Original Article

Comparison of Two Data Mining Techniques in Labeling Diagnosis to Iranian Pharmacy Claim Dataset: Artificial Neural Network (ANN) Versus Decision Tree Model

Ehsan Rezaei-Darzi MSc[1,2], Farshad Farzadfar MD MPH DSc•[2,3], Amir Hashemi-Meshkini PharmD[4,2], Iman Navidi BSc[1,2], Mahmoud Mahmoudi PhD[1], Mehdi Varmaghani MSc[4,2], Parinaz Mehdipour MSc[1,2], Mahsa Soudi Alamdari BSc[2,5], Batool Tayefi MD PhD[2], Shohreh Naderimagham MPH PhD[2,3], Fatemeh Soleymani PharmD PhD[4], Alireza Mesdaghinia PhD[6], Alireza Delavari MD[7], Kazem Mohammad PhD•[1]

Abstract

Background: This study aimed to evaluate and compare the prediction accuracy of two data mining techniques, including decision tree and neural network models in labeling diagnosis to gastrointestinal prescriptions in Iran.

Methods: This study was conducted in three phases: data preparation, training phase, and testing phase. A sample from a database consisting of 23 million pharmacy insurance claim records, from 2004 to 2011 was used, in which a total of 330 prescriptions were assessed and used to train and test the models simultaneously. In the training phase, the selected prescriptions were assessed by both a physician and a pharmacist separately and assigned a diagnosis. To test the performance of each model, a k-fold stratified cross validation was conducted in addition to measuring their sensitivity and specificity.

Result: Generally, two methods had very similar accuracies. Considering the weighted average of true positive rate (sensitivity) and true negative rate (specificity), the decision tree had slightly higher accuracy in its ability for correct classification (83.3% and 96% versus 80.3% and 95.1%, respectively). However, when the weighted average of ROC area (AUC between each class and all other classes) was measured, the ANN displayed higher accuracies in predicting the diagnosis (93.8% compared with 90.6%).

Conclusion: According to the result of this study, artificial neural network and decision tree model represent similar accuracy in labeling diagnosis to GI prescription.

Keywords: Artificial neural networks, decision tree, insurance prescriptions

Cite this article as: Rezaei-Darzi E, Farzadfar F, Hashemi-Meshkini A, Navidi I, Mahmoudi M, Varmaghani M, Mehdipour P, Soudi Alamdari M, Tayefi B, Naderimagham Sh, Soleymani F, Mesdaghinia A, Delavari A, Mohammad K. Comparison of Two Data Mining Techniques in Labeling Diagnosis to Iranian Pharmacy Claim Dataset: Artificial Neural Network (ANN) Versus Decision Tree Model. *Arch Iran Med*. 2014; **17(12):** 837 – 843.

Introduction

Priority setting is one of the fundamental goals of healthcare systems worldwide for equitable and efficient allocation of scarce resources. To meet this goal, policy makers must have access to sufficient information regarding various health-related aspects of their population in national and sub-national levels.[1] These aspects encompass disease patterns,[2] the equity of services,[3] the quality and cost of routine interventions for each disease, adherence and safety issues after intervention.[4]

Various health-related databases may be accessible for this purpose in different countries depending on the structure of the health system. In the burden of diseases studies and especially in Iran, the NASBOD (National and Sub-national Burden of Diseases) project,[5] as a comprehensive study to measure the pattern of diseases in Iran, death and disease registries can be used as an information source of diseases[6–9]; however, they are often limited to only certain types of diseases. Hospital database is another valuable information source[9]; nevertheless. It cannot be considered as a reliable source for outpatient health conditions. Searching for more comprehensive data, administrative healthcare databases such as medical prescriptions claims can be found as one of the valuable sources for getting knowledge about the pharmacoepidemiologic, economic and social patterns of health conditions and quality of care in a health system. In countries with approximately universal health insurance coverage, pharmacy claim data could be a golden proxy for a comprehensive health registry nationwide. As we have access to demographic and socioeconomic information of entitled population of insurance funds, this source can be considered as a retrospective data for many further studies in various fields.

Prescription and other health related databases for estimation of the prevalence of diseases have been used in some studies.[10–13]

Authors' affiliations: [1]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran. [2]Non-communicable Diseases Research Center, Endocrinology and Metabolism Population Science Institute, Tehran University of Medical Sciences, Tehran, Iran. [3]Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran. [4]Department of Pharmacoeconomics and Pharmaceutical Administration, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran. [5]Department of Network Science and Technology, Faculty of New Sciences and Technologies, University of Tehran, Iran. [6]Department of Environmental Health Engineering, School of Public Health and Institute of Public Health Research, Tehran University of Medical Sciences, Tehran, Iran. [7]Digestive Oncology Research Center, Digestive Disease Research Institute, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran.

•**Corresponding authors and reprints:** Kazem Mohammad PhD, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran. E-mail: mohamadk@tums.ac.ir.

Farshad Farzadfar MD MPH DSc, Non-communicable Diseases Research Center, Endocrinology and Metabolism Research Institute, Tehran University of Medical Sciences, Tehran, Iran. Address: No. 4, 4th Floor, Ostad Nejatollahi St., Enqelab Ave., Tehran 1599666615, Iran. Telefax: 98-21-88913543, E-mail: f-farzadfar@tums.ac.ir.

Accepted for publication: 12 November 2014

There are many published studies about how to validate the recorded diagnosis of these prescriptions by different approaches.[14–18] However, in Iran and probably many developing countries, using prescription claims faces many challenges. One of these challenges in Iran is the common prescribing habits of physicians: they evade writing the diagnosis of disease on the prescriptions, preventing any further legal trouble in case of mistakes; therefore, there is no recorded diagnosis for prescriptions. To use this valuable and efficient database in pharmaco-epidemiologic and utilization studies, this obstacle should be circumvented.

In an attempt to find an approach for labeling diagnosis to Iranian prescription, in this study, we evaluated and compared two machine learning techniques, including decision tree and neural network models in term of accuracy. In the present study, we focused on main gastrointestinal diseases to achieve an acceptable pathway for more disease classes and developing a method for more comprehensive studies in future. We selected GI diseases because of their lower overlap with other diseases in terms of prescription items.

Materials and Methods

In this study, we intended to evaluate the accuracy of decision tree and artificial neural network models among machine learning methods. This study was conducted in three parts: data preparation, training-testing phase. The entire process is presented in details in Figure 1.

Data source

A database consisting of 23 million pharmacy insurance claim records from three main Iranian health insurance organizations (SSIO, MSIO, and AFMSIO), from 2005 to 2013, was used, which was provided by the Iran Food and Drug Organization (FDO). The database contained information code, including the name of medicine, potency, dosage form, number of each medicine, total price, insurance type and date of prescription, but no code related to the physician diagnosis.

Sample selection

As we aimed to focus on gastrointestinal disorders, an inclusion criterion of containing at least two gastrointestinal related medicines was defined for sampling and then in the second step, one thousand prescriptions were randomly selected.

Training process

The selected prescriptions were assessed by a physician and a pharmacist separately and assigned a diagnosis. In order to avoid more complexity in the model, the experts ignored the comorbidities. Then, the diagnoses were evaluated and if there were disagreements, they were solved by consulting a third expert (a physician). We used both physician and pharmacist intentionally to track all kinds of common prescribing patterns for each disease, given their different experience and perspectives. In addition, they were asked not to focus on their own knowledge and review all available guidelines and all possible medical treatment types using Internet and textbooks prior to the training phase.

All diagnosed diseases are summarized in Table 1. During the training process, two diseases were eliminated from the list because it was inconceivable to detect them by insurance claim data according to experts' idea: constipation and gallstones. The main reason was that the specific medicines of these diseases are not recorded by pharmacies in the claims, mainly because of not being included in reimbursement lists of insurance.

To reduce the number of input variables, the model was trained based on the existing medical classes in each prescription rather than the medicine names. For this purpose, all medicines with possible indication in GI disorders were extracted from textbooks, guidelines and Internet and were then classified in 18 groups according to their effects and mechanism of action (Table 2). A total of 330 prescriptions were assessed and used to train and test the models simultaneously.

Artificial neural network

An artificial neural network (ANN) can be defined as a sophisti-

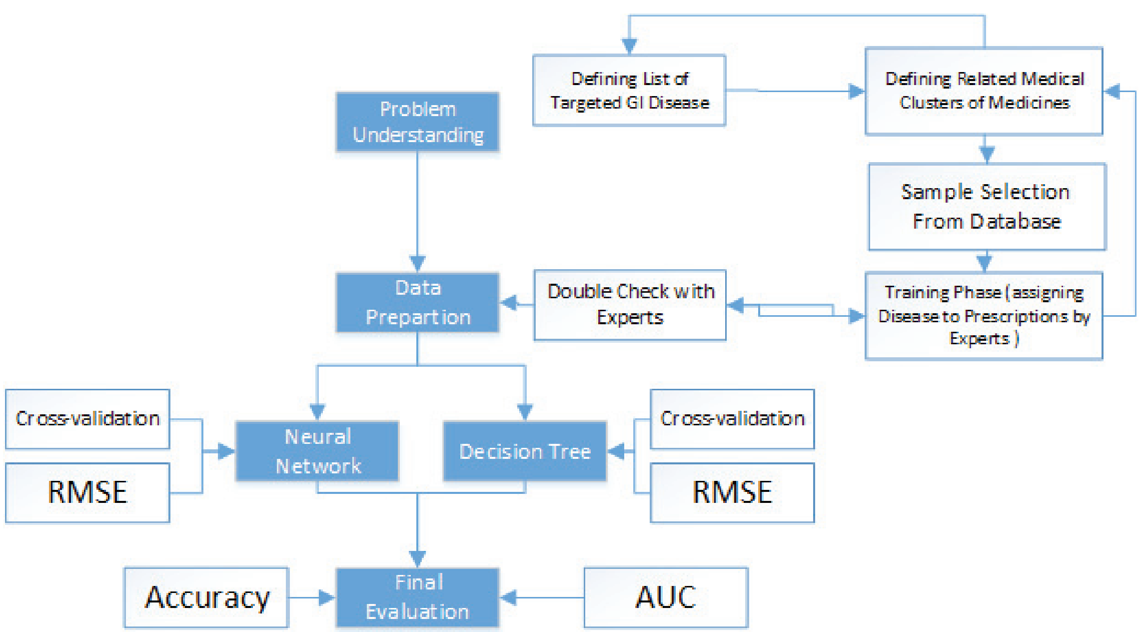

Figure 1. The whole process of the study.

Table 1. The targeted diseases for prediction by data mining models.

Investigated diagnosis	Prior class (%)
Helicobacter Pylori	10.30
Gastritis & Dyspepsia	20.61
Gastroenteritis	25.15
Hemorrhoid & Fissure	8.18
IBD (Crohn's Disease & Ulcerative Colitis)	9.09
IBS	9.09
Parasitic & Fungal infectious	17.58

Table 2. The medical classification for being used in models as input.

Medical cluster	Sub-cluster medicines
Proton pomp inhibitors (PPIs)	Omeprazole, Pantoprazole, …
H2 Blockers	Ranitidine, Famotidine, …
Antibiotics	Amoxicillin, Metronidazole, Bismuth substrate, Clarithromycin, …
Antispasmodics	Hyoscine, Dicyclomine, …
Prokinetic Agents	Metoclopramide
Anti-acid preparations	ALMG
Laxative agents	Bisacodyl, herbal products
Infusion solution	All regular infusion products
Ant parasitic agents	Mebendazole, Albendazole
Digestive Agents	All available Digestive products in Iran
Gastro-protective agents	Sucralfate
Anti-hemorrhoids & Anti-fissure products	Antihemoroid
Antidepressants	Fluoxetine, Imipramine
Immunosuppressive agents	Azathioprine, Methotrexate, ...
Anti-inflammatory agents	Sulfasalazine, Prednisolone, …
Beta Blockers	Propranolol
Supplements	Calcium, Folic acid

cated mathematical model for processing the information of non-linear and complex problems, in which the relationships between various variables are unknown or very complicated. ANNs development is inspired by the structure of human neural system[19] and they are currently widely used in different areas of healthcare research.[20–23] There are many types of ANNs[24]; however, the majority of them have the same general structure, including an "input layer", one or more "hidden layers" (based on the complexity of the model), and an "output layer".[25] The number of neurons comprising input and output layers is equal to the number of predictors and class labels in training data but to reach the optimum number of hidden nodes, many networks must be tried with different units of hidden layer and the net with the minimum error term be chosen. Neurons in each layer are connected to the neurons in the next layer with some weights. Similar to human brain system, the ANN models function is based on both «learning» and «generalizing» abilities.[26] A supervised multilayer perceptron is a feed-forward artificial neural network learning by back-propagation algorithm. In the learning phase, the model transforms the input data through different layers and optimizes the link weights by backward propagation rule to achieve the minimum difference between the predicted and actual values of output.[27]

The feed-forward step can be described for a multi-class response (Y_k) in the following formula.[28]

$$\hat{y}_k = f_k(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{o}, \boldsymbol{x}_0, \boldsymbol{o}_{0k}, \theta)$$

$$= f\left(\sum_{j=1}^{h} o_{kj} \cdot g\left(\sum_{i=1}^{p} w_{ij} x_i + x_{0j}\right) + o_{0k}\right)$$

Where $f(.)$ and $g(.)$ are the logistic activation functions for the output and hidden layer respectively. x is the vector of p attributes or predictors. Vectors of input and hidden weights are called w and o. Constants are considered as biases in the two connection weights (x_{oj} and o_{ok}) .

In the present study, the input layer was defined as our 18 medical classes and the output layer contained the 7 diagnoses, labeled by the experts. The model was trained by back-propagation algorithm to reach the minimum errors changing the tuning parameters and finally it was fixed to the learning rate: 0.3 (at each iteration controls the size of weight and bias changes), momentum value: 0.2 (consider already existing information in current nodes) and number of epochs: 500 (the maximum number of iteration).

Decision tree model

Decision tree is a decision supportive tool, which is widely used as a predictive technique of data mining in healthcare.[29] Decision tree is a simple technique to determine patterns in numerous classification problems, particularly in massive dataset including complex information. Furthermore, its simple approach helps the policy makers to understand it easily. In this model, the relationship between data is represented in a tree structure, starting from a root node (an input attribute detected by model) to different nodes (representing the rules) via multiple branches and finally ending in some terminal nodes (outcome attribute). Decision trees are generated based on an algorithm to split dataset into branches. Indeed, discovery of the rules for making branches depends on the method of extracting the relationship between input and output attributes. In the current study, the C4.5 algorithm[30] was chosen due to its popularity and acceptability in data mining literature. According to the input attributes, each rule assigns an instance (observation) to a node in one branch and this approach would be repeated several times to end the hierarchy of branches in some terminal nodes which are called leaves. The assignment rules follow a mutually exclusive approach; therefore, each leaf only contains a unique path. More details about decision tree model are represented elsewhere.[31] In this study, the 7 diagnoses were considered as leaves, to which the instances (prescriptions) were assigned by some rules that made the hierarchical links between the input attributes (medical classes). Furthermore, the minimum number of instances per leaf was determined as one percent of prescriptions to avoid the effect of small records on rule generation.

Performance assessment of models

Stratified k-fold cross-validation

In order to evaluate and generalize data mining model to a new dataset and limiting the effect of over-fitting, a cross-validation technique was applied in this study.[32] By this technique, the model would be developed using a complementary subset of sample and then run on the remaining to evaluate its accuracy. In the current study, the dataset was divided into a training set (n = 297) and a test set (n = 33) for 10 separated times. Different models were evaluated by changing some special characteristics (i.e., tuning parameters in data mining algorithms) and finally they were compared in terms of performance by the root mean square error (RMSE) criteria. In k-fold cross validation, this approach would be repeated several times by different subsets of dataset to minimize the bias related to random selection of the training subset.[33] In K-fold cross-validation, k equal size and randomly selected subsamples from original sample is used to repeat the procedure. In addition, in the stratified k-fold cross-validation, the folds are selected according to the equality of mean response value in all folds. For comparative assessment of the prediction accuracy of both data mining methods, ROC area (the area under the receiver operating characteristic) and accuracy were taken into account.

Sensitivity and specificity

To evaluate the prediction accuracy of each models, the authors also considered TP rate (sensitivity) and TN rate (specificity).

The software

The WEKA software (Waikato Environment for Knowledge Analysis version 3.7.9) was used to implement and run both data mining models.

Result

General findings

The proportion of each disorder was found to be 10.3%, 20.61%, 25.15%, 8.18%, 9.09%, 9.09%, and 17.58%, respectively for *Helicobacter pylori* infection, Gastritis & Dyspepsia, Gastroenteritis, Hemorrhoid & Fissure, IBD, IBS, and Parasite and Fungal infections. Moreover, Gastroenteritis with 83 (25.15%) instances and IBD with 27 (8.18%) of cases comprised the biggest and smallest proportion of GI disorders of pharmacy claim data, respectively.

Models performance

In *ANN model a* net architecture with eighteen nodes in the input layer and seven nodes in the output layer along with sigmoid activation function in both hidden and output layer was constructed (see figure 2), given its minimum RMSE: 0.2149.

Figure 3 represents the decision tree structure with a size of 35 and 18 leaves. By the C4.5 algorithm, 11 out of 18 medical clusters were selected in the variable selection dataset but in the MLP, all available input attributes were included. The 15th medical class was the first prediction attribute of decision, in which any decisions (or prediction path) were initiated. Additionally, the third and ninth classes were used most for making tree (Figure 3). In the selection phase, the RMSE index of the decision tree was 0.2069.

Generally, two methods had very similar accuracies. Considering the weighted average of TP rate (sensitivity) and TN rate (specificity) considered, decision tree had slightly higher accuracy in its ability for correct classification (83.3% and 96% versus 80.3% and 95.1%, respectively). However, when the weighted average of ROC area (AUC between each class and all other classes) was measured, the ANN displayed higher accuracies in predicting the diagnosis (93.8% compared with 90.6%). The ROC curves are displayed in Figure 4, representing seven comparisons of GI disorders AUC for two data mining methods.

Discussion

This study represented the acceptability of prediction accuracy of both ANN and decision tree models in assigning diagnosis to GI disorder-related prescription in the context of Iran. It means that using either of these two data mining methods could provide valid estimation about the number of prescriptions for GI disorders using pharmacy claim data. Furthermore, it can be useful for the NASBOD study in predicting the epidemiological data of outpatient diseases.[5]

Despite similar performance in predicting GI diagnosis, some practical terms of both methods (speed, interpretability and accuracy) are considerable. ANN models have some drawbacks such as complexity of designing network, lack of relative relevance of independent variables, time-consuming process and less interpretability.[32] On the other hand, the decision tree has some advantages such as easy programming for computer systems, being easy to understand (trees are very interpretable) and providing visual representations of data. Nevertheless, neural networks could be updated with gradient descent for new dataset, unlike the decision tree, which requires inherently batch-learning algorithms.

In this study, the ROC area was used to evaluate the accuracy of each model. The ROC curve is a well-known approach for evaluation of two class pattern recognition problems and given our multi-class problem, the researchers used a weighted average of ROC area to adjust to our model. There are more complicated ways to deal with multi-class problems, including volume under the ROC hyper-surface (VUS) and Pareto front method.[34–35] We are going to include these aspects in our analysis on this dataset in future.

In Iran, using the prescription databases for this purpose is sub-

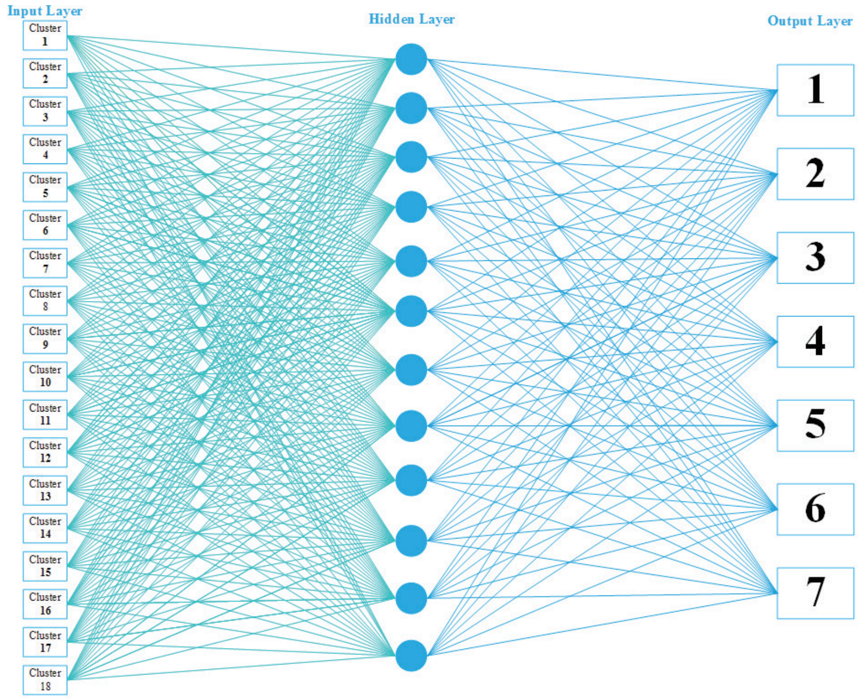

Figure 2. Architecture of ANN model for prediction of GI diseases

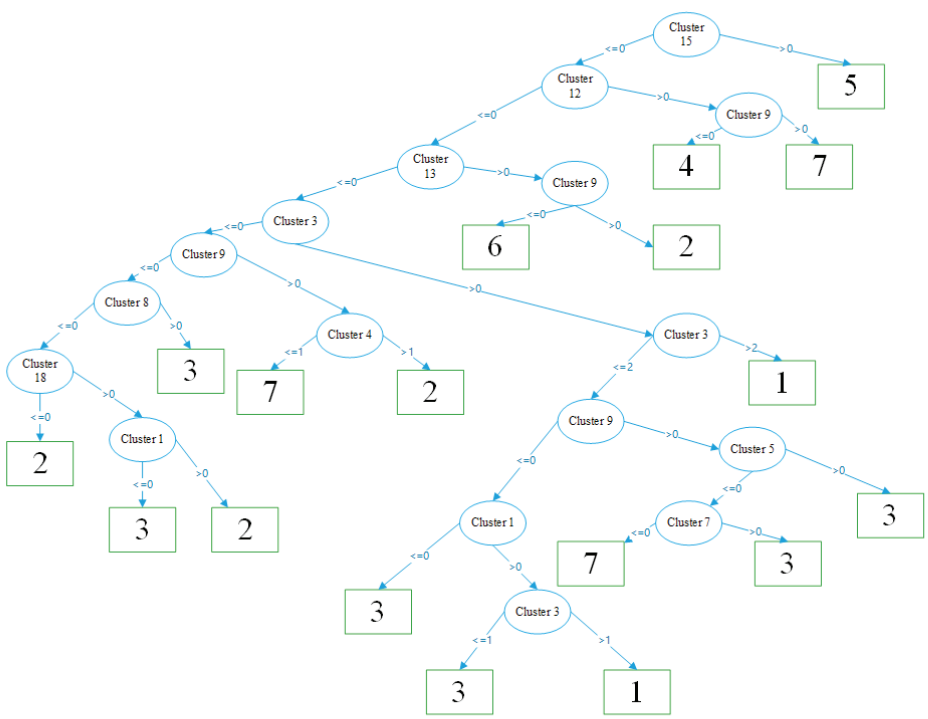

Figure 3. The result of decision tree model for prediction of GI diseases.

ject to many limitations. The fragmented structure of the health insurance system in which there are many funds covering a proportion of population with different coding system is a big obstacle that undermines the quality of data for data mining. Indeed, the regulation and coverage list of these separated funds are different, leading to incomplete recording of uncovered medicines. Another limitation was that we did not have access to the codes of physicians, and sex and age of patients in prescriptions that could undermine the quality of data and therefore the accuracy of prediction.[36]

This was the first study trying to develop an approach for assigning diagnosis to Iranian prescriptions using data mining techniques. Considering the value of pharmacy claim dataset in epidemiological studies, health system assessment and policy evaluation, this approach may be used by other low and middle-income countries, in which disease diagnoses are not recoded in

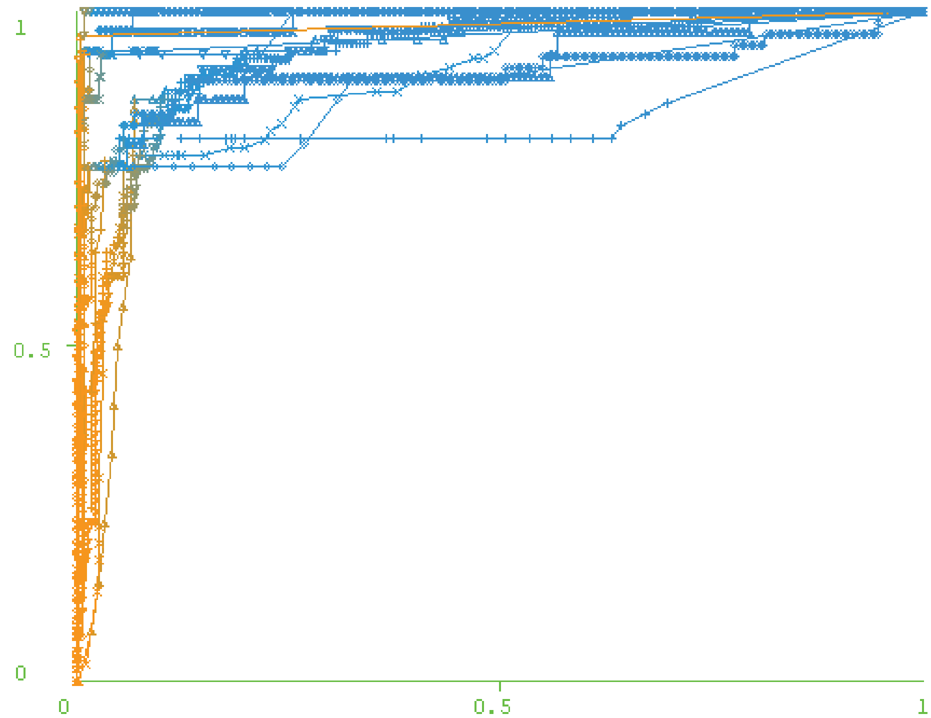

Figure 4. ROC curves for multilayer perceptron and C4.5 algorithm in predicting seven GI disorders.

prescription data for any reason. The results of this study may be used for estimating the economic indices including cost of illness analysis and equity of accessibility to medicines. Using this approach, the prescription database could also be used next to other data sources in estimation of national and sub-national burden of gastrointestinal diseases as well as other disease groups in Iran.[37]

As the result of this study revealed, the artificial neural network and decision tree model represent similar accuracy in labeling diagnosis to GI prescription and may be used in a more comprehensive approach for estimating epidemiologic and economic indices including prevalence and cost of illness.

Competing interests

The authors declare that they have no competing interests.

Authors' contribution

General designing of paper: Farshad Farzadfar, Kazem Mohammad, Ehsan Rezaei Darzi, Amir Hashemi meshkini.

Designing of models: Ehsan Rezaei Darzi, Farshad Farzadfar, Kazem Mohammad, Amir Hashemi Meshkini, Iman Navidi.

Writing primary draft: Amir Hashemi Meshkini, Ehsan Rezaei Darzi.

Manuscript revision: Ehsan Rezaei Darzi, Farshad Farzadfar, Amir Hashemi-Meshkini, Iman Navidi, Mahmoud Mahmoudi, Mehdi Varmaghani, Mahsa Soudi Alamdari, Shohreh Naderimagham, Fatemeh Soleymani, Parinaz Mehdipour, Batool Tayefi, Alireza Mesdaghinia, Alireza Delavari, Kazem Mohammad.

Acknowledgment

Our research was funded by Iran's Ministry of Health and Medical Education of Islamic Republic of Iran and Setad-e-Ejraie Farmane Imam. This study was part of the MSc thesis of Ehsan Rezaei Darzi in biostatistics. The authors would like to thank all colleagues in the non-communicable disease research center (NCDRC) and appreciate the collaboration of organizations and researchers that provided needed data for this project, especially the Food and Drug Organization (FDO). We would also like to express thanks to Dr. Masoud Moradi for his precise editing of the text and Ms. Rosa Haghshenas for her efforts in managing coordinative and administrative processes. We would also like to express our thanks to Reza Esfandiari, PhD Candidate in Translation Studies at University Sains Malaysia, for his precise edit of manuscript.

References

1. Wright J, Williams R, Wilkinson JR. Health needs assessment: Development and importance of health needs assessment. *BMJ*. 1998; **316**: 1310.
2. Armenian HK, Shapiro S. *Epidemiology and Health Services*. Oxford University Press; 1998.
3. Allin S. Does equity in healthcare use vary across Canadian provinces. *Healthc Policy*. 2008; **3:** 83 – 99.
4. Vermeire E, Hearnshaw H, Van Royen P, Denekens J. Patient adherence to treatment: three decades of research. A comprehensive review. *J Clin Pharmather*. 2001; **26:** 331 – 342.
5. Farzadfar F, Delavari A, Malekzadeh R, Mesdaghinia A, Jamshidi H, Sayyari A, et al. NASBOD 2013: design, definitions, and metrics. *Arch Iran Med*. 2014; **17(1):** 7 – 15.
6. Levine MN, Julian JA. Registries that show efficacy: good, but not

good enough. *JCO*. 2008; **26:** 5316 – 5319.
7. Modirian M, Rahimzadeh S, Cheraghi Z, Khosravi A, Salimzadeh H, Kompani F, et al. Quality evaluation of national cancer registry system in iran: study protocol. *Arch Iran Med*. 2014; **17(3):** 193 – 197.
8. Mohammadi Y, Parsaeian M, Farzadfar F, Kasaeian A, Mehdipour P, Sheidaei A, et al. Levels and trends of child and adult mortality rates in the Islamic Republic of Iran, 1990–2013; protocol of the NASBOD study. *Arch Iran Med*. 2014; **17(3):** 176 – 181.
9. Niakan KS, Tayefi B, Noori A, Mearaji M, Rahimzade S, Zandian E, et al. Inpatient data, inevitable need for policy making at national and sub-national levels: a lesson learned from NASBOD. *Arch Iran Med*. 2014; **17(1):** 16 – 21.
10. Zhao Y, Connors C, Wright J, Guthridge S, Bailie R. Estimating chronic disease prevalence among the remote Aboriginal population of the Northern Territory using multiple data sources. *ANZJPH*. 2008; **32:** 307 – 313.
11. Wiréhn A-BE, Karlsson HM, Carstensen JM. Estimating disease prevalence using a population-based administrative healthcare database. *Scand J Public Health*. 2007; **35:** 424 – 431.
12. Wigertz A, Westerling R. Measures of prevalence: which healthcare registers are applicable? *Scand J of Public Health*. 2001; **29:** 55 – 62.
13. Maio V, Yuen E, Rabinowitz C, Louis D, Jimbo M, Donatini A, et al. Using pharmacy data to identify those with chronic conditions in Emilia Romagna, Italy. *J Health Serv Res Policy*. 2005; **10:** 232 – 238.
14. Muhajarine N, Mustard C, Roos LL, Young TK, Gelskey DE. Comparison of survey and physician claims data for detecting hypertension. *J Clin Epidemiol*. 1997; **50:** 711 – 718.
15. Robinson JR, Young TK, Roos LL, Gelskey DE. Estimating the burden of disease: comparing administrative data and self-reports. *Med Care*. 1997; **35:** 932 – 947.
16. Worth R, Mytinger R. Medical insurance claims as a source of data for research: accuracy of diagnostic coding. *Hawaii Med J*. 1996; **55:** 9 – 11.
17. Katz JN, Barrett J, Liang MH, Bacon AM, Kaplan H, Kieval RI, et al. Sensitivity and positive predictive value of Medicare Part B physician claims for rheumatologic diagnoses and procedures. *Arthritis Rheum*. 1997; **40:** 1594 – 1600.
18. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol.* 2004; **57:** 131 – 141.
19. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev.* 1958; **65:** 386.
20. Anagnostou T, Remzi M, Lykourinas M, Djavan B. Artificial neural networks for decision-making in urologic oncology. *Eur Urol*. 2003; **43:** 596 – 603.
21. Lancashire L, Schmid O, Shah H, Ball G. Classification of bacterial species from proteomic data using combinatorial approaches incorporating artificial neural networks, cluster analysis and principal components analysis. *Bioinformatics*. 2005; **21:** 2191 – 2199.
22. Zupan J, Gasteiger J. *Neural Networks in Chemistry and Drug Design*. John Wiley & Sons, Inc.; 1999.
23. Lisboa PJ, Taktak AF. The use of artificial neural networks in decision support in cancer: a systematic review. *Neural Net*. 2006; **19:** 408 – 415.
24. Gevrey M, Dimopoulos I, Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecol Model*. 2003; **160:** 249 – 264.
25. Wei JT, Zhang Z, Barnhill SD, Madyastha KR, Zhang H, Oesterling JE. Understanding artificial neural networks and exploring their potential applications for the practicing urologist. *Urology*. 1998; **52:** 161 – 172.
26. Amato F, López A, Peña-Méndez EM, Vaňhara P, Hampl A, Havel J. Artificial neural networks in medical diagnosis. *J Appl Biomed*. 2013; **11:** 47 – 58.
27. Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J Pharm Biomed Anal*. 2000; **22:** 717 – 727.
28. Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I, de Mendonça A. Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Res Notes*. 2011; **4:** 299 313 .
29. Milovic B. Prediction and decision making in health care using data mining. *IJPHS*. 2012; **1:** 69 – 78.
30. Quinlan JR. *C4. 5: programs for machine learning*. Morgan kaufmann; 1993.
31. Podgorelec V, Kokol P, Stiglic B, Rozman I. Decision trees: an overview and their use in medicine. *J Med Syst*. 2002; **26:** 445 – 463.
32. Parsaeian M, Mohammad K, Mahmoudi M, Zeraati H. Comparison of Logistic Regression and Artificial Neural Network in Low Back Pain Prediction: Second National Health Survey. *Iranian J Publ Health*. 2012; **41:** 86 – 92.
33. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI*. 1995: **14(2):** 1137 – 1143.
34. Landgrebe T, Duin R, editors. A simplified extension of the area under the ROC to the multiclass domain. *Seventeenth Annual Symposium of the Pattern Recognition Association of South Africa*; 2006.
35. Fernandez Caballero JC, Martínez FJ, A simplified extension of the area under the ROC to the multiclass domain. Seventeenth annual symposium of the pattern recognition association of South Africa. *IEEE Transactions on Neural Networks*. 2010; **21:** 750 – 770.
36. Burke HB, Goodman PH, Rosen DB, Henson DE, Weinstein JN, Harrell FE, et al. Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer*. 1997; **79:** 857 – 862.
37. Salimzadeh H, Ardeshir LF, Abedian S, Kalantar MS, Malekzadeh M, Mohaghegh H, et al. The trend of national and sub-national burden of gastrointestinal and liver diseases in Iran 1990 to 2013; study protocol. *Arch Iran Med*. 2014; **17(1):** 33 – 53.