Performance Evaluation of Decision Tree Classifiers for the Prediction of Response to treatment of Hepatitis C Patients

AbuBakr Awad
Department of Computer Science
Faculty of Computers and Information, Cairo University
Cairo, Egypt
+201000195098
bakr.awad@gmail.com

Mahasen Mabrouk, Tahany Awad,
Naglaa Zayed, Sherif Mousa, Mohamed Saeed
Department of Endemic Medicine and Hepatology
Faculty of Medicine, Cairo University
Cairo, Egypt
mahasen.mabrouk@gmail.com,
tahany.awad@gmail.com,
naglaazayed@yahoo.com,
sheriefmusa@yahoo.com,
m7mad_said@yahoo.com

ABSTRACT

This study included 2962 Egyptian patients with chronic hepatitis C virus (HCV) infection. Different decision-tree models were used to explore baseline predictors of response to Peginterferon plus Ribavirin therapy to discriminate HCV patients who are likely to respond. We have developed simple software that generates different possible combination of parameters for each model; using this software we were able to assort the decision-tree model according to the best performance, and to find the best classifier. The three models were comparable as regards to accuracy (about 69%); however REP Tree has shown fast and well performance for classification on medical data sets of increased size. Various pre-treatment decision tree algorithms have demonstrated that low level of Alpha-Fetal Protein (AFP) is associated with high response rate; and has the prospective to support clinical decisions regarding the proper selection of patients for therapy without imposing extra costs for additional examinations.

Categories and Subject Descriptors

H.2.8 [**DATABASE MANAGEMENT**]: Database Applications – *Data mining;* J.3 [**LIFE AND MEDICAL SCIENCES**]: Medical information systems.

General Terms

Experimentation. Algorithms.

Keywords

AFP, gender, decision tree, C4.5, CART, REP tree, HCV, pegylated interferon, ribavirin.

1. INTRODUCTION

The population of Egypt has a heavy burden of liver disease due to chronic infection with hepatitis C virus (up to 10%). The high prevalence of HCV and its eventual consequences that can develop into potentially fatal cirrhosis and hepatocellular carcinoma have motivated the Egyptian Ministry of Health and Population (MOHP) to execute a national approach against this major health burden [1]. Since 2006, the Egyptian national strategy for the control of viral hepatitis has implemented more efficient, better-tolerated, cost-effective strategies. Patients with chronic HCV can now receive specialized medical service and treatment according to standardized guidelines in one of the treatment centers which were initiated in several governorates.

The current standard of care in chronic HCV is combined Pegylated Interferon Alpha (Peg-IFN) plus Ribavirin (RBV) therapy which has to be given for 48 weeks. Taking into account the moderate efficacy of IFN/RBV therapy in patients with HCV, the long-term duration of treatment (i.e. 48 weeks), the relative high incidence of severe side effects and its high cost, it is not preferable to initiate treatment in all patients infected with HCV [2, 3]. Thus there is a need for an innovative technique to identify the characteristics and predictive factors of patients who will respond to treatment.

Data mining is a method of predictive analysis which can explore tremendous volumes of data to discover hidden patterns and relationships in highly complex dataset and thus can enable the development of predictive models. Data Mining with Decision trees plays a vital role in the field of medical diagnosis and prediction [4]. The decision-tree analysis tool; a core component of data mining analysis; is a tree-shaped structure that represent decision sets which can classify data. Decision trees are so popular because they produce human readable classification rules and easy to interpret than other classification methods

Decision-tree analysis is a favored technique for building understandable and predictive models that could facilitate the simple allocation of patients into subgroups which can identify the possibility of outcome. It was applied for the prediction of early virological response (EVR) to PEG-IFN and RBV combination therapy in chronic hepatitis C [5] and sustained virological response (SVR) [6].

PervasiveHealth 2014, May 20-23, Oldenburg, Germany

DOI 10.4108/icst.pervasivehealth.2014.254946

There are several ways to evaluate performance of model, each evaluation method consider its point of view. The evaluation matrix used to evaluate the performance of algorithms, is usually based on: Accuracy, Precision, Recall, F-Score, ROC curve, and time complexity.

In medical field, recall (sensitivity) refers to the proportion of people with disease who have a positive test result, while precision (specificity) refers to the proportion of people without disease who have a negative test result, while the area under the curve (AUC) serves as an indicator of the overall performance of the algorithm.

2. OBJECTIVE

- Evaluating the performance of decision tree classifiers to select the best algorithm to determine the important predictors of therapeutic outcome to therapy in chronic HCV Egyptian patients.
- Helping decision maker to select a strategy that will help minimize the overall treatment cost.

3. METHODOLOGY

3.1 Study population

This retrospective study included 2962 adult Egyptian patients with chronic HCV infection of both sexes. All patients had been treated with Peg-IFN alpha 2a or 2b plus weight-based RBV between the years 2008 and 2012 at Cairo-Fatemic Hospital, Egypt. The study was approved by the ethical committee of the Egypt Ministry of Health and Population (MOHP).

3.2 Feature Selection

Feature selection based on routine data was used to minimize the dimensionality of the problem to be amenable for analysis. The features were classified as categorical or numerical (see Table 1).

Table 1. Table captions should be placed above the table

	Attribute	**Represented As**
1	Gender	Categorical (Male, Female)
2	Age	Numeric
3	Body Mass Index (BMI)	Numeric
4	Glucose	Numeric
5	Albumin	Numeric
6	Alkaline phosphatase	Numeric
7	Aspartate transaminase	Numeric
8	Alanine transaminase	Numeric
9	Total bilirubin	Numeric
10	White blood cells	Numeric
11	Hemoglobin	Numeric
12	Platelet count	Numeric
13	Prothrombin	Numeric
14	Antinuclear antibody	Categorical (Negative, Positive)
15	HCV_RNA_IU_ML	Numeric
16	Alfa fetoprotein	Numeric
17	Histological activity	Categorical (A1, A2, A3)
18	Fibrosis	Categorical (F1, F2, F3, F4)
19	Final Status	Categorical (Responders, non-Responder)

3.3 Experiment Structure

In the present study, we used Weka software (a collection of machine learning algorithms for data mining tasks in Java language) implementation of:

- C4.5, which was published by Ross Quinlan in 1993 [7].
- CART which was published by L.Breiman in 1984 [8].
- REP Tree as a Fast decision tree learner (FDTL) which was published by Mehta, Agrawal, & Rissanen in 1996 [9]

The algorithms used above are examples of commonly used decision trees of high accuracy in medical classification as illustrated before in the review. They handle both categorical and continuous attributes and also handle missing values to build a decision tree. These models built were used to explore baseline predictors of response to PEG-IFN plus RBV therapy among clinical, biochemical, virological and histological pretreatment variables and to select a pre-treatment algorithm to discriminate HCV patients who are likely to respond to PEG-IFN plus RBV therapy.

The experiments above run on the same machine specs in a machine with a processor model Intel® Core I5-460M, 2.53 MHz with 3MB Cache and with memory module of 4GB.

3.4 Validation of the Decision Tree

The calculated algorithms were validated using the k-fold cross-validation approach with test mode value 10-folds. This approach is considered to be a powerful methodology to overcome data over-fitting. Briefly, the original sample is divided into k sub-samples. The cross-validation process is repeated k times (folds) and each of the k sub-samples is used once as the validation data. The k results obtained from the k-folds could then be averaged to produce a single estimation of model performance.

3.5 Algorithms Optimization

We have developed simple software that generates different possible combination of parameters (e.g. confidence factor, minimum number of instances per leaf) that affect the performance of the algorithm. Using this software we were able to assort the decision-tree model according to the best performance, and to find the best classifier for prediction of response to HCV treatment.

4. RESULTS AND DISCUSSION

What is unique to the present study is the big cohort of 2962 patients (2416 males/546 females) chronically infected with HCV. End of treatment response at week 48 was 61%, and sustained virological response at week 72 was 54%.

4.1 Performance results of models using standard parameters

The three models were comparable as regards to accuracy about 69%. REP Tree was much faster as the time taken to build model was only 0.07 seconds compared to 0.25 seconds needed to build C4.5 tree. The CART tree was very slow compared to REP tree or C4.5, the time taken to build the CART model was 11.19 seconds. We concluded that CART decision tree is too slow to be included in the comparison study (see Table 2).

Table 2. Comparison between C4.5, CART, REP Tree after using standard parameters

	C4.5	CART	REP tree
Time taken to build model	0.25 seconds	11.19 seconds	0.07 seconds
Correctly Classified Instances	68.97%	69.45%	68.84%
Precision	0.655	0.664	0.652
Recall	0.69	0.694	0.688
F-Measure	0.631	0.641	0.627
ROC Area	0.594	0.603	0.614

4.2 Performance results of models after algorithm optimization

The three models were comparable as regards to accuracy about 69%. REP Tree was much faster as the time taken to build model was only 0.07 seconds compared to 0.25 seconds needed to build C4.5 tree. The CART tree was very slow compared to REP tree or C4.5, the time taken to build the CART model was 11.19 seconds. We concluded that CART decision tree is too slow to be included in the comparison study (see Table 3).

Table 3. Comparison between C4.5, REP Tree after algorithm optimization

	C4.5	REP tree
Time taken to build model	0.25 seconds	0.07 seconds
Optimum Correctly Classified Instances	70.22%	70.7%
Optimum Precision	0.685	0.689
Optimum Recall	0.702	0.707
Optimum F-Measure	0.65	0.659
Optimum ROC Area	0.636	0.64

Out of 19 attributes the decision tree models showed that AFP was selected as the variable of initial split (most decisive); and that patients with low AFP levels have higher response rates (see Figure 1,2). AFP is usually used for the screening or the diagnosis of hepatocellular carcinoma, but the current study has highlighted that low AFP level is a decisive predictor of response , and was able to confirm previous reports that Peg-IFN/RBV therapy was associated with a decrease in serial AFP levels irrespective to the virological response to treatment [10,11,12]. Other attributes as age, male gender, BMI, ALT, total bilirubin, albumin, and blood glucose have less decisive role for prediction of response. These finding were confirmed using multivariate analysis.

Using this model, an estimate of the response before treatment can be rapidly obtained, which may facilitate clinical decision making. The high probability group may be suitable candidates for PEG-IFN/RBV therapy, however, the estimation of low probability should not be used to preclude patients from therapy, and the final decision should be made on a case-by-case basis.

For a tight budget, a decision maker may choose to select a predictive model with a high precision over other models. This strategy will help minimize the overall treatment cost. The drawback of this strategy is that some patients who are classified wrongly as not respondent to the antiviral drugs will not get the treatment and therefore they may suffer severe illness.

For a high budget, a decision maker may choose to select a predictive model with a high recall over other models. This strategy will help treatment a great number of patients. The drawback of this strategy is that some patients who are classified wrongly as respondent to the antiviral drugs will get the treatment and therefore they will not get cured beside the loss of resources

For a moderate budget, a decision maker may choose to select a predictive model with a suitable F-measure (which is a combination of recall and precision)

5. CONCLUSION

We built a pre-treatment model for the prediction of virological response to PEG-IFN/RBV. We have developed simple software that generates different possible combination of parameters to find the best classifier. REP tree has fast and well performance for classification on medical data sets of increased size. Because this decision tree model was made up of simple variables, it can be easily applied to clinical practice. This model may have the potential to support decisions about patient selection for PEG-IFN/RBV based on a possibility of response weighed against the potential risk of adverse events or costs. Various pre-treatment decision tree algorithms have demonstrated that low level of AFP is associated with high response rate. To our knowledge this study has highlighted that low AFP level is a decisive predictor of response regardless to other factors.

6. ACKNOWLEDGMENTS

This study was supported by a grant from the Science and Technology Development Fund (STDF) through supporting the project entitled: Bioinformatics in predicting the response to interferon-ribavirin combination therapy in patients with HCV genotype-4 (Bio-IN-therapy).

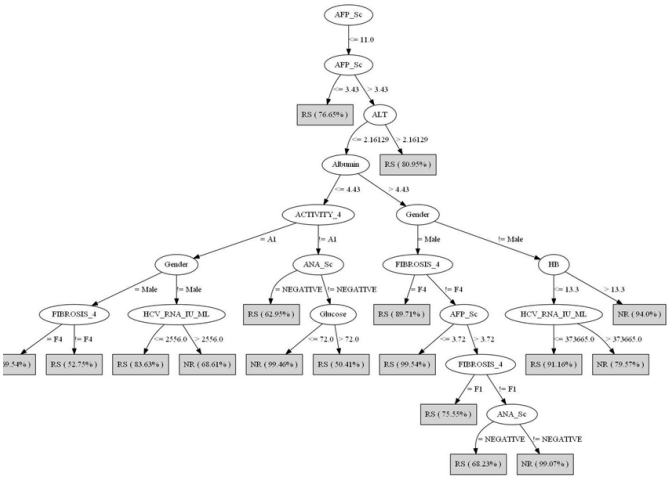

Figure 1(a). C4.5 algorithm tuning to perform best ROC Area - Part (a)

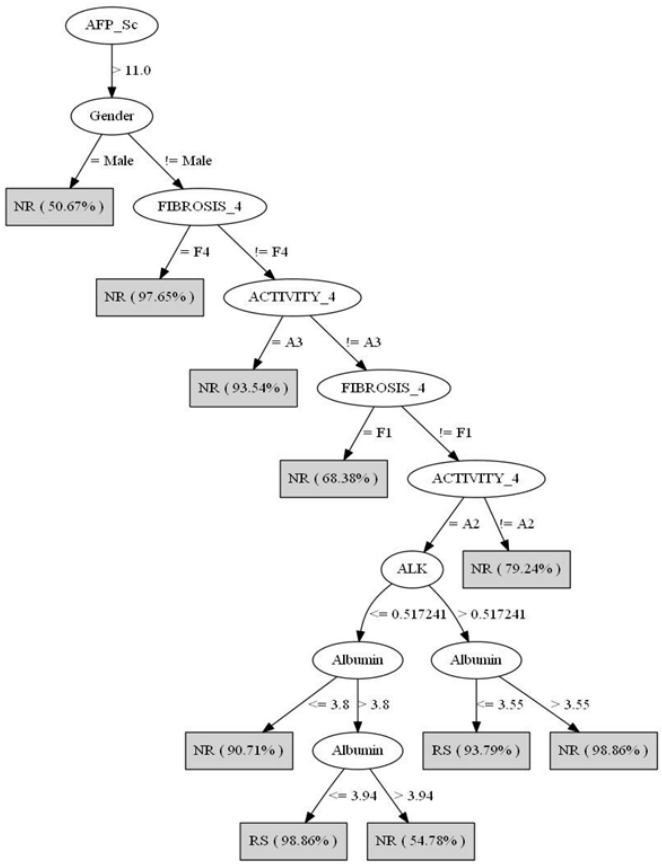

Figure 1(b). C4.5 algorithm tuning to perform best ROC Area - Part (b)

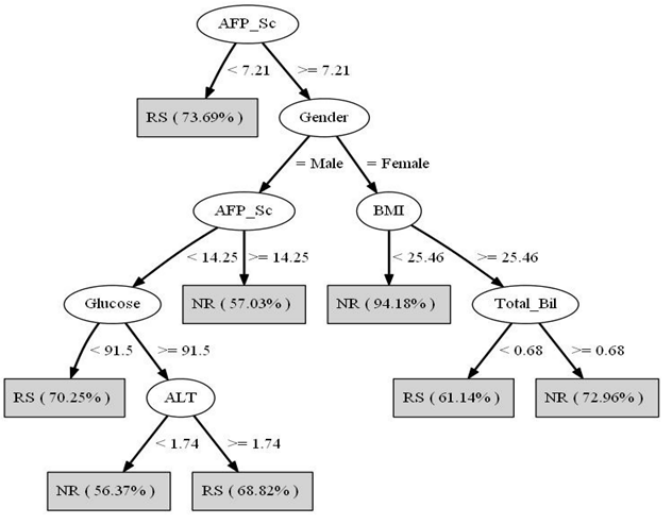

Figure 2. REPTree algorithm tuning to perform best ROC Area

7. REFERENCES

[1] El-Zanaty, Fatma and Ann Way. 2009. Egypt Demographic and Health Survey 2008. Cairo, Egypt: Ministry of Health, El-Zanaty and Associates, and Macro International.

[2] Dienstag JL., Mchutchison JG (2006) *American gastroenterological association medical position statement on the management of hepatitis C: Gastroenterology, 130 (1)*, 225-226.

[3] Khattab M., Ferenci P., Stephanos J. Hadziyannis, Colombo M., Manns P. Almasio L., Rafael Esteban, Ayman A. Abdo, Stephen A. Harrison9, Nazir Ibrahim, Cacoub P., Eslam M., Samuel S. Lee. Management of hepatitis C virus genotype 4: Recommendations of An International Expert Panel. *Journal of Hepatology*, *54(6)*. 1250-1262

[4] Bellazzi, R., and Zupan, B. Predictive Data Mining in Clinical Medicine: Current Issues and Guidelines. *International Journal of Medical Informatics, 77(2)*, 81 – 97.

[5] Kurosaki M, Sakamoto N, Iwasaki M, et al. Pretreatment prediction of response to peginterferon plus ribavirintherapy in genotype 1 chronic hepatitis C using data mining analysis. *J Gastroenterol 2011, 46(3)*, 401-409.

[6] Khairy M, Fouad R, Mabrouk M, El-Akel W, Awad AB, Salama R, Elnegouly M, Shaker O. The impact of interleukin 28b gene polymorphism on the virological response to combined pegylated interferon and ribavirin therapy in chronic HCV genotype 4 infected egyptian patients using data mining analysis. *Hepatitis Monthly, 13(7),* E10509. DOI= http://dx.doi.org/10.5812/hepatmon.10509.

[7] Ross Quinlan J. *C4.5:Programs for Machine Learning*. Morgan Kaufmann, San Mateo - CA, 1993.

[8] Breiman LJH, Friedman RA, Olshen CJ, Stone CM. *Classification and regression trees*. Wadsworth Publishing, California, 1980.

[9] Agrawal, Rakesh and Srikant, Ramakrishnan and others. Fast algorithms for mining association rules. *Proc. 20th int. conf. very large data bases, VLDB, 1215*, 487-499.

[10] Murashima S, Tanaka M, Haramaki M, Yutani S, Nakashima Y, Harada K, Ide T, Kumashiro R, Sata M. A decrease in AFP level related to administration of interferon in patients with chronic hepatitis C and a high level of AFP. *Dig Dis Sci (2006), 51*, 808-812.

[11] Tamura Y, Yamagiwa S, Aoki Y, Kurita S, Suda T, Ohkoshi S, Nomoto M, Aoyagi Y. Serum alpha-fetoprotein levels during and after interferon therapy and the development of hepatocellular carcinoma in patients with chronic hepatitis C. *Dig Dis Sci (2009), 54*, 2530-2537.

[12] Mahasen Mabrouk, Wahid Doss, Naglaa Zayed, Shimaa Afify, Gamal Esmat. Impact of Serum Alpha-fetoprotein Levels on the Response to Antiviral Therapy in Egyptian Patients with Chronic Hepatitis C. *J hepatogastroenterology (2013), 2(5).*