement by modifying the old method of "What does the teacher teach and what does the student learn?" to "what do students need, what does the system provide?" are studied in [11] comprehensively. In this study, it has been mentioned that clustering users in an RS are very important and fuzzy method can be used to classify users such that quality of service is increased.

Huang and Yin [12] studied the application of RS in television network. First, a combination of descriptive algorithms reduces the effective region of users through cluster analysis. Then, the accuracy of the descriptive algorithms is improved through credit mechanism such that recommendation results are better for target users.

Yang et al. [13] proposed a dress-up RS (wearing clothes and etc.) to describe customer features. The main dress-up features include a collar, number of buttons, material, and style. The proposed algorithm builds rules and preferred model for the customer.

Ragab Pietrasieński et al. [14] proposed an incentive governmental plans of economical establishments in international. In this study, approaches based on RS are studied considering economical groups. Ragab et al. [15] proposed a college admission system using hybrid recommenders based on data-mining techniques and knowledge-exploration rules to solve college-admission prediction. The proposed method is called HRSPCA which includes two hybrid recommenders which collaborate and employ college predictor to increase efficiency.

LIAN et al. [16] proposed a RS for TV programs. Considering previous data, this system obtains a regular pattern of watching TV and a current list of programs for users. Analyses and experiments have been done using real data-set and efficiency of this algorithm are compared with previous methods.

Abbasi et al. [17] indicated that trust is useful for reducing the error of recommendations. The authors have defined trust at profile level and item level as the percentage of correct guesses from public profile and specific items point of view. Trust relationships are described by the user and they are more accurate than implicit ones; because this type of trust in another user is detected and clarified by a human agent, that is confider.

Jamali et al. [18] designed a Trust-walker method to select neighbours randomly inside the social network comprised of users and trusted users. This method is integrated with a content-based method to predict rate of items. Music recommendation system [19] which uses users' interactions inside social networks and other data propagated inside open link data area and semantic web technology, extracts RDF from music websites and performs the semantic query on them to make its recommendation is also one of these RSs.

The RS method proposed by Kazienko et al. [20] provides preferences of other users of the multimedia common system based on knowledge explored in the multidimensional social network. This system considers users' activities in separate layers of the multidimensional social network. Burke et al. [21] presented a review on hybrid recommender systems and provided some excremental results.

Ghazanfari et al. [22] used evaluation of GMDH and MLP networks for prediction of compressive strength and workability of concrete. They have aimed to assess and compare the prediction accuracy of these two methods in modeling concrete slump flow and compressive strength of concrete incorporating slag, fly ash, and super plasticizer. The simulation results have shown that the GMDH algorithm is superior to the MLP algorithm.

Ebtehajv et al. [23] used GMDH for the purpose of predicting the discharge coefficient of rectangular sharp-crested side weirs. The GMDH model is compared with the Feed-Forward Neural Network (FFNN) model in terms of performance. The simulation results have shown that the GMDH is superior to the FFNN model.

Shaghaghi et al. [24] presented a comparative analysis of GMDH Neural Network based on genetic algorithm and particle swarm optimization in stable channel design. They used GA to improve the multi-objective Pareto optimal design of GMDH results. In other word, GA is used as a encoding scheme to generalize the structure of GMDH. They also extend the Particle Swarm Optimization (PSO) algorithm to GMDH for a better comparison of the models.

Ahmadi et al. [25] proposed an intelligent model to predict the output power and torque of a Stirling heat engine. This model employs the GMDH to develop an accurate predictive tool. Also, the GA is used in a model to specify the complete structure of the GMDH and Singular Value Decomposition (SVD) is used to identify the optimum constants of quadratic formulations for predicting of torque and power.

The following subsections will discuss the detailed review in the RS.

A. User Rate Prediction

In a RS, when producing recommendations set for the target user, predicting the rate the to be given the item is the main task, because the error in recommendations and users' satisfaction from the system completely depend on this part and the employed method. For this reason, the algorithms proposed in this research field try to increase the accuracy and validity of the system's recommendations. In order to achieve this purpose, the system's recommendations should be improved. In fact, the smaller is a difference among predicted rates, prediction errors would be less, and the validity of the system's recommendations would be higher.

Rate prediction could be performed by calculating the similarity between users and items. Thus, rate prediction algorithms based on user or item can accomplish their task. Thus, rate prediction algorithms in RSs are categorized into user based and item based systems [1].

One of the rates predicting methods is the collaborative filtering method. In simple user-based collaborative filtering, first the similarity of users (who has rated the target item) with the target user should be measures and then their rates to the target item should be employed to predict the target user's rate to the weighted target item. The weight of each user's rate is its similarity with the target user. One of the methods for measuring the similarity of two users in the RS is using the Pearson correlation coefficient [2].

In the proposed method, we focus on user-based class and use data out of RS for improving recommendations and solving the user cool start problem. The user-based algorithm employs user trust data in social networks. In this paper, a collaborative filtering method based on trust without considering the similarity of people is proposed to recommend films to users.

B. GMDH Network

Evolutionary algorithms like genetic algorithm are widely applied in different stages of designing neural networks [26] such that they have unique capabilities in finding optimal values and searching in unpredictable spaces [26]. Thus, in this research, the genetic algorithm is used to design the Neural Network and define its coefficients. GMDH Neural Network based on genetic algorithm is considered as a tool with high capability in modelling complex dynamic nonlinear systems.

GMDH Neural Network consists of a set of neurons which are formed with different pairs through a quadratic polynomial. Assume there is a set of m variables including x_1, x_2, …, x_m and variable y. Data related to each variable x_i and target variable y also exist for each time period. In other words, each variable is a vector consisting time series numbers related to that variable [22-26]. Primary information which should be collected for constructing GMDH algorithm is a set of n observations which is shown as a matrix in Figure 1 [26].

There are two problems for initializing the algorithm:

I. Detecting the relation which generates the output based on input variables x_i.
II. predicting y for known values of x_is.

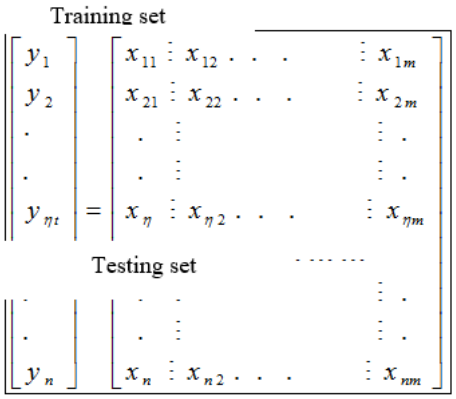

Figure 1. GMDH constructing matrix

In other words, recognizing the model and the relations among variables (modelling) can be used to predict future values of the target variable [27].

The basis of the GMDH algorithm is a process for constructing a polynomial with a high order which is known as Voltra function and it is presented as below (this polynomial is also called Ivakhennco):

$$\hat{y} = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{i=1}^{m}\sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \dots \quad (1)$$

For this purpose, we decompose Voltra functions into quadratic double variable polynomials in the GMDH algorithm.

$$G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j \quad (2)$$

In this decomposition, Voltra series is converted into a set of recurrent equations chain such that by algebraic replacements of each recurrent equation in this equation, Voltra series is established again.

$$y_i = F(x_{i1}, x_{i2}, x_{i3}, \cdots, x_{im}) \quad i = 1, 2, 3, \cdots, m \quad (3)$$

And if function f is described as below:

$$\hat{y} = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{i=1}^{m}\sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \dots \quad (4)$$

Function f can be decomposed as follows:

$$\left.\begin{array}{lll}
\hat{y}_k = G(u_i, u_j) \quad i, j = 1, 2 (i \neq j) & & k = 1 \\
\hat{u}_k = G(s_i, s_j) \quad i, j = 1, 2, k, F_1 (i \neq j) & F_1 \leq C_{F_2}^2 & k = 2 \\
\hat{s}_k = G(p_i, p_j) \quad i, j = 1, 2, k, F_2 (i \neq j) & F_2 \leq C_{F_3}^2 & k = 3 \\
\vdots & & \\
\hat{z}_k = G(w_i, w_j) \quad i, j = 1, 2, k, F_l (i \neq j) & F_l \leq C_m^2 & k = F_{l+1} \\
\hat{w}_k = G(x_i, x_j) \quad i, j = 1, 2, k, F_m (i \neq j) & & k = F_m
\end{array}\right\} \quad (5)$$

In fact, the purpose of this algorithm is to find the unknown coefficients of α in Voltra function series. It should be mentioned that all partial models will have a similar structure to the following equation:

$$\hat{f}(x_i, x_j) = v_0 + v_1 x_i + v_2 x_j + v_3 x_i^2 + v_4 x_j^2 + v_5 x_i x_j \quad (6)$$

Since our purpose in this algorithm is to model the primary system, the main model of the system can be achieved by combining the partial system model and repeating this operation.

$$\hat{y} = v_0 + \sum_{i=1}^{m} v_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} v_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{i=1}^{m}\sum_{k=1}^{m} v_{ijk} x_i x_j x_k + \ldots \quad (7)$$

In the second stage, partial models are selected, in other words, the ideal systems formed in the previous stage are combined in pairs to form new partial systems with minimum five and maximum six input variables. This way, the combination operation is continued by selecting and removing a number of partial models to obtain a relatively ideal model.

In general, GMDH Neural Network maps an input vector (X) to a numerical number (y). Figure 2 shows an example of the GMDH neural network. In this example, the network has four inputs and one output. Evolution of GMDH Neural Network is a simple feed-forward method (like perceptron). However, GMDH neural networks are not completely connected. Neurons of the first layer constitute value supply units of the first hidden layer. Output (y) might be a polynomial of order 2(k-1) where k is the total number of layers in the network.

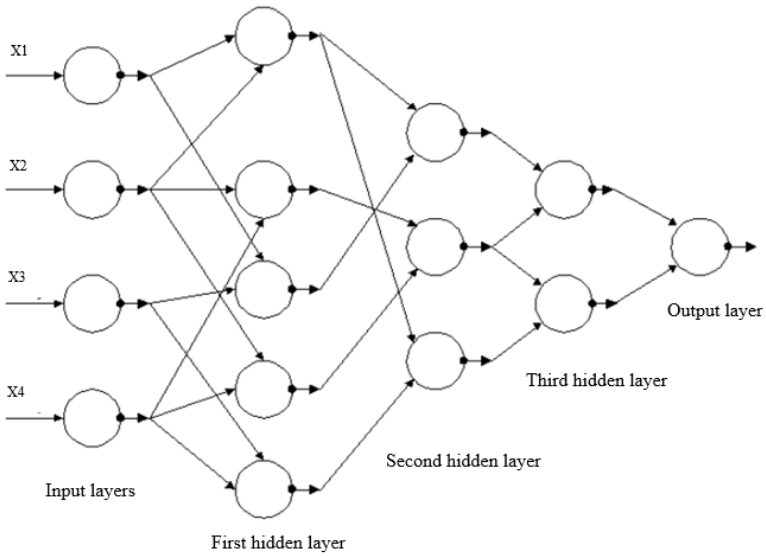

Figure 2. An example of GMDH Neural Network

Constructing GMDH network begins from the input layer and grows towards the output layer gradually as a layer at a time. Each subsequent layer begins operation with the maximum number of possible neurons (means, $C(M_{k-1},2)$) and it is adjusted by the elimination of improper neurons and determination of weights and then it remains constant. This technique is different from backpropagation technique in which all layers might collaborate in learning process simultaneously.

The main idea of GMDH is that each neuron wants to generate y at its output (desired output of the network). In other words, each neuron of the polynomial network adapts its output to the desired value of y for each input vector X of the training dataset. This approximation is performed through linear regression. Each sample in the training dataset is a linear equation on six unknowns. Then, the mean squares technique is used to explore the best combination of six weights.

The training set is used to guide adjustment of six weights of each neuron in a layer of the network which is being constructed. Usually, mean square error y of one neuron is different from other neurons. Next step is to eliminate neurons of the layer in which their error is larger than a defined "large threshold" by the user. Network construction is continued layer by layer until a stopping criterion is achieved.

One of the important issues that arises in multi-layer artificial neural networks is the design of a network structure. The number of layers as well as the inner structure, such as the number of weights and their initial values are the main factors that need be designed. In addition, the function of each neuron, should be selected appropriately that can be resulted an ideal mapping between the data incoming and outgoing.

In the design of the GMDH neural networks, the main goal is to prevent the divergence of the network and to link the shape and structure of the network to one or more numerical parameters. Changing that parameter causes the changes of the networks structure. For this purpose, the Evolutionary Algorithm (EA) is used to design the structure of the GMDH network.

In this method, a Genetic Algorithm (GA) is used for convergence of neural networks. The violation of the constraints is taken as the criterion for determining the structure of the grid. All of the neurons are given the same chances of participation in the formation of the neural network. In fact, there is no limit for creation of the network, and all process are performed with initial random inputs with purposeful process during the iterations, to obtain the optimum solution of the given problem. The maximization problem with fitness criterion for optimum selection can be discovered with total number of neurons in the entire network as well as the network output error rate compared with the tested value.

C. The dataset of the study

In this study, the MovieLens dataset of GroupLens project in Minnesota University is used. This set contains more than 1

million (1000209) rates by 6040 users on 39000 films in which each user has rated for at least 20 films. Mentioned data are collected from movielens.umn.edu for three months. Users' rates are in the form of numbers from 1 to 5 where 1, 2, 3,4 and 5 stands for bad, medium, good, very good and excellent, respectively.

Sparsity in a dataset is obtained from Eq. (8). In this equation, $|R|$, $|U|$ and $|I|$ are number of rates, number of users and number of items in the dataset.

$$sparsity = 1 - \frac{|R|}{|I| \text{ ? } |} \qquad sparsity = 1 - \frac{1000000}{3900 ? \ \ 040} \approx 0.957 \tag{8}$$

The sparsity of MovieLens dataset considering a number of rates, users, and its items and based on Eq. (8) is 0.937. Since this number is close to 1, the sparsity of this dataset is very high and suitable for evaluating descriptive algorithms which aim to overcome the sparsity of rates matrix.

III. The Proposed Model

The proposed model is based on trust factor without considering the similarity of people employed in the design and exploration of implicit trust between users and Neural Network using GMDH. The proposed system is comprised of three general sections:

1) *Preparing data*
2) *Pre-processing data*
3) *Finding implicit trust between users.*

Figure 3 shows the flowchart of the proposed model. In the following, details of the proposed model are described.

A. Data Preparation

This study uses the dataset which provided in MovieLens website (http://movielens.org). MovieLens is a web-based recommender system and virtual community that recommends movies for its users to watch, based on their film preferences using collaborative filtering of members' movie ratings and movie reviews. This recommender system was created in 1997 by GroupLens Research, a research lab in the Department of Computer Science and Engineering at the University of Minnesota, in order to gather research data on personalized recommendations. The dataset (ratings.dat) used in this study, which provided in [28], contains more than one million records in the following form:

$$UserID :: MovieID :: Rate :: TimeStamp$$

Fields of each record include user identity, film identity, user rate and time tag. First, it is required to prepare data for transmission to a proper format for applying pre-processing and processing operations on data. In order to prepare and pre-process data, MATLAB is used. Data preparation steps are as follows:

i) Dataset file, means ratings.dat, is first converted to a CSV file (for example, ratings.csv).
ii) ratings.csv contains four columns (1.user identity, 2. Film identity, 3. User rate, 4. Time tag). The fourth column, time tag is eliminated because, in the proposed system, the similarity of users and implicit trust of users' rate are used. The resulting file has more than 1 million rows wherein each row, there are user identity, film identity, and user rate.

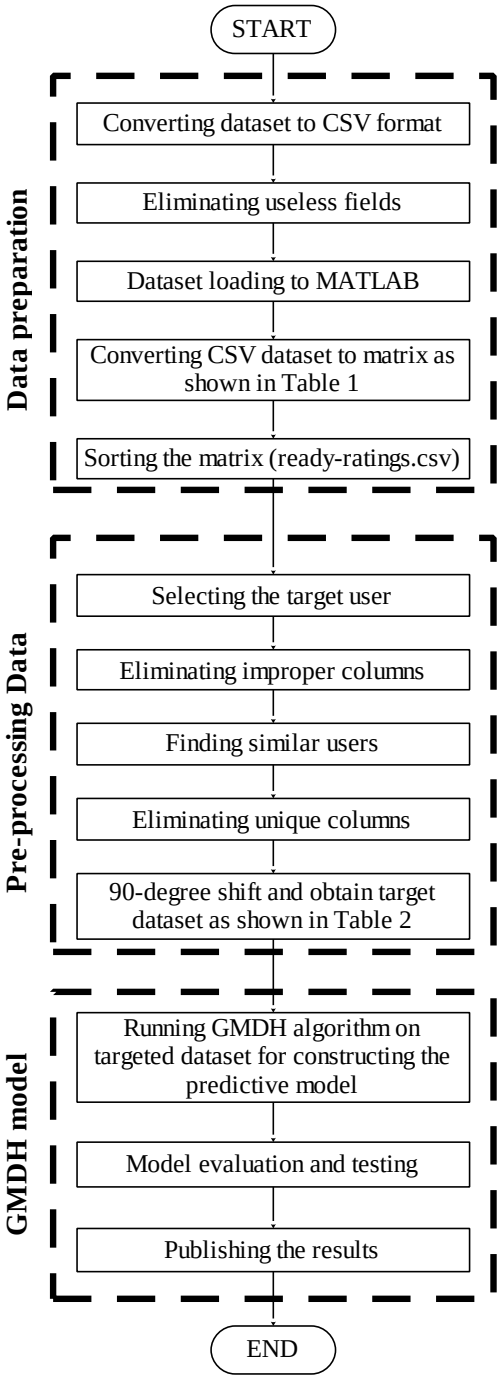

Figure 3. The flowchart of the proposed model

iii) The file ratings.csv is loaded in MATLAB.
iv) A matrix in the form of Table 1. is prepared using the dataset. This matrix has 3900 columns (total number of films) and 6040 rows (total number of users). For each user, the number of rates for each film is written in each column. If a user has not rated for a film, the corresponding column is set to zero.

Movie ID

User ID	x	x	x	x	x
	x	x	x	x	x
	x	x	x	x	x

Table 1. Data preparation matrix

v) prepared matrix is stored (entitled as ready-ratings.csv) in the form of a CSV file as a dataset .

B. *Pre-processing Data*

In order to increase the efficiency of the proposed model, important pre-processing is performed on the studied dataset. pre-processing is performed by MATLAB. In order to perform pre-processing, first, ready-ratings.csv is loaded in MATLAB and then pre-processing is applied. The pre-processing procedure is as follows:

i) Selecting the target user: among all users, a user which has the maximum number of rates is selected as the target user. In the dataset, the user with UserID=4169 is selected as the target user. Target user has rated for 2314 films.

ii) Eliminating improper columns: In this step, columns or films which target user has not rated for are eliminated from the dataset.

iii) Finding similar users: in this step, 10 users which have maximum similarity to the target user are selected among all users. In the proposed model, selecting similar users includes two phases:

a. In the first phase, 50 users which have the maximum number of rates compared to the target user are selected.

b. In the second phase, 10 users with maximum similarity are selected among 50 users selected in the first phase using Pearson correlation coefficient.

iv) Eliminating unique columns: finally, corresponding column for each film to which none of the users have rated for is eliminated.

v) 90-degree shift: the obtained dataset up to this step has 11 rows (target user and 10 similar users) and 322 columns (number of films which target user and at least one of the similar users have rated for).

The purpose of the proposed system is to find implicit trust between rates of 10 similar users compared to rate of the target user. For each film *X*, considering the number of rates to the film by 10 similar users, the number of rate for film *X* by the target user can be predicted. Thus, for each film *X*, rates issued by 10 similar users are considered as input features and rate of the target user as output field. Therefore, it is required that the dataset matrix obtained up to this point to be rotated 90 degrees clockwise or counter clockwise. The resulting dataset is converted to the form of Table 2. In this table, U1-U10 are rates of similar users and *F* is the rate of the target user.

U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	F
x	x	x	x	x	x	x	x	x	x	x
x	x	x	x	x	x	x	x	x	x	x
x	x	x	x	x	x	x	x	x	x	x

Table 2. Pre-processed data matrix

Dataset obtained from pre-processing is used as the target dataset. Various learning algorithms explore implicit relationships and recommend users based on this dataset.

C. *Finding Implicit Trust*

In the last step, relationships between comments of similar users (trusted users) and comments of the target users are explored. To this end, in the proposed model, GMDH Neural Network is used.

Using GMDH algorithm, a model can be proposed as a set of neurons in which various pair of neurons in each layer are connected together using a second-order polynomial and new neurons in the subsequent layer are generated. This representation can be used in modeling mapping inputs to outputs.

The official definition of the problem is, in fact, finding a function $\hat{}$ such that it can be used as real function *f* to predict output $\hat{y}$ for an input vector $X(x_1, x_2,\ldots, x_m)$ so that the output is as close as possible to actual *y*.

Therefore, in the proposed system, GMDH algorithm tries to select $\hat{f}$ such that it can predict the rate of the target user ($\hat{y}$) for rates of 10 active users to any of 322 films. For each film *v* from the final set of 322 films, vector $X(x_1, x_2,\ldots, x_m)$ is the rates of 10 active users to film *v* and $\hat{y}$ is the rate of the target user to film *v*. Thus, knowing *M* (in the proposed system *M*=322) observations of multiple-input-single-output (in the proposed system *n*=10 inputs and one output) such that:

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \cdots, x_{in}) \quad (i = 1, 2, 3, \cdots, M) \tag{9}$$

GMDH Neural Network can be trained to predict output values of (rates of the target user) for each given input vector $X(x_1, x_2,\ldots, x_m)$, that is:

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \cdots, x_{in}) \quad (i = 1, 2, 3, \cdots, M) \tag{10}$$

Now, the problem is to determine GMDH Neural Network such that the square difference between real output and predicted output is minimized. That is:

$$\sum_{i=1}^{M} \left[\hat{f}(x_{i1}, x_{i2}, \cdots, x_{in}) - y_i \right]^2 \rightarrow \min \tag{11}$$

In other words, the difference between the predicted rate using function $\hat{f}$ and rate of the target user should be minimized. General connection between input and output variables might be represented using a complicated discrete framework of Voltra functional series:

$$\hat{y} = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{i=1}^{m}\sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \ldots \tag{12}$$

which is known as Kolmogrov-Gabor polynomial. This complete framework of mathematical description can be represented as a system of partial second-order polynomials including only two variables (neurons) as follows:

$$\hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j \tag{13}$$

In this step, for each row (film) of the final dataset, for two users *i* and *j*, a 2-degree polynomial as Eq. (13) is created and coefficients of this polynomial are selected such that the output of this polynomial is a number close to the voting of the target user.

In this method, such partial second-order description is

used recursively in a network of connected neurons so that global mathematical relationship of input and output variables in Eq. (12) is constructed. Coefficient a_i in Eq. (13) is calculated using regression techniques such that difference between the actual output (y) and calculated output ($\hat{y}$) for each pair of x_i and x_j as input variables is minimized. It is obvious that a tree of polynomials is constructed using 2-degree equations where their coefficients are obtained in the minimum-error state. In this method, coefficients of each 2-degree equation G_i are obtained to optimize output in all set of input-output pairs.

$$E = \frac{\sum_{i=1}^{M}(y_i - G_i)^2}{M} \rightarrow \min \tag{14}$$

In the basic GMDH algorithm, all pairs of independent variables might be selected from n input variables to create a regression polynomial in the form of Eq. (13). Therefore, $\binom{n}{2} = \frac{n(n-1)}{2}$ neurons are created in the first hidden layer of the feed-forward network from $\{(y_i, x_{ip}, x_{iq});\ (i = 1,2,...M)\}$ observations for different $p,q \in \{1,2,...,n\}$. In other words, M rows of $\{(y_i, x_{ip}, x_{iq});\ (i = 1,2,...M)\}$ data can be constructed from observation using $p,q \in \{1,2,...,n\}$ in the form of the following matrix:

$$\begin{bmatrix} x_{1p} & x_{1q} & ... & y_1 \\ x_{2p} & x_{2q} & ... & y_2 \\ ... & ... & & y_M \\ x_{Mp} & x_{Mq} & ... & y_M \end{bmatrix}$$

Using 2-degree equations for each row of M data row, the following matrix equation can be obtained simply:

$$Aa = Y \tag{15}$$

where a is a vector of unknown coefficients in the 2-degree equation (16).

$$a = (a_0, a_1, a_2, a_3, a_4, a_5) \tag{16}$$

and $Y = (y_0, y_1, y_2, \ldots, y_n)^T$ is output observation vector which can be easily obtained as follows:

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & {x_{1p}}^2 & {x_{1q}}^2 \\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & {x_{2p}}^2 & {x_{2q}}^2 \\ ... & & & & & \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}x_{Mq} & {x_{Mp}}^2 & {x_{Mq}}^2 \end{bmatrix} \tag{17}$$

In the proposed system, the SNE method is used to calculate coefficients of polynomials, Eq. (17). In this method, the unknown coefficients matrix (a) is obtained as follows [6]:

$$a = (A^T A)^{-1} A^T Y \tag{18}$$

After decomposition of the main system to $\binom{n}{2}$ partial systems, a model with two variable inputs is calculated for each system. Then the partial systems are combined two by two which results in $\frac{\binom{n}{2} \times \left(\binom{n}{2} - 1\right)}{2}$ systems or a new partial model with at least three and four input variables.

Indeed, the number of model-dependent variables or number of system inputs is not important and the real estimation accuracy of the main system is important; thus, considering this rule for reducing duplex calculations and increasing efficiency and modelling accuracy, a number of the partial models which have high accuracy and estimation compared to other models are selected and others are eliminated.

In the second step, the combination of partial models is selected; in other words, ideal systems formed in the previous system are combined again to form new partial systems with at least five and six input variables. In subsequent steps, some of the partial models are selected and some are eliminated and their combination is continued until a relatively ideal model is obtained.

The objective of the combination process is to obtain a model in which all system variables play a significant role and another objective for continuous combinations is to obtain a model in which output error is less than other models.

D. Structure design models for GMDH networks

One of the important problems in multiple layer artificial neural networks (perceptron and etc.) is network structure design. In this design, the number of layers and inner structure including the number of weights and their initial values and excitation function of each neuron should be selected properly to obtain a proper and ideal mapping between input and output data.

In the proposed system, the evolutionary genetic algorithm is used to design the GMDH network structure. In this method, limitation caused by considering error as the criterion for network structure determination is removed and all neurons have the same chance to collaborate in formation of the neural network. In fact, creating such a network assumes no limitation and all operations are performed randomly and targeted for finding the most optimal structure. The only fitness measure for selection might be the total number of neurons and output error compared to the experimented value [6].

One of the advantages of the GMDH Neural Network as a result of which it has given desirable results in this study is its high learning speed when the number of training samples is low. In the target dataset, there are 322 training samples. Having such dataset, most learning algorithms like perceptron neural network, the Bayesian network, Decision Tree and support vector machine would not be able to offer a prediction model with the desired accuracy. But GMDH Neural Network does give desirable results. In the next section, simulation results of the proposed model are given and its efficiency is compared with other algorithms.

IV. Simulation Results

In order to execute the proposed model, GEvoM [29] is used; this software is designed in University of Gilan. In order to develop this software, .Net Framework is used. Other compared algorithms are evaluated using Weka.
As mentioned before, in this study, MovieLens dataset used in the GroupLens project in Minnesota University is used [28].

A. Evaluation Metrics

Metrics evaluated in this study include:

i) Root Mean Square Error (RMSE): This error is one of the most well-known measures used in the evaluation of rate prediction validity. This type of error is calculated using Eq. (19). Here, $r_{u,i}$ are real rates and $pred_{u,i}$ are predicted rates of user u for all user-item pairs. In addition, U is the set of test users.

$$RSME = \sqrt{\frac{\sum_{u \in U} \sum_{i \in testset_u} (pred_{u,i} - r_{u,i})^2}{\sum_{u \in U} testset_u}} \quad (19)$$

ii) Mean Absolute Error (MAE): Mean Absolute Error (MAE) sums the difference between the predicted rate and real rate and normalizes it. But, RMSE takes the error root before summation and normalization. RMSE assigns the higher weight to higher error; therefore, it gives a better measure of error. Eq. (20) shows the calculation of this error.

$$MAE = \frac{\sum_{u \in U} \sum_{i \in testset_u} | pred_{u,i} - r_{u,i} |}{\sum_{u \in U} testset_u} \quad (20)$$

iii) Precision: Another measure in prediction is precision. Simply, precision is the ratio of total correct predicted rates (observations) to total rates. Precision is calculated using Eq (21).

$$\frac{TP}{TP + FP} \quad (21)$$

Here, TP is the number of rates which are predicted positive correctly and FP is the number of rates which are predicted positive incorrectly.

The efficiency of the proposed algorithm is compared with Multi-Layer Perceptron (MLP) Neural Network [30], Radial Basis Function (RBF) network [31], Naïve Bayesian network [31], Bagging [31], Sequential Minimal Optimization (SMO) [32], Decision Tree (J48) [31], Logistic [33], and Random Forest [33].

Weka is used to run existing algorithms. Also, *k*-fold technique is used to construct and test the predicting models. In this technique, the original sample is randomly partitioned into k equal size subsamples. Of the *k* subsamples, a single subsample is retained as the validation data for testing the model, and the remaining *k*-1 subsamples are used as training data. The cross-validation process is then repeated *k* times (the folds), with each of the *k* subsamples used exactly once as the validation data. The final results were averaged from the *k* results of the folds. The advantage of this technique is that all observations are used for both training and validation, and each observation is used for validation exactly once.

B. Experimental Results

The proposed system is executed in GevoM tool [] and prediction results are presented in the form of a diagram as in Figure 4.

As can be seen, the proposed system could not predict rate number 1 as none of the 10 selected active users have rated for the films which the target user has rated for. In addition, in predicting rate number 5, its efficiency is low because none of the active users has rated for 5 (like 1) or most of them have selected a score other than 5. For instance, one has rated for 5 but 3 others have rated for 2 (rates of active users fluctuate much). But, its efficiency in the prediction of rates number 2, 3 and 4 is desirable.

In this system, rates for films are integer values between 1 to 5. But GMDH Neural Network performs regression and its output is a decimal between 1 to 5. That is, the output of this Neural Network might not be a tag (or nominal value). While other compared algorithms like Bayesian network, Decision Tree, Logistic and Random Forest are executed on data which their output field is nominal, MLP and RFB network can be implemented for numerical and nominal outputs. Thus, in order to compare the proposed algorithm with other algorithms, the output of the proposed system is converted to nominal values between 1 to 5. The simplest approach is to round outputs of the proposed model to integer numbers where any of these integer numbers can be interpreted as nominal values.

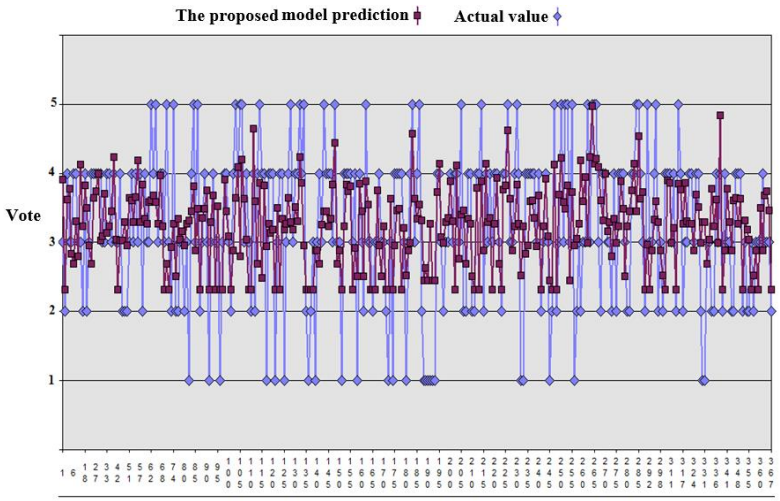

Figure 4. The output of the Proposed Model in compared to actual values

Other compared algorithms are executed in Weka and their results are compared with the proposed model in terms of RMSE, MAE, and precision.

Experiment results in Figure 5 show that the proposed model outperforms J48, Random Forest, Naïve Bayesian network, SMO network, MLP network and Logistic by a precision of 33%, 30%, 27%, 32% and 34%. As we mention prior, GMDH network offers a high learning speed even when there are few numbers of training data. The result of this experiment show that the precision rate of the proposed method in predicting users' rates is extremely superior to other algorithms. Because, in the cases which the target dataset has not enough number of training data, many algorithms such as J48, Random Forest, Naïve Bayesian, SMO, MLP, and Logistic cannot learn well. But, GMDH due to using the evolutionary genetic algorithm for the optimal design of the network structure, can learn much faster.

In addition, Figure 5-7 compare the proposed model with other algorithms in terms of RMSE and MAE, respectively. The results show that the proposed model outperforms other algorithms in terms of both RMSE and MAE. MAE of the proposed model is 0.273 while the value of this measure for other algorithms is larger than 0.276.

In terms of RMSE, the proposed model outperforms other algorithms with a value of 0.276 while other algorithms have an error rate higher than 0.38.

As mentioned before, one of the most important advantages of GMDH Neural Network is that its prediction precision is high even when the number of samples is low. While other algorithms including Decision Tree, Random Forest, Naïve Bayesian, SMO network, MLP network and Logistic do not have such advantage and a dataset including a large number of samples is required for achieving a model with high precision.

In GMDH neural network, prediction precision is high because the evolutionary genetic algorithm is used to design optimal structure of the network and maximum neighborhood techniques are used to estimate parameters; in cases, where its prediction is incorrect, the difference of the predicted rate compared to real rate is low.

In GMDH neural network, during each training cycle, synaptic weights of each neuron which minimizes the normal error between real values and predicted values are calculated and branches (neurons) with minimum collaboration in the calculation of the output are eliminated and other branches are preserved as a result of which their synaptic weight remains constant. Thus, MAE and RMSE of the proposed model are much lower than other compared models. As a recommendation, the proposed method can be applied to conventional intelligent bots [27] to enhance the learning process.

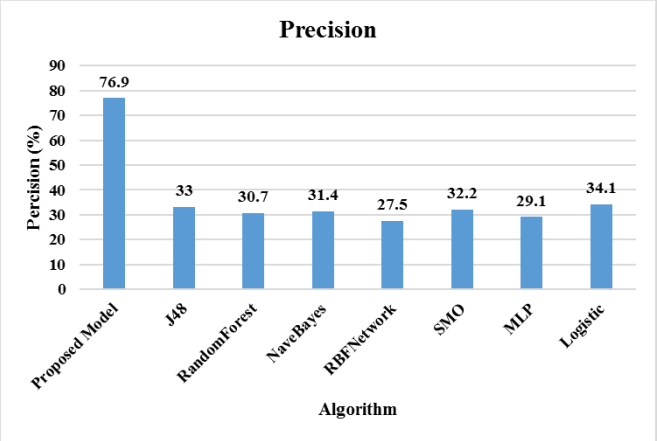

Figure 5. Comparing the Proposed Model with other Algorithms in terms of Precision

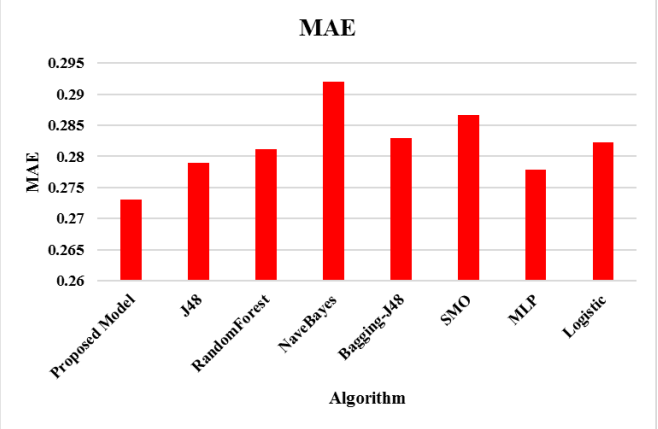

Figure 6. Comparing the Proposed Model with other Algorithms in terms of MAE

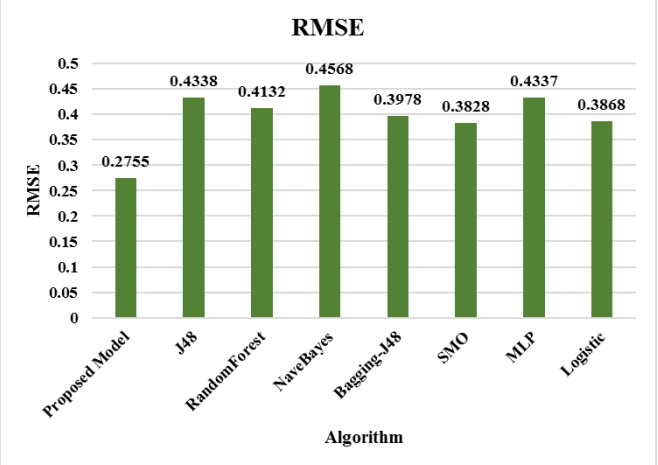

Figure 7. Comparing the Proposed Model with other Algorithms in terms of RMSE

V. Conclusion

In this paper, a RS based on GMDH Neural Network is proposed to recommend films to users. In the proposed method, a trust-based approach is used to solve the collaborative filtering problem. Implicit trust networks among users are used to reduce prediction error of the improved

user-oriented collaborative filtering algorithm. Prediction results of the proposed model are extracted in terms of precision and error and the results are compared with several common algorithms like MLP, Bayesian network, and Decision Tree. Results show that the proposed model is superior with a precision of 76% and an MAE of 0.273.

As future works, it can be suggested a hybrid model consisting of three GMDH, J48 and SMO algorithms to predict users' rates. The performance of this hybrid model is such that each algorithm trained the predictive model based on the data set, individually. Then, the test data set is given in parallel as input to each of these three predictive models. The prediction results from these three models are combined with either Voting or other methods such as Mean Value or Weighting methods to obtain the final prediction. This approach can increase the accuracy of the proposed model in predicting user ratings, because of the parallel operation of the three algorithms.

Acknowledgment

This research work is supported by Universiti Teknikal Malaysia Melaka (UTeM) and Big Data Technology Pty Ltd, Australia (GLuar/BIGTECH/2017/FTMK-CACT/A00009).

References

[1] Lu, J., Wu, D., Mao, M., Wang, W. and Zhang, G., 2015. Recommender system application developments: a survey. *Decision Support Systems*, *74*, pp.12-32.

[2] Sharma, L. and Gera, A., 2013. A survey of recommendation system: Research challenges. *International Journal of Engineering Trends and Technology (IJETT)*, *4*(5), pp.1989-1992.

[3] Massa, P. and Avesani, P., 2007, October. Trust-aware recommender systems. In *Proceedings of the 2007 ACM conference on Recommender systems* (pp. 17-24). ACM.

[4] Francesco Ricci, et al., Recommender Systems Handbook, Springer, 2011.

[5] Jannach, D., Zanker, M., Felfernig, A. and Friedrich, G., 2010. *Recommender systems: an introduction*. Cambridge University Press.

[6] http://www.gmdh.net/

[7] Witczak, M., Korbicz, J., Mrugalski, M. and Patton, R.J., 2006. A GMDH neural network-based approach to robust fault diagnosis: Application to the DAMADICS benchmark problem. *Control Engineering Practice*, *14*(6), pp.671-683.

[8] Guy, I., Zwerdling, N., Ronen, I., Carmel, D. and Uziel, E., 2010, July. Social media recommendation based on people and tags. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 194-201). ACM.

[9] Sigurbjörnsson, B. and Van Zwol, R., 2008, April. Flickr tag recommendation based on collective knowledge. In *Proceedings of the 17th international conference on World Wide Web* (pp. 327-336). ACM.

[10] Guy, I., Ronen, I. and Wilcox, E., 2009, February. Do you know?: recommending people to invite into your social network. In *Proceedings of the 14th international conference on Intelligent user interfaces* (pp. 77-86). ACM.

[11] Zhang, L. and Liu, W., 2014. Research on user clustering of recommended system based on fuzzy clustering. *Applied Mechanics & Materials*.

[12] Huang, Q. and Yin, S., 2013. Network TV Recommended System Framework and Research of Collaborative Filtering Algorithm. *Video Engineering*, *9*, p.038.

[13] QI, Yang and ZHU, X.J., 2010. An apparel recommended system based on data mining [J]. *Journal of Xi'an Polytechnic University*, *4*, p.009.

[14] Pietrasieński, P., 2011. The evolutionary character of supporting the internationalisation processes: recommended system solutions. *Polish Journal of Management Studies*, *4*, pp.193-199.

[15] Ragab, A.H.M., Mashat, A.F.S. and Khedra, A.M., 2012, November. HRSPCA: Hybrid recommender system for predicting college admission. In *Intelligent Systems Design and Applications (ISDA), 2012 12th International Conference on* (pp. 107-113). IEEE.

[16] LIAN, B., WANG, L. and PEI, Y., 2013. A Recommendation Methods of Television Programs Based on sub-communities Mining [J]. *Journal of Taiyuan University of Technology*, *3*, p.020.

[17] Abbasi, M.A., Tang, J. and Liu, H., 2014. Trust-aware recommender systems. *Machine Learning book on computational trust, Chapman & Hall/CRC Press*.

[18] Jamali, M. and Ester, M., 2009, June. Trustwalker: a random walk model for combining trust-based and item-based recommendation. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 397-406). ACM.

[19] Passant, A. and Raimond, Y., 2008, October. Combining Social Music and Semantic Web for music-related recommender systems. In *Social Data on the Web Workshop*.

[20] Kazienko, P., Musial, K. and Kajdanowicz, T., 2011. Multidimensional social network in the social recommender system. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *41*(4), pp.746-759.

[21] Burke, R., 2002. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, *12*(4), pp.331-370.

[22] Ghazanfari, N., Gholami, S., Emad, A. and Shekarchi, M., 2017. Evaluation of GMDH and MLP Networks for Prediction of Compressive Strength and Workability of Concrete. *Bulletin de la Société Royale des Sciences de Liège*, Vol. 86, pp. 855-868.

[23] Ebtehaj, I., Bonakdari, H., Zaji, A.H., Azimi, H. and Khoshbin, F., 2015. GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs. *Engineering Science and Technology, an International Journal*, *18*(4), pp.746-757.

[24] Shaghaghi, S., Bonakdari, H., Gholami, A., Ebtehaj, I. and Zeinolabedini, M., 2017. Comparative analysis of GMDH neural network based on genetic algorithm and particle swarm optimization in stable channel design. *Applied Mathematics and Computation*, *313*, pp.271-286.

[25] Ahmadi, M.H., Ahmadi, M.A., Mehrpooya, M. and Rosen, M.A., 2015. Using GMDH neural networks to model the power and torque of a stirling engine. *Sustainability*, *7*(2), pp.2243-2255.
[26] Onwubolu, G. and Onwubolu, G., 2015. *GMDH-methodology and implementation in MATLAB*. Imperial College Press.
[27] Pradana, A.D.I.T.Y.A., Goh, O.S. and Kumar, Y.J., 2018. Intelligent Conversational Bot for Interactive Marketing. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *10*(1-7), pp.1-4.
[28] http://grouplens.org/datasets/movielens/1m/
[29] Ahmadi, H., Mottaghitalab, M. and Nariman-Zadeh, N., 2007. Group method of data handling-type neural network prediction of broiler performance based on dietary metabolizable energy, methionine, and lysine. *Journal of Applied Poultry Research*, *16*(4), pp.494-501.
[30] Takács, G., Pilászy, I., Németh, B. and Tikk, D., 2009. Scalable collaborative filtering approaches for large recommender systems. *Journal of machine learning research*, *10*(Mar), pp.623-656.
[31] Amatriain, X., Jaimes, A., Oliver, N. and Pujol, J.M., 2011. Data mining methods for recommender systems. In *Recommender systems handbook* (pp. 39-71). Springer, Boston, MA.
[32] Di Noia, T., Mirizzi, R., Ostuni, V.C. and Romito, D., 2012, September. Exploiting the web of data in model-based recommender systems. In *Proceedings of the sixth ACM conference on Recommender systems* (pp. 253-256). ACM.
[33] Zhang, H.R. and Min, F., 2016. Three-way recommender systems based on random forests. *Knowledge-Based Systems*, *91*, pp.275-286.

Author Biographies

Meysam Shamshiri was born in Iran in 1983. He received his B.Eng. in Electrical & Electronic Engineering from Islamic Azad University, Toyserkan, Iran in 2008. He received his M.Eng, 2013 and PhD, 2017 from Universiti Teknikal Malaysia Melaka (UTeM) in electrical engineering. He is currently doing his post-doctoral in the faculty of electrical engineering at UTeM. His research interests include distribution network planning, demand response application, smart grid development, integration of renewable energy, data mining and machine learning.

Goh Ong Sing started his career as an academician since 1990 at Universiti Sains Malaysia (USM). In August 2002, he joined Universiti Teknikal Malaysia Melaka (UTeM) and in 2015 he was appointed as a Director of Industry Liaison Centre and Lead Samsung IoT Academy. Now he is working as Assistant Vice Chancellor at Office of Industry and Community Network. His main research interest is in the development of intelligent agent, natural language processing and speech technology to facilitate graceful human-computer interactions, conversational robot and mobile services. He is the author of more than 100 peer-reviewed scientific journal, books, conference papers, and book chapters. He has led and worked on research grants funded by Malaysian Government's Intensified Research, Malaysia Technology Development Corporation, Big Data Technology Pty Ltd, Murdoch University, and Australian Research Centre for Medical Engineering at University of Western Australia.

Yogan Jaya Kumar is a Senior Lecturer at the Department of Intelligent Computing and Analytic in the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). He earned both his bachelor degree and master degree from Universiti Sains Malaysia (USM), in the field of Mathematical Science in year 2003 and 2005. He completed his PhD studies at Universiti Teknologi Malaysia, in 2014 in the field of Computer Science. Currently, his research involves in the field of Text Mining, Information Extraction and AI applications.