ISSN: 2345-2897; Health Education and Health Promotion. 2019;7(3):139-145

The Application of an Evolutionary Model using Fuzzy Logic on Health Literacy Data

ARTICLE INFO

Article Type
Descriptive Study

Authors
Enayati E.*[1] *MSc,*
Hassani Z.[2] *MSc,*
Moodi M.[3] *PhD*

How to cite this article
Enayati E, Hassani Z, Moodi M. The Application of an Evolutionary Model using Fuzzy Logic on Health Literacy Data. Health Education and Health Promotion. 2019;7(3):139-145.

[1]Computer Sciences Department, Basic Sciences Faculty, University of Bojnord, Bojnord, Iran
[2]Computer Sciences Department, Basic Sciences Faculty, University of Kosar, Bojnord, Iran
[3]"Social Determinant of Health Research Center" and "Health Education & Health Promotion Department, Health Faculty", Birjand University of Medical Sciences, Birjand, Iran

Correspondence
Address: Basic Sciences Faculty, University of Bojnord, Esfarayen Road, Bojnord, Iran. Postal Code: 9453155111
Phone: +98 (58) 32240017
Fax: +98 (58) 32410700
E.Enayati@ub.ac.ir

Article History
Received: May 7, 2019
Accepted: June 8, 2019
ePublished: July 21, 2019

ABSTRACT

Aims Health literacy (HL) is the main factor shows health literate level of people in a certain society. Discovering and understanding affective factors on HL level could lead experts to improve these factors in the target community. This study aimed to Health Literacy classification of population and find a major component with data mining approaches.

Instruments and Methods In this paper, we have acquired more details about major factors on the health literacy level of target society by assessing evolutionary methods. We benefit of Particle Swarm Optimization (PSO) and KNN and fuzzy KNN algorithm for classification and use wrapper technique for feature selection by our model. Feature selection are done as weighted features and selects the most effective features of health literacy. Our proposed model evaluates a data set of Health Literacy by two classifiers with/without fuzzy logic. Applied data set is a real data gathered from a descriptive-analytic cross-sectional study on adult population include 2133 record with 74 attributes in 2016 at South Khorasan province. We have gained effective factors on HL level of the population according to regions and total population without using any statistical analysis tools with the lowest human interference by an evolutionary method.

Findings Proposed model have found effective factors on the health literacy level of population in South Khorasan province. Results are obtained 92.02% accuracy for the total population and 97.99% for regions population.

Conclusion Simulations demonstrate the evolutionary method is a suitable way for extracting results from health data sets and also shows the superiority of the proposed method.

Keywords Health Literacy; Data Mining; Machine Learning; Fuzzy Logic

CITATION LINKS

[1] Beyond reading and understanding: Health literacy as the capacity to ... [2] Healthy ageing: A challenge for ... [3] Health literacy: A prescription to end ... [4] Low health literacy: Implications for national health ... [5] The costs of limited health literacy: A systematic ... [6] Health literacy in Europe: Comparative results of the European Health Literacy Survey ... [7] Health literacy and the influencing factors: A study in five provinces of ... [8] Analyzing diabetes datasets using data ... [9] Data-mining technologies for diabetes: A systematic ... [10] Systematic review of data mining applications in patient-centered mobile-based information ... [11] Privacy preserving distributed association rule mining approach on vertically partitioned healthcare ... [12] Prediction of benign and malignant breast cancer using data mining ... [13] Application of data mining: Diabetes health care in young and ... [14] Breast cancer diagnosis via data mining: performance analysis of seven different ... [15] Data mining techniques and applications - a decade review from 2000 to ... [16] Data mining and medical world: Breast cancers' diagnosis, treatment, prognosis and ... [17] Data mining application for exploring the relationship between addiction and ... [18] Data mining algorithms and techniques in mental health: A systematic ... [19] Applying data mining techniques in ... [20] Particle swarm optimization: A survey of historical and recent developments with hybridization ... [21] A comprehensive survey on particle swarm optimization algorithm and its ... [22] Particle swarm ... [23] Compressed kNN: K-nearest neighbors with data ... [24] Application of the weighted k-nearest neighbor algorithm for short-term load ... [25] Fuzzy k-nearest neighbor method to classify data in a closed ... [26] Hybrid intelligent system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and neural networks combined with a fuzzy ... [27] Health literacy status and understanding of the prescription instructions in diabetic ... [28] A new adaptive testing algorithm for shortening health literacy ...

Introduction

Health Literacy

Nowadays, it is important a health system has suit health and treatment facilities but it is not enough. Actually, having literate individuals which manner in a healthy way and have a good lifestyle and could have good functionality in sick situations is more important. Health awareness level is very effective in reducing treatment costs, social costs, Psychological costs and economic costs in sick conditions [1]. The ones who are educated in school or university aren't literate individuals in this area. It is Literacy in health information and their applications in daily life. This type of literate named as Health Literacy (HL).

Health Literacy definition is an individual ability to obtain, interpretation and understanding necessarily health recommendations and services which all of them help to suit determination [2]. It is a related concept to reading and listening ability, educational, social and governmental agents, individual behavior, and many others. Adequate HL level in a population means they could extract required health information and instruction from health learning media like pamphlets, newspaper, videos, TV programs and etc., which is prepared by the health system. These lead them to the correct decisions and health services. Also, HL signifies one could or couldn't follow the doctor's instructions as well as. Health Literacy defined first time by Kick Busch in 1997 [3].

Cost of low health literacy is for all such as people, governments, and health system. In a study [4] author estimate low HL cost in American is between $106 billion and $238 annually. In another research conducted by Klaus Eichler *et al*. [5], it is showed that the consequences of low Health literacy cost for the health care system are 3-5% extra cost per year. It cost for each person with low HL level than high HL level in the US is $143 to 7,798. The European Health Literacy Survey (HLS-EU) in a 2011 research showed that 12% of the population suffer from low health literacy and about 35% have limited health literacy [6]. In Iran, the results of a study of Tehrani Banihashemi *et al*. [7] showed the level of health literacy in 5 provinces. It was adequate HL in 28.1%, limited in 15.3% and inadequate in 56.5% of the population.

Data Mining (DM) in Health system

In today's world, on one hand, a large volume of data is produced and gathered in various health centers and hospitals. Knowledge extraction from these data sets is a complex process. These processes may change the fate of a population.

On the other hand, data mining techniques are a process of pattern recognition and knowledge extraction from data. In many pieces of researches [8-13], DM approaches are used in different level of health care systems. Therefore, these techniques could efficient in each of stages such as prevention, diagnosis, treatment or health education.

Prevention: Data mining could find hidden patterns of health behaviors and predict the result of these behavioral patterns [12].

Diagnosis: Data mining techniques assist to discover a health problem through association rules and hidden pattern recognition. It is a way for early detection of special health subjects like diseases [14].

Treatment: In this stage medical data warehouse which is a collection of previous treatment approaches and results could select the best therapy plan.

Health promotion and education: It is serious to recognize Target community and the impression of health educations on it. So, data mining is a perfect way for occurring knowledge about society.

In all cases mentioned above, data mining methods are applicable [15-18]. Implementation of DM techniques is achievable with artificial intelligence algorithms, evolutionary algorithms, and machine learning. Many data scientist researchers around the world are investigating on health and medical field [19].

Aims of the study: In this study, we mine a health literacy data set for discovering extra knowledge about effective factors on HL level in South Khorasan province which ordinal statistical approach may couldn't find them. We proposed an evolutionary model that benefits of Particle Swarm Optimization algorithm (PSO) algorithm and K-Nearest Neighbor Algorithm (KNN) algorithm as a cost function and fuzzy logic for improving our model in classification.

Instruments and Methods

The data collection is used from a descriptive-analytic cross-sectional study on adult population (Over 18 years of age) in South Khorasan province in 2016. Individuals included referrals to all health centers covered by health centers in the South Khorasan cities. We have assigned a code from 1 to 11 in order to Birjand, Boshrooyeh, Tabas, Sarayan, Zirkouh, Darmian, Khosf, Ferdows, Sarbisheh, Nehbandan, Qaen. Information was collected using multi-stage random cluster sampling. The Iranian health literacy questionnaire was used to collect data. This research project which already was conducted by the Ministry of Health with the Health Education and Health Promotion center in cooperation with the Modeling Research Center in Health Sciences University of Medical Sciences Kerman has been volunteered in seven selected provinces and in some provinces. We only used collected data in South Khorasan province.

The questionnaire contains 69 questions in 9 dimensions of health literacy, which includes questions regarding access to information resources (Q1:1-5), use of information resources (Q2:6-11),

ability to read Persian (Q3:12-16), ability to understand the material (Q4:17 -24), ability to interpret, judge and evaluate (Q5:25-30), decision-making and communication ability (Q6:31-38), health knowledge (Q7:46-53), individual empowerment (Q8:58-65), and social empowerment (Q9:66- 69). The final health literacy score was range from 0 to 20: Less than 10 were weak, 14-10 average and over 14 were adequate. The total sample size was 2133 people. Each city based on their population have specific sample portion. Samples were individuals referred to all health centers covered by the health centers of Birjand University of Medical Sciences.

Questioner used in descriptive-analytic cross-sectional study are in the table 1.

Table 1) Data features (Questioner used in descriptive-analytic cross-sectional study) and statistical analysis of population

Num	Question	Mean±SD
	Birth date; Sex; Resident city; Education level; Occupation	
Q1	Type of health information resource accessible in previous month; Physician, internet, radio and TV, journals and newspapers, pamphlets - brochures-booklets	10.7±0.1
Q2	Type of health information resource used in previous month; Physician, internet, radio and TV, journals and newspapers, pamphlets or brochures or booklets	9.3±0.2
Q3	Reading ability of health information; Pamphlets- brochures- booklets; Written orders of doctors, dentists or health personnel; Medical and dental forms; Guidance boards in hospitals or health centers; Guidance forms before surgery, sonography or radiology	12.4±0.1
Q4	The understanding ability of health information; Guidance boards in hospitals or health centers; Guidance forms before surgery, sonography or radiology; Health information on the radio and TV; Health information on the internet or electronic resources; Health information in medical professional books; Health information in newspapers, booklets, pamphlets and brochures; Written drugs order by pharmacist or factory; Health personnel recommendations	13.6±0.07
Q5	Interpreting ability of health resources information; Health information in the radio and TV; Health information on the internet; Health personnel recommendations; Health information in booklets, pamphlets, and brochures; Health information in newspapers and journals; Health recommendations of friends and family	13.6±0.08
Q6	Evaluation of ability in: Filling out medical forms; Explaining health information to others; Using prescribed drugs to the end; Visiting doctor if diagnosed as cancer in the family; Doing yearly medical check-up; Caring about self-health; Ask question anymore about health subject if don't understand; Take care about the Nutritional Value of foods; Having sports activities 3 times each week for 30 minutes; Eating vegetarian and fruits daily; Eating high-fat and fried foods daily; Eating milk, cheese and yogurt daily; Teeth brushing at least one time daily; Adding salt to the food when feeding; Measuring weight 2 times monthly	14.0±0.08
Q7	- Is 80-kilogram weight suitable for a person with 40 years old and 165 cm? - Is 90 milligram/liter fasting sugar suitable for a person with 30 years old? - Is blood pressure (15/9) normal for a person with 60 years old? Having information about the normal range of: Weight; Blood glucose; Blood pressure; How many should 20 milligrams tablets use daily when you must use 80 milligrams every day; How do you use antibiotic capsules 3 times daily	11.4±0.1
Q8	How much do you do these activities: Measuring blood pressure;- Measuring blood glucose; Measuring body temperature; Measuring heart plus; Doing intramuscular injection; Doing perfusion; Doing first aid in emergency; Doing first aid in the road accident	5.9±0.1
Q9	It is easy for me to membership in the sport clubs; I will join to the eco-friendly activities; I will join the meeting about health problems; I will select my political or social candida with respect to the health priorities	11.5±0.1

Particle Swarm Optimization algorithm (PSO)

PSO is an evolutionary optimization algorithm (EA). Evolutionary algorithms take into account a type of metaheuristic algorithms inspired biological evolution. PSO is a population-based algorithm with outstanding performance. PSO algorithm obeys the social behaviors of the population. PSO similar to the other EA has a primary population, mutation, fitness function, and ending conditions. The population named swarm. In the swarm are special points be named particles which are potential solutions [20, 21].

Simply, each particle could move in the problem space with an initial velocity. Velocity in each iteration is updated. Particles try to find the best value of cost function every time. Each particle has a personal best solution ($pBestx_i^k$) which recorded it and all of the particles have a global best solution ($gBest$). Generally, particle velocity is a mixed value of self-velocity (v_i^k) and personal best velocity and global best velocity where i is a particle number and k is an iteration number. Also, particles be shown as $x_i^k =< x_{i,1}^k, x_{i,2}^k, \ldots\ldots, x_{i,d}^k$ where d is dimension space for each particle in kth iteration. The velocity of the next iteration obtains of eq.1 [22].

$$v_i^{k+1} = Wv_i^k x_i^k + C_1R_1(pBestx_i^k - x_i^k) + C_2R_2gBest \quad (1)$$

In eq.1, $C_1, C_2, \in R$ are constants weighting. These coefficients impact on the local and global best solutions role, respectively. $R_1, R_2 \sim U$ [0, 1] are random values. W is inertia weight; whose goal is to supervise the effect of the past velocity of a particle.

After moving a particle in each iteration, cost function evaluates the new cost of each particle and update $pBestx_i$and $gBest$. The algorithm is continued as long as the stop condition is met. In this situation, PSO has found the best solution for the problem.

K-Nearest Neighbor Algorithm (KNN)

The KNN algorithm is an instance-based method. This algorithm is based on that the closest instances have similar attributes [23]. KNN has a simple structure for implementation with few algorithm parameters and runs in an acceptable time. In the training set, the algorithm process finds k-nearest samples to a specific instance. Moreover, it assigns an object to a class by k training samples, where k is the number of training samples. In the classifying operation, target instances are included in a class in which the almost k-nearest samples belong to that class [24].

Fuzzy K-Nearest Neighbor algorithm (FKNN)

Fuzzy K-Nearest Neighbor algorithm is a type of KNN algorithm. In FKNN is be used fuzzy membership function as distance function in KNN [25]. KNN specifies a vector for a special class but FKNN assigns class membership to a sample vector. FKNN classifies data by membership function and this assigning process improves KNN performance. FKNN search the classes of samples set for the K-nearest neighbors and then finding k-samples. This process is similar to the crisp version [26].

Let $X = \{x_0, x_1, \dots, x_N\}$ is instances which belong C classes. Also, let u_{ij} be the membership in the ith class of the jth instance and u (x) is membership of x as defined:

$$u_i(x) = \frac{\sum_{j=1}^{k} u_{ij}(1/\|x-x_j\|^{2(m-1)})}{\sum_{j=1}^{k}(1/\|x-x_j\|^{2(m-1)})} \quad (2)$$

Where m is a real number greater than 1.0 that sets the "strength" of the fuzzy distance function and ||x-x_j|| is the L-norm distance between the set x, and the x_j as the k-th nearest neighbor vector x_j.

Proposed model

One of the most popular methods for feature selection is optimization algorithms. In this study, effective features on HL dataset are selected by PSO. Two classification algorithms (KNN and FKNN) is considered with PSO for finding effective and efficient features. The accuracy of classification algorithm determines the value of fitness function which is Evaluation value for PSO. This process could select impressive features. In this method, the impact of every feature in each class is considered too. Features weights are defined by PSO. So, the chosen features play the most role in classes. In the following, the process of the proposed algorithm is explained.

Firstly, for achieving the suitable results, data is normalized as:

$$X = (x - x_{MIN})(x_{MAX} - x_{MIN}) \quad (3)$$

Where x_{MIN} and x_{MAX} are the most and least values of features.

For finding main health literacy factors in South Khorasan province and its cities be used the proposed method. In the database, An instance is a set of features ($f_1, f_2, .., f_n$) and our method select r features such ($f'_1, f'_2, .., f'_r$) where r < n and ($w_1, w_2, .., w_r$) is continuous value of PSO, In each PSO iterations. Then, the data set is altered with f'_i and w_i ($0 \leq i \leq r$). Classification algorithms classify altered data set with 10-Fold cross-validation.

The pseudocode of proposed model is described as follows:

1) Preprocess: Data normalization
2) Initialization PSO parameters
3) Generating initial PSO population
4) Evaluation: Evaluating the population with the training samples and fitness function.
4-1. Selecting features
4-2. features weighting
4-3. classification with KNN and FKNN by cross validation (10 k-fold)
1) Stop condition: If the optimal model is found, the algorithm stops else go to step 4 (Figure 1).

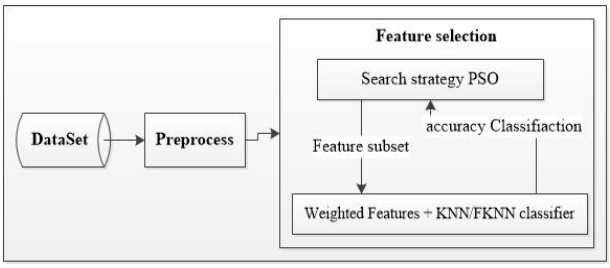

Figure 1) Model Diagram of proposed method

The evaluation metrics like accuracy, precision, and recall are used for measuring performance of model. These are defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

Our used platform for running proposed model was windows 10 OS and Matlab R2017b software. Where population sizes are 10 and 15, maximum iterations are run 50 and 100, W, C1, C2 and Alpha are 1, 2, 2 and 0.065.

Findings

We had evaluated the proposed model with 2 experiments. In the first experiment, introduced data set was divided into 3 groups of HL level (Inadequate, borderline, and adequate) and we had acquired the most effective attributes (Q1-Q9) on the health literacy level of all population with the highest

accuracy. In the second experiment, the data set was divided into 3 groups based on the geographic region (12 cities). Our proposed model had acquired the most effective attributes (Q1-Q9) on the health literacy level of all population with the highest accuracy for each city too.

Two experiments are done with 50, 100 iterations and 10, 15 agents. Data set had classified by 2 algorithms PSO-KNN and PSO-FKNN (Table 2).

Table 2) Simulation results of data set in 3 groups with PSO-KNN and PSO-FKNN

Agent	Iteration	Accuracy	Precision	Recall
PSO-KNN				
10	50	89.78	78.26	86.07
	100	89.22	75.22	84.65
15	50	89.50	73.48	81.43
	100	89.74	77.39	78.76
PSO-FKNN				
10	50	92.02	78.72	92.50
	100	91.31	54.54	92.31
15	50	91.55	77.42	92.31
	100	91.78	70.59	92.31

Experiment 1:

The goal of experiment 1 is identification of main factors (Q1-Q9) on health literacy level in South Khorasan province. Simulation results are showed in table 2. It is considerable in all algorithm runs which PSO-FKNN has had better outcomes than PSO-KNN. The best accuracy was obtained with PSO-KNN 89.78% and this value for PSO-FKNN has been 92.02%. It shows improvement in PSO-KNN with fuzzy logic used in the KNN cost function. Selected domain by our proposed model is like this:

1) PSO-KNN: Q1 (2, 3, 5), Q7 (46, 47, 49, 50, 52, 53), Q8 (59, 61, 62, 63, 64)

2) PSO-FKNN: Q1 (2, 3, 5), Q7 (46, 49, 50, 51, 53), Q8 (59, 60, 61)

These results demonstrate exact question number of each category selected by proposed model. These question numbers selected because they were the most frequent questions in the selected domains.

Experiment 2:

The aim of experiment 2 is identifying major factors (Q1-Q9) on health literacy level in each city of South Khorasan province. Table 3 represents average and best accuracy of PSO-KNN and PSO-FKNN algorithm for each of 11 cities in the South Khorasan province (Table 3).

Finally, selected attributes based on cities are extracted. These major categories selected with two different runs of algorithm: PSO-KNN and PSO-FKNN. In the most cases, the features selected are joined by both methods. Table 3 display features selected for each city.

Furthermore, Statistical analysis of this study was performed using SPSS 22 software. Descriptive statistics including relative and absolute frequency distribution, mean and standard deviation were used. The result Using t-test (Either Mann-Whitney), ANOVA (Kruskal-Wallis), and Chi-square tests at the level of 0.05, A=0, has been analyzed.

Generally, statistical results of this study show HL level in South Khorasan province is inadequate and high percent of the population have had low access to the information resource.

Table 3) Simulation accuracy results of data set with PSO-KNN and PSO-FKNN for 11 cities in South Khorasan province

City Code	PSO-FKNN			PSO-KNN		
	Accuracy		Features	Accuracy		Features
	Mean	Max	Selected	Mean	Max	Selected
1	93.96	95.30	Q2, Q4, Q5, Q7, Q9	91.07	91.68	Q1, Q3, Q4, Q8
2	94.30	95.30	Q2, Q5, Q7, Q9	90.57	91.15	Q5, Q7, Q9
3	95.30	95.97	Q2, Q3, Q5, Q7, Q9	92.18	93.16	Q2, Q3, Q5, Q9
4	96.48	97.99	Q2, Q3, Q7, Q8	92.79	93.97	Q2, Q3, Q5
5	95.30	97.32	Q1, Q6, Q7, Q8	92.15	92.61	Q1, Q3, Q7, Q8, Q9
6	94.80	95.97	Q2, Q5, Q6, Q7	92.85	93.42	Q5, Q6, Q7, Q8
7	96.31	97.32	Q2, Q3, Q5, Q6, Q7	92.05	92.48	Q2, Q3, Q5, Q7
8	94.80	95.97	Q2, Q4, Q6	92.13	93.02	Q2, Q4, Q6
9	95.47	95.97	Q1, Q4, Q5, Q6	91.61	91.95	Q2, Q4, Q7
10	96.31	97.99	Q2, Q4, Q6, Q7, Q9	92.12	92.62	Q4, Q6, Q7, Q8, Q9
11	94.13	95.30	Q2, Q4, Q6, Q7, Q8	91.58	92.20	Q2, Q5, Q8

Discussion

Mostly, outcomes of health literacy researches were presented by statistical analysis. For example in a study, authors investigated health literacy status of diabetic patients by analytical approaches [27].

In contrast, machine learning techniques are used rarely. In 2011, Kandula *et al*. have presented an adaptive testing algorithm for classification of health literacy data [28]. They applied principles of measurement decision theory and Shannon's information theory. They used two data sets which include 52 participants and 60 questions in the first data set and 165 participants, 8 questions in the second data set. Moreover, our proposed method could detect major factors on health literacy status.

The results obtained are showed domains Q1, Q7, Q8 are the most effective domains on HL level of the studied population. According to the statistical results, weakness in access to the information resources (Mean=10.7), ability to interpret, judge and evaluate (Mean=11.4), individual empowerment (Mean=5.9) are cause of coming down health literacy level in the South Khorasan province. In other word statistical analysis supported model results too.

By paraphrasing outcomes of our model, below facts could have been achieved correspondent to the obtained question numbers of proposed model:

1) Low level of access to the information resources like internet, radio, TV, booklets, pamphlets, and etc. have been attained a weak score of level HL in this population (Question numbers 2, 3, 5 in Q1 domain).
2) Lak of popular health information such as awareness of natural weight range, normal glucose range and using drugs in a right way is restrictions in domain Q7 which is the ability to interpret, judge, and evaluate in this study (Question numbers 46, 49, 50, 53 in Q7 domain).
3) Individual empowerment domain has had the lowest average in the studied region. Inadequate knowledge and practical experiments about some self-health aids like measuring blood glucose, body temperature, heart plus, doing profusion, and first aid is a problem in it (Question numbers 59, 60, 61, 63, 64 of Q8 domain).
All of cases mentioned above are areas that must promote for health literacy promotion. Discussion about health literacy promotion in the South Khorasan province cities based on consequence of model outcome is possible, too.
Note: Extracting more details by evaluating the results of the model could be achievable. This is by using finding question numbers in each city and then; acquiring their meanings by Questionnaire.

Conclusion

Today, optimization algorithms have an important role in finding effective factors in a certain data set. In this study, we have used the PSO algorithm for selection of major factors and have used KNN and FKNN classifier on HL data set. In our proposed method, the algorithm identify weighed features that have the main role in classes. According to data have height dimensions, feature selection by traditional analytical methods is complex and difficult. So, the proposed model could achieve this aim by artificial techniques.
Recognizing main agents on Health literacy level of a society basis on geographic regions and finding relationships between these provide a suitable background for improving HL level and lead health scope to use fast and reliable data mining techniques. These could discover hidden knowledge in the health data which is in the hidden level for health experts.

Acknowledgements: The authors would like to extend their gratitude to all persons, who participated in the study and shared their experiences with the researchers.
Ethical permissions: The data used in this article is presented in Ethics Committee of Birjand University of Medical Sciences and are been received Ethic code ir.bums.rec.1394.328.
Conflict of Interest: The authors declare that there is no conflict of interest.
Authors' Contribution: Enayati E. (First author), Methodologist/ Original researcher/Discussion author (50%); Hassani Z. (Second author), Assistant researcher (25%); Moodi M. (Third author), Methodologist/ Statistical analyst /Assistant researcher (25%)
Funding/Support: The present study did not have any funding source.

References

1- Rademakers J, Heijmans M. Beyond reading and understanding: Health literacy as the capacity to act. Int J Environ Res Public Health. 2018;15(8).pii:E1676.
2- The Swedish National Institute of Public Health. Healthy ageing: A challenge for Europe. Stockholm: The Swedish National Institute of Public Health; 2006.
3- Institute of Medicine, Board on Neuroscience and Behavioral Health, Committee on Health Literacy. Health literacy: A prescription to end confusion. Kindig DA, Panzer AM, Nielsen-Bohlman L, editors. Washington DC: National Academies Press; 2004.
4- Vernon JA, Trujillo A, Rosenbaum SJ, DeBuono B. Low health literacy: Implications for national health policy. Report. Washington DC: Department of Health Policy, School of Public Health and Health Services, The George Washington University; 2007.
5- Eichler K, Wieser S, Brügger U. The costs of limited health literacy: A systematic review. Int J Public Health. 2009;54(5):313-24.
6- Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, et al. Health literacy in Europe: Comparative results of the European Health Literacy Survey (HLS-EU). Eur J Public Health. 2015;25(6):1053-8.
7- Tehrani Banihashemi SA, Amirkhani MA, Haghdoost AA, Alavian SM, Asgharifard H, Baradaran H, et al. Health literacy and the influencing factors: A study in five provinces of Iran. Strides Dev Med Educ. 2007;4(1):1-9. [Persian]
8- Hina S, Shaikh A, Abul Sattar S. Analyzing diabetes datasets using data mining. J Basic Appl Sci. 2017;13:466-71.
9- Marinov M, Mosa AS, Yoo I, Boren SA. Data-mining technologies for diabetes: A systematic review. J Diabetes Sci Technol. 2011;5(6):1549-56.
10- Fallah M, Niakan Kalhori SR. Systematic review of data mining applications in patient-centered mobile-based information systems. Healthc Inform Res. 2017;23(4):262-70.
11- Domadiya N, Rao UP. Privacy preserving distributed association rule mining approach on vertically partitioned healthcare data. Procedia Comput Sci. 2019;148:303-12.
12- Chaurasia V, Pal S, Tiwari BB. Prediction of benign and malignant breast cancer using data mining techniques. J Algorithms Comput Technol. 2018;12(2):119-26.
13- Aljumah AA, Ahamad MG, Siddiqui MK. Application of data mining: Diabetes health care in young and old patients. J King Saud Univ Comput Inf Sci. 2013;25(2):127-36.
14- Senturk ZK, Kara R. Breast cancer diagnosis via data mining: performance analysis of seven different algorithms. Comput Sci Eng Int J. 2014;4(1):35-46.
15- Liao SH, Chu PH, Hsiao PY. Data mining techniques and applications - a decade review from 2000 to 2011. Expert Syst Appl. 2012;39(12):11303-11.
16- Jamili Oskouei R, Moradi Kor N, Abbasi Maleki S. Data mining and medical world: Breast cancers' diagnosis, treatment, prognosis and challenges. Am J Cancer Res. 2017;7(3):610-27.

17- Jamili Oskouei R, Farokhbalaghi Sh. Data mining application for exploring the relationship between addiction and depression. Int J Comput Inf Technol. 2016;4(2):43-7.
18- Alonso SG, De La Torre-Díez I, Hamrioui S, López-Coronado M, Barreno DC, Nozaleda LM, et al. Data mining algorithms and techniques in mental health: A systematic review. J Med Syst. 2018;42(9):161.
19- Ioniţă I, Ioniţă L. Applying data mining techniques in healthcare. Stud Inform Control. 2016;25(3):385-94.
20- Sengupta S, Basak S, Peters II RA. Particle swarm optimization: A survey of historical and recent developments with hybridization perspectives. Mach Learn Knowl Extr. 2018;1(1):157-91.
21- Zhang Y, Wang Sh, Ji G. A comprehensive survey on particle swarm optimization algorithm and its applications. Math Probl Eng. 2015;2015:931256.
22- Kennedy J, Eberhart R. Particle swarm optimization. Proceedings of ICNN'95 International Conference on Neural Networks, 27 Nov-1 Dec 1995, Perth, WA, Australia. Piscataway: IEEE; 1995.
23- Salvador-Meneses J, Ruiz-Chavez Z, Garcia-Rodriguez J. Compressed kNN: K-nearest neighbors with data compression. Entropy. 2019;21(3):234.
24- Fan GF, Guo YH, Zheng JM, Hong WC. Application of the weighted k-nearest neighbor algorithm for short-term load forecasting. Energies. 2019;12(5):916.
25- Amirfakhrian M, Sajadi S. Fuzzy k-nearest neighbor method to classify data in a closed area. Int J Math Model Comput. 2013;3(2):109-14.
26- Castillo O, Melin P, Ramírez E, Soria J. Hybrid intelligent system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and neural networks combined with a fuzzy system. Expert Syst Appl. 2012;39(3):2947-55.
27- Singh Sh, Acharya SD, Kamath A, Ullal SD, Urval RP. Health literacy status and understanding of the prescription instructions in diabetic patients. J Diabetes Res. 2018;2018:4517243.
28- Kandula S, Ancker JS, Kaufman DR, Currie LM, Zeng-Treitler Q. A new adaptive testing algorithm for shortening health literacy assessments. BMC Med Inform Decis Mak. 2011;11:52.