Deep Learning-based Evolutionary Recommendation Model for Heterogeneous Big Data Integration

Hyun Yoo[1], and Kyungyong Chung[2*]
[1]Contents Convergence Software Research Center, Kyonggi University
154-42, Gwanggyosan-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, 16227, South Korea
[e-mail: rhpa0916@gmail.com]
[2]Division of Computer Science and Engineering, Kyonggi University
154-42, Gwanggyosan-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, 16227, South Korea
[e-mail: dragonhci@gmail.com]
*Corresponding author: Kyungyong Chung

Abstract

This study proposes a deep learning-based evolutionary recommendation model for heterogeneous big data integration, for which collaborative filtering and a neural-network algorithm are employed. The proposed model is used to apply an individual's importance or sensory level to formulate a recommendation using the decision-making feedback. The evolutionary recommendation model is based on the Deep Neural Network (DNN), which is useful for analyzing and evaluating the feedback data among various neural-network algorithms, and the DNN is combined with collaborative filtering. The designed model is used to extract health information from data collected by the Korea National Health and Nutrition Examination Survey, and the collaborative filtering-based recommendation model was compared with the deep learning-based evolutionary recommendation model to evaluate its performance. The RMSE is used to evaluate the performance of the proposed model. According to the comparative analysis, the accuracy of the deep learning-based evolutionary recommendation model is superior to that of the collaborative filtering-based recommendation model.

Keywords: Data Mining, Deep Learning, Recommendation, Multimedia, Data Integration

1. Introduction

$\mathbf{A}$s new technologies have developed in the convergence of technologies, and as knowledge from various fields triggers another convergence, user lifestyles have been and will continue to be influenced and changed. In order to respond to the environmental changes brought about by the Fourth Industrial Revolution (4IR) era of aging populations and low birth rates, there is an urgent need to achieve smart health innovations based on this convergence technology. Along with changes in social and living habits, disease patterns have also changed. At present, chronic diseases have become a major cause of death, as environmental changes react with prolonged living habits in complex ways [1,2]. Various types of smart health applications have been developed in combination with Ambient Intelligence (AmI), the Internet of Things (IoT), considerably high communication-network infrastructures, and cloud computing. The convergence of different digital devices and the development of bio-signal measurement devices have made it possible to easily identify one's health conditions and provide them with personalized healthcare service [3]. For the operation of smart health services, big data are received from various digital devices and combined, and then the potential values and meaningful knowledge are extracted and analyzed. The optimized artificial neural network (ANN) that uses supervised-learning based machine learning frequently shows better performance in various areas than the conventional methodology.

With the rapid development of deep learning technology, products that have used this technology in their development have started to be released. In the areas of healthcare, precision medicine, and new drug clinics, machine-learning based research is continually being conducted, leading to the development of potential new therapies. As such, deep learning technology is crucial for developments in the medical field. A service user can obtain a variety of information using deep learning technology. Watson, the artificial intelligence platform developed by IBM, uses a machine-learning algorithm and AI to provide the most appropriate medical examinations and accurate treatments [4]. Using its past experience and knowledge of various situations, it judges and treats diseases. By learning the medical big data that comprise the patient data as well as the examination results and causal relations regarding treatments, along with analyzing the current symptoms and medical records on the basis of medical histories, such a platform generates the most likely selection, medical judgment, and examination. Accordingly, it is possible to apply such medical big data-based AI in the medical industry to create new values, such as customized medical services and disease-prevention initiatives, thereby changing the smart-health paradigm. The information extraction of such a new paradigm makes it possible to provide more accurate and proper information than that provided by classical statistics-based research [5,6]. This will be helpful by saving medical costs, increasing the quality of medical services, and improving national health.

The heterogeneous big data-based machine-learning deep neural network (DNN) model combines multiple neural-network layers and obtains a meaningful result through efficient reasoning and learning. Based on large learning data, a DNN learns from items that have been properly observed, and can thereby create a proper approximation function. The technology is applied to solve problems in various areas, including image processing, object recognition, tracking, and textual and vocal recognition, and it exhibits excellent performance. Through a deep learning-based lifelog analysis, it is possible to find the life habits or patterns of similar users, which can then provide more diverse and personalized services [7]. Recently, for a more intelligent service provision, numerous studies applying the combination of either

heterogeneous big data and lifestyles or a machine-learning algorithm and AI have been conducted. In order to provide personalized services, a recommendation system is applied to diverse fields [8]. A recommendation system employs an algorithm that uses a particular user's personalized information to recommend an item that is most likely to be the preferred option. Collaborative filtering, the representative algorithm of a recommendation system, generates a group based on the similarity of certain items, thereby creating reliable expected information depending on the similarity. This method has been comprehensively evaluated in the areas of news searches and shopping item recommendations, both of which are based on evaluation processes.

As one of the various techniques for smart healthcare, medical big-data analysis has been applied. This technique is used to extract meaningful information from medical big data for the judgment of a health condition, and it can provide a proper healthcare service through the use of data mining and an AI algorithm [2,6,9]. In order to improve the accuracy of different models, it is necessary to take the characteristics of each algorithm into account. In addition, through a verified performance evaluation, it is necessary to obtain results that are more improved than those obtained from the conventional model and to complete the verification. Therefore, in research on mining-based smart healthcare, accurate data analysis can be expected when completing the following four steps: data collection, data transmission, data analysis, and feedback [10].

This study is organized as follows: Chapter 2 describes the related studies on the heterogeneous big data-based recommendation technology and the recent directions in the deep neural network technology. Chapter 3 describes the proposed deep learning-based evolutionary recommendation model for heterogeneous big data integration. Then, Chapter 4 describes the performance evaluation and Chapter 5 provides a conclusion.

2. Related Work

2.1 Heterogeneous big data-based recommendation technology

A recommendation system uses information filtering to analyze a user's preference so as to then provide a customized item that is likely to be of interest to that user. This item could be statistics, bioinformatics, business administration, medical services, marketing, weather information, or security management. The big data-based information filtering techniques use collaborative filtering, content-based filtering, image filtering, machine learning based filtering, and hybrid filtering [8,11]. Each method is used to increase the recommendation accuracy of a target item, with each method having its own advantages and disadvantages, so they can be properly combined to generate hybrid filtering [12]. In a recommendation system, collaborative filtering is divided into memory-based filtering and model-based filtering according to the heterogeneous big data type.

In a social network, recommendation technology is used to analyze the consumers' big data, and it is applied to product advertising, exhibitions, and marketing. In order to dynamically extract the explicit and implicit preferences, deep learning and machine learning are used. Deep learning uses a convolution neural network (CNN), autoencoder, or recurrent neural network (RNN) as an ANN algorithm. Machine-learning techniques can be classified into supervised learning and nonsupervised learning depending on whether or not learning data are present. Depending on the specific problem, these techniques learn through either a regression analysis, classification, or clustering. Regression analysis is a technique that is applied to extract the output of input data. Classification is a technique that can extract the attribute or

type of input data. Clustering is a technique that conducts grouping based on the output of input data. A recommendation system applies a machine-learning algorithm and AI in the big-data application areas. By entering training data in a learning machine, it is possible to create a learning model that can be used for processing [13]. At present, for both structured or unstructured data, big data-based systems are applied in spam email filtering, stock price prediction, Facebook advertising, Amazon.com recommendations, Google Trends, profiling, data sports, presidential election forecasting, and lifestyle analysis.

As a development environment, a Hadoop Distributed File System (HDFS) that runs in a cloud environment is used for large data. Hadoop's architecture consists of an HDFS cluster and MapReduce. An HDFS cluster consists of multiple data nodes, each of which is processed through several servers. MapReduce distributes tasks to multiple machines, thereby enabling the concurrent processing of the distributed tasks in parallel. A MapReduce task consists of Map and Reduce functions. The Map function disassembles forms that consist of a key-value and a value–value, and also creates a new key-value as well as a new value–value. The Reduce function integrates the same key value into one index for the purpose of computation. Through this process, the HDFS cluster and MapReduce develop a distributed-processing framework based on big data.
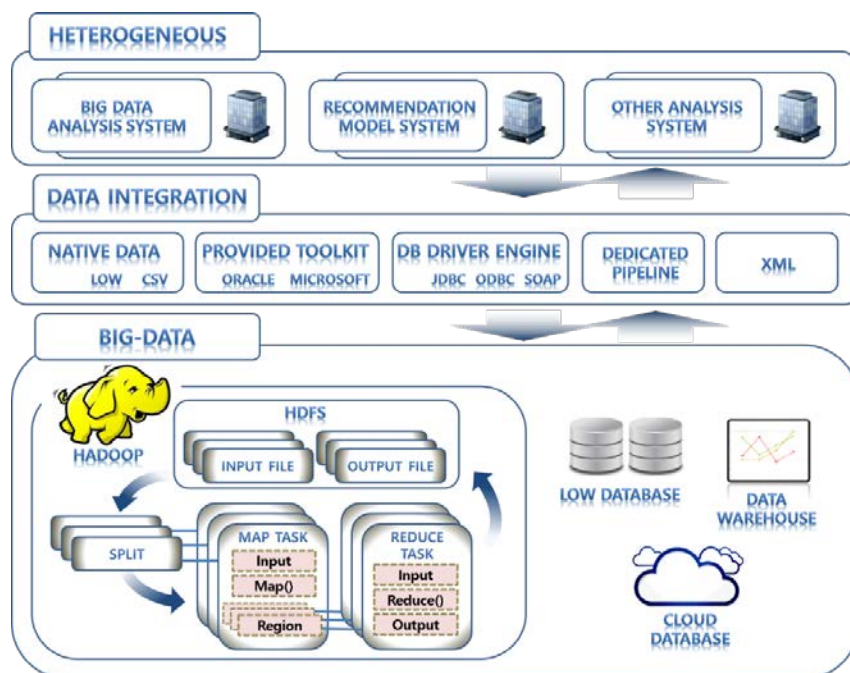
Fig. 1. Heterogeneous Big Data Integration Configuration

Fig. 1 shows the heterogeneous big data integration configuration. A big data-based learning model configured in this way is interlocked with the recommended systems through heterogeneous big-data integration. Obtaining improved functional results in the integrated configuration requires effective integration, and the data integration configuration offers high analysis speed and real-time analytic capacity. Models that are diversely configured have unstructured data such as medical imaging data and biometric log data. In the case of the unstructured data, the effectiveness of a real-time analysis applying big-data processing is low.

Therefore, in external terms, it is necessary to configure an exclusive data pipeline that can be quickly connected in a cloud sourcing environment. Internally, the problem regarding heterogeneous big data integration can be solved through various methods, such as the effective design of the predefined scenario models and proper structuring of the data analysis model. Further, the HDFS process enables distributed data processing of the big data through cluster computing. This process is more suitable for big data processing than real-time processing, as the former, in which a single big data set is divided into several pieces for its configuration into the map phase, requires more time. The processed results are then collected through a merger, followed by the Reduce execution phase, in order to obtain the results.

Regarding the research trend of the sensibility-engineering design recommendation, a sensibility-engineering design-recommendation system is applied to fashion, automobiles, textile, bags, shoes, makeup, bicycles, and glasses so that personalized services can be provided based on the user sensibility [14,15]. Sensibility consists of sense-based psychological experiences and the perceptions of physical stimuli, and it is generated using a combination of various feelings including all five senses, comfort, discomfort, and inconvenience. In order to measure the user sensibility, peer-to-peer (P2P) hybrid networking applies sensibility-engineering interactions to connect human sensibility with design as well as for contextual recognition [16]. This networking technology solves the problem of the central-server-based data transmission, and it uses model-based collaborative filtering to overcome the problems associated with memory-based correlation-analysis technology. The developed system uses convergence technology for human power to provide specialized hybrid-emotion contents. Further, it is applied in the operation of sensibility-science experience programs in terms of living so as to contribute to the improvement of the self-directed research ability and the problem-solving ability [17].

Regarding the research trend of healthcare technology, a clinical decision-making support system provides diagnostic information to enable medical staff to access clinical information and medical knowledge that may be helpful for understanding a patient's situation [2]. In terms of clinical settings, the examination and treatment of patients in big data-based remote medical services is actively being researched [13]. Further, deep learning-based studies on the base causes of brain diseases and the brain-structure function and connection network regarding Alzheimer's disease are being conducted as well. For example, for knowledge-based dietary-nutrition recommendation [18], the context of an obese adolescent is used to create a similarity group that shares a high correlation. In the similarity group, the similarity weight for the {user-diet} merger matrix is applied, and then collaborative filtering is applied in order to recommend customized recipes and the corresponding nutritional information. Using mobile devices, it is possible to provide personalized services at any time and in any place. Recommendations for obese adolescents, who have emerged as a social issue, have been introduced at home and abroad, and the results show that they are clinically effective. Medical big data, including bio data and crowdsourcing data obtained from social media, are analyzed. Using knowledge-based crowdsourcing [11], the index service solved the rarity of collaborative filtering and the problem of the initial evaluation, thereby saving time and money; furthermore, it improved the accuracy and reproducibility of the index service which makes different recommendations depending on the user [19,20].

2.2 Recent directions in Deep Neural Network technology

A neural-network algorithm is a learning algorithm developed from the idea of the biological neural-network structure. The basic ANN is a mathematical model, and its design is based on the nerve-cell structure of the biological brain, including a network structure that is akin to the nerve-cell synapses [2]. In the ANN model, it is possible to change the coupling intensity of the artificial neurons using neural-network learning, and this provides the model with the problem-solving ability. In the previously mentioned network structure, the independent functionality of each node provides each node with excellent parallelism, and the nodes are distributed in the connection lines with a large amount of information; therefore, it is possible to correct the errors that arise in some of the nodes. An ANN comprises multiple input nodes and one output node, and the sum of the received data from multiple nodes is used for processing; this sum leads to the creation of the intensity of the output data depending on the setup-value criterion. The learning structure of an ANN comprises nodes that each have a link with a weight. In the full mode, each weight is adjusted depending on the input intensity, and this leads to repeated learning. A weight is a means for the nodal memory that represents each node's importance. If a new input value comes in after the learning has been completed, a proper output value is extracted depending on the link weight. Generally, in the same way multiple biological nerve cells are linked to each other, the nodes in an ANN can be linked through their inputs and outputs in order to create multiple layers. The link weight of each layer is presented with a weight that can be changed during the operation of the ANN. The ANN with these multiple layers and a link weight is defined as a DNN that shows favorable results in the modeling of a complex nonlinear relation, and this type is frequently applied in the learning and recognition areas. Specifically, the structure of a DNN is divided into three parts: the first and second parts are the input layer and the multiple hidden layers for the data input, respectively, while the third is the output layer for the final output. It is possible to create a DNN structure for the object-identification model or for vocal recognition. In this structure, the objects can be expressed in terms of the hierarchical composition of the basic factors, while the additional layers can be used to collect the characteristics of the lower layers; due to this advantage, as compared to an ANN with a similar structure, the ANN with the DNN structure can model various types of data with smaller units (nodes). The operation type of universal neural network modeling is called Back Propagation. The process of updating weights from the Input layer to the Output layer and taking the results from activation function is called the 'forward' process. In the forward process, there is an error between the result generated in the output layer and the actual value. Back Propagation sends this error in a backward manner so as to calculate a weight. One time of weight calculation is one epoch. With a rise in epochs, it is possible to learn weights continuously and gradually reduce an error. In a deep neural network, myriad nodes are connected, so their weights need to be calculated. Accordingly, an efficient method of calculating these weights is required. A universal weight calculation method shows Gradient Descent, in which it is necessary to repeat a stepwise weight correction process according to the gradient of errors [8,9].

However, this ANN is hampered by the problem of a long calculation time in terms of the learning, a usage difficulty that is caused by local optimization, and an overfitting problem that is caused by excessive fitting to the previous training data. Since the research on neural networks has continued and the competency of graphics processing unit (GPU)-based high-performance parallel computing has improved, research on DNN-based deep learning algorithms has led to substantial improvements in the corresponding calculation amount, for which a high-speed parallel calculation based on the General Purpose of Graphic Processor Unit (GPGPU) technology was applied. Recently, the problems of the deep learning-based

learning and reasoning model have been solved, along with improvements in its algorithm and the security of the big data involved, as well as the remarkable development of its hardware, and as such it is now applied in various convergence areas. The development of deep learning technology has been rapid, as shown by the launches of products that have been developed using this technology. Further, in the deep learning-based AI systems, the big data analysis, vocal recognition, voice synthesis, image recognition, and image feature extraction have been improved in terms of performance. In particular, the three-dimensional (3D) anatomical brain-structure-splitting based on deep learning has shown improved accuracy and reproducibility, meaning that it is useful in the exploration of the brain structure that is not revealed in brain images, and it is also effective in the exploration of not only normal brain structures, but also abnormal brain structures.

Currently, with the use of deep learning-based data analysis, spam mail filters, stock price prediction, Facebook advertising, recommendations at Amazon.com, Google trends, profiling, data sports, presidential election prediction, and other predictive services are continuously being developed and applied. In addition, voice recognition speaker and interpreter devices are being released. As such, relevant technologies are being developed quickly. In the future, the healthcare system in combination with IoT and a big data processing algorithm will be able to understand users' health information and dietary habit patterns, and thereby provide more diverse and personalized healthcare services.

3. Deep learning-based Evolutionary Recommendation Model for the Heterogeneous Big Data Integration

3.1 Data preprocessing for the heterogeneous big data integration

For the basic data on the healthcare services of the Republic of Korea (ROK), the raw data of the Korea National Health and Nutrition Examination Survey (KNHANES) are used. The KNHANES was conducted by the Korea Centers for Disease Control & Prevention, and these data are used as the fundamental data for the establishment and evaluation of the overall public-health policies [21]. As the timely preparation of the evidence materials for the establishment and evaluation of the national health policy is becoming increasingly necessary, the annual national statistics as well as the city and provincial statistics of every three-year period are provided. The raw material of the National Health and Nutrition Examination Survey includes an examination survey, a health-questionnaire survey, and a nutrition survey. The examination survey consists of a basic survey table, family history, thyroid-disease examination, lung-function examination, tuberculosis (TB) examination (chest X-ray), oral examination, eye examination, chromoscopy, otolaryngologic examination, bone-density examination, osteoarthritic examination, and muscular examination. **Fig. 2** shows the process of the data preprocessing and selection.
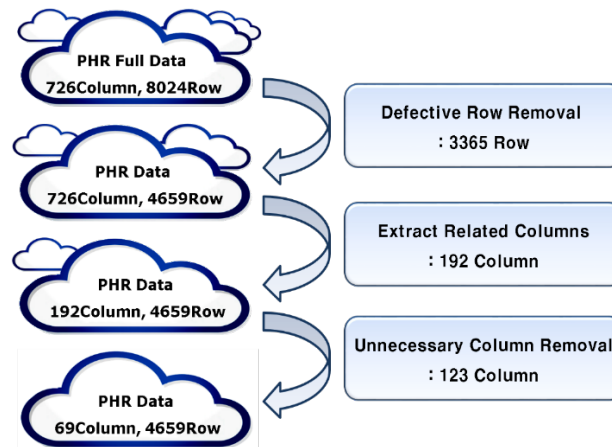
**Fig. 2.** Data Preprocessing and Selection Process

The respondents to the health questionnaire survey are composed of adults, adolescents, and young children, and its nutrition survey includes a dietary life survey, food-intake survey, food-intake frequency survey, and food-stability survey. The present study uses the seventh statistical data of 2018 among the datasets from 2016 to 2018 [21]. This statistical material consists of 726 items and 8,024 data points. In order to determine the significance of the healthcare data, 534 subitems that have been deemed to bear no relation to the health pattern and the dietary lives have been excluded [7,22]. Additionally, the items related to pregnancy and childbirth, specifically, 61 simple data regarding early childhood and adolescence and 62 data on the dependent or associated relations are excluded, along with 3,365 data that are missing answers and contain errors.

Ultimately, 69 items and 4,659 records were selected. Of the 69 items, one is the current diabetes disease type, which is used as a fixed factor when the significance is determined. For the data selection and preprocessing, the previous study [23,24] was used as a guideline. **Table 1** presents the selected attribute information.

Table 1. Selected Attribute Information

No.	Item	Attribute	Unit
1	I_LF_DIABETES	Current status of diabetes	Boolean
2	I_LF_S14_1	Frequency of not eating due to lack of food (including family)	Integer
3	I_SEX	Gender	Boolean
4	I_BP8	Average sleeping time per day	Integer
5	I_BE5_2	Flexible exercise days per week	Integer
6	I_BS1_1	Current smoking status	Boolean
7	LF_S8_1	Frequency of not eating during the day due to dieting	Integer
8	I_PA_WALK	Walking practice	Boolean
9	I_BE3_31	Frequency of walking for a week	Integer
10	I_LF_SECUR_G	Degree of food stability by levels	Integer
:	:	:	
69	I_N_FE	Iron intake (*mg*)	Integer
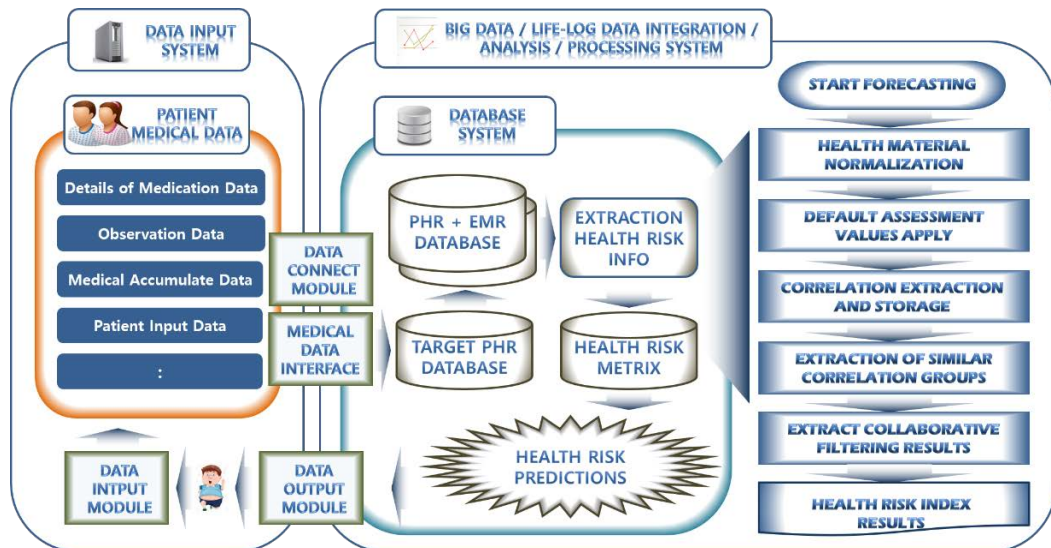
Fig. 3. Process of Health Status Decision using Collaborative Filtering

The selected data were applied to the clustering operation for various chronic diseases with the use of collaborative filtering. The collaborative filtering was performed on the results that were drawn from the users' different environmental information. The clustering operation was executed on the similar users' information, and the values of the associated clustered items were then obtained. In health services, collaborative filtering is used to create a cluster that is similar to the past user information when the user environment changes, and it is also used to judge the most appropriate similar cluster for the new information. This filtering process is used as an algorithm to determine different health-risk situations depending on the individual user. A new user can obtain information on the disease that is the closest to his/her input information as well as the risk priority [16]. By learning autonomously and creating a pattern based on the evaluation data, collaborative filtering is used to judge a health condition and provide personalized healthcare service to individual users [23,24]. **Fig. 3** illustrates the health-status decision-making process for which the collaborative filtering technique is applied.

3.2 Deep learning-based evolutionary recommendation model

The DNN algorithm of machine learning is useful in data classification and clustering, and its data change identification power is strong. Therefore, the algorithm is useful for analyzing the feedback data and re-evaluating the forecast information. The method of reanalyzing a user's evaluation data based on an algorithm can ensure the accuracy and reproducibility of a mining-based health platform [25,26]. The evolutionary recommendation model uses a DNN algorithm and feedback information collected from the users. Therefore, it is possible to learn from the input data drawn from the evaluation results and create a proper approximation function with high accuracy [27]. The evolutionary recommendation model is designed with collaborative filtering and the learning algorithm of a DNN. Further, the data are observed, and the learning occurs on the basis of the observed data. The DNN learning is classified into supervised learning and autonomous learning, each of which is properly selected depending on the given issue. In order to evaluate the feedback repeatedly, a favorable reinforcement-learning method is used. As previously stated, the DNN structure consists of the input layer, output layer, and hidden layer [28]. The internal structure of the evolutionary recommendation model is shown in **Fig. 4**.
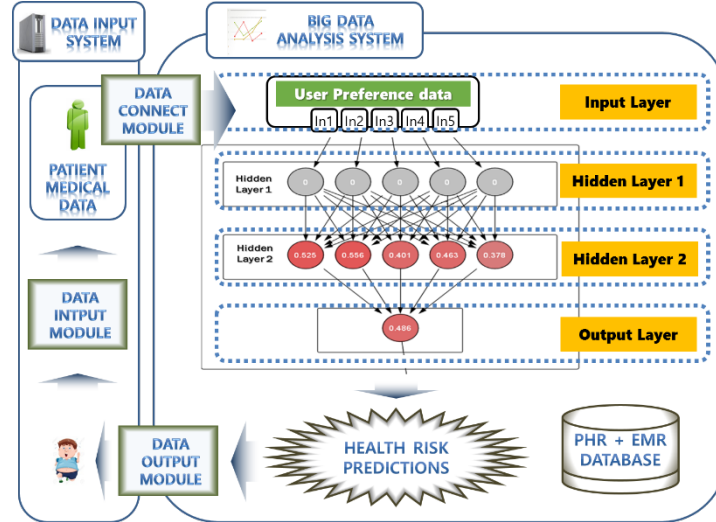
Fig. 4. Internal Structure of an Evolutionary Recommendation Model

In this study, the input layer is composed of five nodes, the hidden layer can include various hidden nodes depending on the test results, and the output layer consists of one output node. The feedback-evaluation step comprises a five-step scale ranging from 1 to 5 points. With 3 points representing a neutral condition, the evaluation with lower points means a negative condition, whereas the evaluation with higher points means a positive condition [24].
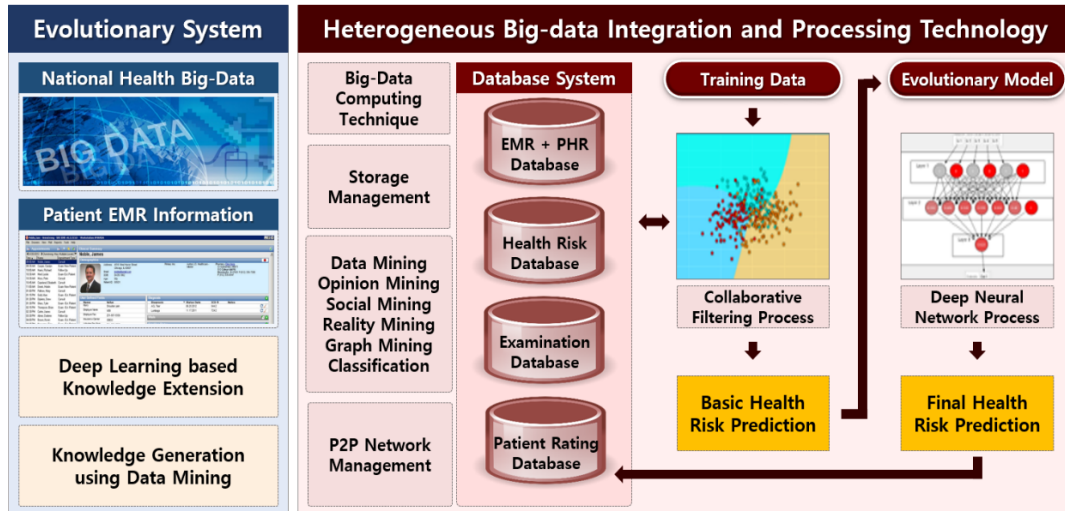
For the evolutionary recommendation model, a backpropagation algorithm is employed for the inter-node learning. The backpropagation algorithm uses the preference that occurs in the output layer in order to repeatedly calculate the error of the hidden layer. Then, the corrected value becomes a backpropagation to the input layer. Accordingly, the backpropagation operation is repeatedly executed until the output-layer error reaches a proper level. The evolutionary recommendation model generally uses the error backpropagation. The error backpropagation is a technique used to adjust the weight of each node, which enables the achievement of a proper output value in the same input layer [28,29]. Typically, a weight is updated by the gradient descent. The gradient descent is used to find the error of the input values using the gradient and a continual movement toward the smaller gradient. The algorithm performs repeated operations until the error reaches its minimum value, then updates the link weight in the direction of the error minimization. A cost function is needed to find the minimum error. The cost function E_p is calculated using Equation (1); O is the output value of the function, d is the target output, and p is the pattern data sequence number.

$$E_p = \frac{1}{2}\sum_j (d_{pj} - o_{pj})^2 \tag{1}$$

d_{pj} is the target output value of the jth element included in the pth pattern data, o_{pj} is the function result of the jth element included in the pth pattern data. In order to minimize the error of the function $f(x)$, the formula for expressing the relationship between the target value x_{i+1} and x_i is as follows in Equation (2).

$$x_{i+1} = x_i - \alpha\nabla f(x_i) \tag{2}$$

is the slope and α is the learning coefficient. Larger values of α approach the resultant value quickly, but the error sometimes oscillates.

Fig. 5. Evolutionary Recommendation System for a Heterogeneous Big-data Integration

Fig. 5 shows the evolutionary recommendation system for the heterogeneous big data integration, wherein the preprocessing of the initial data, the collaborative filtering, and the DNN algorithm are shown. In order to develop an evolutionary recommendation model, it is necessary to change the size of the layers, the complexity of the backpropagation algorithm operation, and the learning rate of each node [30]. Based on the results from repeated tests, it is possible to design the most appropriate recommendation model. In proportion to the number of nodes that are used depending on the size of the evolutionary recommendation model, it is possible to remarkably increase the amount of data as well as the learning count, thereby leading to exponential increases in the learning-time cost. In particular, the greater the number of hidden layers that are present, the higher the cost of the connection weight that is increased exponentially [31,32]. Therefore, in this study, a hidden layer of a high importance and the learning count are set as the evaluation criteria. The DNN function for comparing the learning count is designed with five steps. In this way, the performance and accuracy are compared depending on the number of hidden layers and the learning count.

4. Performance evaluation

In order to evaluate the performance of the deep learning-based evolutionary recommendation model, a server–client model for health platforms was designed. The hardware (H/W) components of the server were the Intel(R) i5-6600 3.30 GHz CPU with an 8-GB memory, and its software (S/W) components were the Apache Tomcat 9.0, which served as the Hypertext Transfer Protocol (HTTP) server, and the MySQL 5.7 database management system (DBMS). For the client device, an LG G3 LTE-A Cat. 6 Android device (LG Electronics, ROK) was used. The H/W of the client device included the Snapdragon 805 mobile processor with 2.5 GHz and 3 GB memory (Qualcomm Technologies, USA). The server received and saved the user's environmental information, then calculated the association with a group of users with chronic diseases and the weight of the feedback evaluation based on the basic data [27,33].

The root mean square error (RMSE) was used to evaluate the proposed model [12]. The RMSE, which represents the difference between an actual value (observed value) and an estimated value, is generally used to find the difference between an estimated value or a model's predicted value and an actual value; its definition is similar to that of the standard deviation [31]. In this study, the RMSE is specifically defined as the difference between the predicted value of the collaborative filtering model and the predicted value of the deep learning-based evolutionary recommendation model. The RMSE can be used to evaluate the accuracy of the proposed model, and it is given by Equation (3) as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \qquad (3)$$

Where n is the total extraction count, O_i is the actual result, P_i is the algorithm-result value, and i is each extraction step. In the experiment, a virtual user with a particular preference is created so as to evaluate the data. The preference tendency of the virtual user is randomly irregular, and 465 search and feedback occasions, which accounts for 10 % of the total data, is set as the basic unit. For the learning-count comparison, the accuracy over three steps is compared according to a unit of 10 times. The four types model consists of one collaborative filtering model and three evolutionary recommendation models with 10, 100, and 1000 learning counts. A total of four types of evolutionary recommendation models are compared, so the matching accuracy between the change in the resulting value and the virtual user's preference tendency is measured depending on the user preference. so the matching accuracy between the change in the resulting value and the virtual user's preference tendency is measured depending on the user preference. The final performance evaluation results are presented in **Table 2**.

Table 2. Final Performance Evaluation Results

Collaborative Filtering (RMSE)	Virtual Users		Evolutionary Recommendation Model		
	User Number	RMSE	Learning count 10	Learning count 100	Learning count 1,000
0.15	1	1.24	0.12	0.03	0.01
	2	1.73	0.58	0.15	0.03
	3	0.99	0.90	0.23	0.05
	4	0.67	0.60	0.15	0.03
	5	1.37	0.23	0.07	0.02
	:	:	:	:	:
	465	0.33	0.30	0.08	0.02
Average	-	1.46	0.56	0.14	0.03

According to the results of this experiment, the accuracy of the deep learning-based evolutionary recommendation model is greater than that of the collaborative filtering-based recommendation model, and this is attributed to the fact that the DNN structure was applied to the user's evaluation value. Regarding the performance, with a rise in the learning count, the accuracy was higher, but the operation cost was also higher. In addition, as the hidden layers were increased, the cost became higher, but the accuracy increased further. Given the results, it

became evident that the algorithm combined with the DNN and collaborative filtering showed better performance than the conventional collaborative filtering which only used data analysis.

5. Conclusion

In the health field, research on improving recommendation systems has been conducted using heterogeneous big-data analysis. In particular, the algorithms used to increase the user accuracy and reproducibility have been developed in various ways. The diverse types of algorithms for big data analysis present both advantages and disadvantages. If the characteristics of each algorithm are understood, and a combination of the algorithms in consideration of the advantages and disadvantages is used, the algorithm's performance can be improved. The algorithmic convergence can result in higher accuracy and enhanced results, thereby making a complementary design possible. This study proposes an evolutionary recommendation model using a deep learning-based DNN algorithm, while a gradient descent-based backpropagation algorithm was applied to create an inter-node model. In order to experimentally measure the accuracy, recommendation information was extracted from the preprocessed data of the Korea National Health and Nutrition Examination Survey (KNHANES). For the health platform, this study developed an algorithm using a combination of the collaborative filtering-based recommendation and the deep learning-based evolutionary recommendation model. In order to evaluate the performance, the RMSE was applied to compare the resulting values of the collaborative filtering method and the modified value obtained using the deep learning-based evolutionary recommendation model. According to the evaluation, the prediction method using the deep learning-based evolutionary recommendation improved the accuracy. Therefore, given that the recommendation accuracy was improved in the health-platform environment, the proposed model is found to be effective. This complex algorithm can improve the performance of a recommendation system, thereby making it possible to provide a variety of users with highly satisfactory personalized services in the health platform, such as exercise management, lifestyle management, emergency management, sleep care, stress care, cancer care, and obesity care. These are new added-value health services, and can lead to savings in medical costs as well as an improved quality of life (QOL).

References

[1] B. R. Wang, J. Y. Park, K. Chung, I. Choi, "Influential Factors of Smart Health Users according to Usage Experience and Intention to Use," *Wireless Personal Communications*, Vol. 79, No. 4, pp. 2671-2683, December, 2014. Article (CrossRef Link)

[2] J. C. Kim, K. Chung, "Mining based Time-Series Sleeping Pattern Analysis for Life Big-data," Wireless Personal Communications, Vol. 105, No. 2, pp. 475-489, March, 2019. Article (CrossRef Link)

[3] DoCoMo Healthcare, http://www.d-healthcare.co.jp/english/.

[4] IBM Blumix Service, https://www.ibm.com/cloud-computing/bluemix/.

[5] R. C. Park, H. Jung, K. Chung, K. H. Yoon, "Picocell based Telemedicine Health Service for Human UX/UI," *Multimedia Tools and Applications*, Vol. 74, No. 7, pp. 2519-2534, April, 2015. Article (CrossRef Link)

[6] H. Yoo, K. Chung, "Heart Rate Variability based Stress Index Service Model using Bio-Sensor," *Cluster Computing*, Vol. 21, No. 1, pp. 1139-1149, March, 2018. Article (CrossRef Link)

[7]   H. Yoo, K. Chung, "PHR based Diabetes Index Service Model using Life Behavior Analysis," *Wireless Personal Communications*, Vol. 93, No. 1, pp. 161-174, March, 2017. Article (CrossRef Link)

[8]   J. C. Kim, K. Chung, "Mining based Time-Series Sleeping Pattern Analysis for Life Big-data," *Wireless Personal Communications*, Vol. 105, No. 2, pp. 475-489, March, 2019. Article (CrossRef Link)

[9]   J. C. Kim, K. Chung, "Emerging Risk Forecast System using Associative Index Mining Analysis," *Cluster Computing*, Vol. 20, No. 1, pp. 547-558, March, 2017. Article (CrossRef Link)

[10] R. C. Chen, C. F. Hsieh, W. L. Chang, "Using Ambient Intelligence to extend Network Lifetime in Wireless Sensor Networks," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 7, No. 6, pp. 777-788, December, 2016. Article (CrossRef Link)

[11] J. C. Kim, K. Chung, "Depression Index Service using Knowledge based Crowdsourcing in Smart Health," *Wireless Personal Communication*, Vol. 93, No. 1, pp. 255-268, March, 2017. Article (CrossRef Link)

[12] K. Chung, J. H. Lee, "User Preference Mining through Hybrid Collaborative Filtering and Content-based Filtering in Recommendation System," *IEICE Transaction on Information and Systems*, Vol. E87-D, No. 12, pp. 2781-2790, December, 2004.

[13] K. Chung, J. C. Kim, R. C. Park, "Knowledge-based Health Service considering User Convenience using Hybrid Wi-Fi P2P," *Information Technology and Management*, Vol. 17, No. 1, pp. 67-80, March, 2016. Article (CrossRef Link)

[14] H. Jung, K. Chung, "P2P Context Awareness based Sensibility Design Recommendation using Color and Bio-signal Analysis," *Peer-to-Peer Networking and Applications*, Vol. 9, No. 3, pp. 546-557, May, 2016. Article (CrossRef Link)

[15] K. Chung, H. Yoo, D. Choe, H. Jung, "Blockchain Network based Topic Mining Process for Cognitive Manufacturing," *Wireless Personal Communications*, Vol. 105, No. 2, pp. 583-597, March 2019. Article (CrossRef Link)

[16] J. C. Kim, K. Chung, "Mining Health-Risk Factors using PHR Similarity in a Hybrid P2P Network," *Peer-to-Peer Networking and Applications*, Vol. 11, No. 6, pp. 1278-1287, November, 2018. Article (CrossRef Link)

[17] K. Chung, Y. Na, J. H. Lee, "Interactive Design Recommendation using Sensor based Smart Wear and Weather WebBot," *Wireless Personal Communications*, Vol. 73, No. 2, pp. 243-256, November, 2013. Article (CrossRef Link)

[18] H. Jung, K. Chung, "Knowledge-based Dietary Nutrition Recommendation for Obese Management," *Information Technology and Management*, Vol. 17, No. 1, pp. 29-42, March, 2016. Article (CrossRef Link)

[19] Health Insurance Review and Assessment Service (HIRA), http://opendata.hira.or.kr/.

[20] J. C. Kim, K. Chung, "Associative Feature Information Extraction using Text Mining from Health Big Data," *Wireless Personal Communications*, Vol. 105, No. 2, pp. 691-707, March, 2019. Article (CrossRef Link)

[21] Korea Centers for Disease Control and Prevention, https://knhanes.cdc.go.kr/.

[22] H. Jung, H. Yoo, K. Chung, "Associative Context Mining for Ontology-Driven Hidden Knowledge Discovery," *Cluster Computing*, Vol. 19, No. 4, pp. 2261-2271, December, 2016. Article (CrossRef Link)

[23] H. Yoo, K. Chung, "Mining-based Lifecare Recommendation using Peer-to-Peer Dataset and Adaptive Decision Feedback," *Peer-to-Peer Networking and Applications*, Vol. 11, No. 6, pp. 1309-1320, November, 2018. Article (CrossRef Link)

[24] I. Mashal, O. Alsaryrah, T. Y. Chung, "Testing and Evaluating Recommendation Algorithms in Internet of Things," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 7, No. 6, pp. 889-900, December, 2016. Article (CrossRef Link)

[25] H. Jung, K. Chung, "Life Style Improvement Mobile Service for High Risk Chronic Disease based on PHR Platform," *Cluster Computing*, Vol. 19, No. 2, pp. 967-977, June, 2016. Article (CrossRef Link)

[26] K. Chung, H. Yoo, D. E. Choe, "Ambient Context-based Modeling for Health Risk Assessment Using Deep Neural Network," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11, pp. 1387-1395, 2020. Article (CrossRef Link)

[27] HL7 Health Level Seven International, http://www.hl7.org/.

[28] D. E. Rumelhart, G.E. Hinton, R.J. Williams, "Learning Representations by Back-propagating Errors," *Nature*, Vol. 323, pp. 533-536, October, 1986. Article (CrossRef Link)

[29] D. Wang, W. Ding, X. Ma, H. Jiang, F. Wang, J. Liu, "MiFo: A Novel Edge Network Integration Framework for Fog Computing," *Peer-to-Peer Networking and Applications*, Vol. 12, No. 1, pp. 269-279, January, 2019. Article (CrossRef Link)

[30] T. Chen, H. R. Tsai, "Application of Industrial Engineering Concepts and Techniques to Ambient Intelligence: A Case Study," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 9, No. 2, pp. 215-223, April, 2018. Article (CrossRef Link)

[31] F. Orciuoli, M. Parente, "An Ontology-driven Context-aware Recommender System for Indoor Shopping based on Cellular Automata," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 8, No. 6, pp. 937-955, November, 2017. Article (CrossRef Link)

[32] K. Chung, R. C. Park, "PHR Open Platform based Smart Health Service using Distributed Object Group Framework," *Cluster Computing*, Vol. 19, No. 1, pp. 505-517, March, 2016. Article (CrossRef Link)

[33] K. Dhir, A. Chhabra, "Automated Employee Evaluation using Fuzzy and Neural Network Synergism through IoT Assistance," *Personal and Ubiquitous Computing*, Vol. 23, No. 1, pp. 43-52, February, 2019. Article (CrossRef Link)
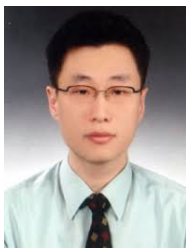
Hyun Yoo has received B.S., M.S., and Ph.D. degrees in 1999, 2011, and 2019, respectively, all from the Department of Computer & Information Engineering, Sangji University, South Korea. He has worked for Fujitsu and XCE (SK). He has been a researcher at Data Mining Lab., Kyonggi University, South Korea. He is currently a research professor in the Contents Convergence Software Research Center, Kyonggi University, South Korea. His research interests include Deep Learning, Artificial Intelligent, Big Data Mining, Ambient Intelligence, Medical Decision System, Healthcare, Recommendation, and HCI.

Kyungyong Chung has received B.S., M.S., and Ph.D. degrees in 2000, 2002, and 2005, respectively, all from the Department of Computer Information Engineering, Inha University, South Korea in 2000, 2002, and 2005, respectively. He has worked for the Software Technology Leading Department, Korea IT Industry Promotion Agency (KIPA). From 2006 to 2016, he was a professor at the School of Computer Information Engineering, Sangji University, South Korea. Since 2017, he is currently a professor in the Division of Computer Science and Engineering, Kyonggi University, South Korea. He was named a 2017 Highly Cited Researcher by Clarivate Analytics. His research interests include data mining, artificial intelligence, healthcare, knowledge systems, HCI, and recommendation systems.