The Irrelevant Values Problem of Decision Tree in Medical Examination

Kuang-Yi Chou[3], Huan-Chao Keh[1], Nan-Ching Huang[1]*, Shing-Hwa Lu[2],
Ding-An Chiang[1] and Yuan-Cheng Cheng[1]

[1]*Department of CSIE, Tamkang University,
Tamsui, Taiwan 251, R.O.C.*
[2]*Department of Urology, School of Medicine, National Yang-Ming University and Department of Urology,
Zhong Xiao Branch, Taipei City Hospital, Urological Center, Taipei, Taiwan 115, R.O.C.*
[3]*Center of General Education, National Taipei University of Nursing and Health Sciences,
Taipei, Taiwan 112, R.O.C.*

Abstract

Data mining technique is extensively used in medical application. One of key tools is the decision tree. When a decision tree is represented by a collection of rules, the antecedents of individual rules may contain irrelevant values problem. When we use this complete set of rules to medical examinations, the irrelevant values problem may cause unnecessary economic burden both to the patient and the society. We used a hypothyroid disease as an example for the study of irrelevant values problem of decision tree in medical examination. Hypothyroid disease is used to associate to the mechanism of thyroid-stimulating hormone (TSH). Physicians will combine lots of information; such as patient's clinical records, medical images, and symptoms, prior to the final diagnosis and treatment, especially surgical operation. Therefore, to avoid generating rules with irrelevant values problem, we propose a new algorithm to remove irrelevant values problem of rules in the process of converting the decision tree to rules utilizing information already present in the decision tree. Our algorithm is able to handle both discrete and continuous values.

Key Words: Decision Tree, Classification, Irrelevant Values, Missing Branches, Medical Examination

1. Introduction

The decision tree is based on the application of recursive formula and the algorithm of top-down and divide-and-conquer. The irrelevant values problem may be generated during the construction of decision tree. The structure of the decision tree may be altered at the cut-off point while irrelevant values are removed. On the other hand, the decision tree can handles continuous values through a local discretization which the dependency among all attributes must relate to each other; if the dependency does not exist, a global discretization is a better choice.

*Corresponding author. E-mail: 132310@mail.tku.edu.tw

The ID3 algorithm is commonly brought to solve the irrelevant values problem of a decision tree. When a decision tree does not have abundance of data, the grafting technique is mostly applied to its algorithm. As a result, a best test result can be obtained from a cut-off point where its irrelevant values problem has been removed. In this paper, it proposes an algorithm which removes irrelevant values problem and uses association rules to integrate both classification rules and a cut-off point with global discrete attributes. At same time, it creates a classifier to integrate all classification rules. A decision tree, moreover, is built up by selecting the best test attribute as the root of the decision tree. Then, the same procedure is operated on each branch to lead on the remaining levels of the decision tree until all examples in a leaf belong to

the same class. The decision tree; however, creates a branch for each value which appears in the training data without considering whether the value is relevant to the classification. In addition the resultant tree may have the over-specialization problems [1–3]. Without losing generality, we only consider ID3-like algorithm in our study.

Cheng et al. [4] and Bohanec and Rajkovic [5] pointed out that the irrelevant values problem and the missing branches problem are two major causes of over-specialization in the decision tree. The missing branches problem is due to the reduced subsets at the non-leaf nodes which do not necessarily contain examples of every possible value of the branching attribute. Consequently, the decision tree may fail to classify some instances. Since some values of that attribute may be irrelevant to the classification; the resultant rules of the decision tree may have irrelevant conditions, which extra information must be supplied [5,6]. Extra information means more examinations and costs are required to the patient and the society. When the decision tree is applied to medical applications we have to deal with the irrelevant conditions in the decision tree to save medical resources and avoid unnecessary examinations.

For example, let us consider the decision tree in Figure 1; we are going to focus on the irrelevant values in the branch Br. As shown in Figure 1, branches Br_1 and Br_2 can be represented by $a_1 \wedge b_1 \rightarrow c_1$ and $a_2 \wedge b_1 \rightarrow c_1$; however, values a_1 and a_2 are irrelevant to each other respectively. Consequently, these rules can be generalized by deleting these irrelevant values without affecting its accuracy; leaving the more appealing rule, $b_1 \rightarrow c_1$. The resultant rule is more concise and comprehensible than the original rule. However, its drawback is likely to suffer from missing value problem. This kind of rules may be useful in many applications. For example, the doctor can examine item B first to make sure whether or not this patient needs to take item A. This process can reduce some burdens; for example, expense, inconvenience or harm to the patient.

Fayyad proposed two algorithms, GID3 and GID3*, to solve the over-specialization problem of the decision tree constructed using ID3 [4,7–9]. Both algorithms over-whelm the irrelevant values problem at attribute phantomization step before attribute selection. However, the problems found in branches of GID3 and of GID3* is

longer than the branches found in ID3.

Quinlan pointed out that the tree is represented by a collection of rules and each leaf in the tree would not result in much simpler than the tree which one rule is for every leaf [6,10]. Therefore, Quinlan used a pessimistic estimation on the rule accuracy to generalize each rule by removing not only irrelevant conditions but also conditions that hardly discriminated nominated class from other classes [6,11]. Since the antecedents of a rule may contain irrelevant conditions, the deletion of superfluous conditions is result in a generalized rule regardless of its correctness and is simpler than the tree. Jerez-Aragones et al. [1] and Chiang et al. [12] provided another view for the solution of over-specialization problem from a decision tress which its attributes contain discrete values only. They eliminated irrelevant values in the process of converting the decision tree to the rules for classification according to the information on the decision tree. Our algorithm does not use example cases to guide the conversion process. It has clear computational advantages which can be integrated into any existing decision tree system easily. The empirical results present evidences that our algorithm overcomes not only the irrelevant values problem, but also the missing branch problems with negligible cost of concerning the construction of decision tree [13–15]. Consequently, the average error rate of the decision tree is also reduced. The new algorithm can not only deal with discrete values but also with continuous ones.

The decision tree is one of the key data mining techniques in medical application. As a matter of fact, decision tree of nowadays does not have a specific system for certain types of data, because the algorithms generated from a decision tree are varied and depends on type of
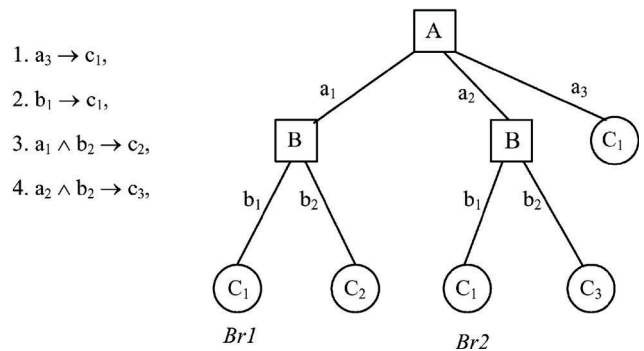
1. $a_3 \rightarrow c_1$,

2. $b_1 \rightarrow c_1$,

3. $a_1 \wedge b_2 \rightarrow c_2$,

4. $a_2 \wedge b_2 \rightarrow c_3$,

Figure 1. A decision tree with irrelevant values.

input data. Because of digitization of medial information, any diagnosis has to correlate with various medial records from different database and the process may cause unnecessary economic burden both to the patient and the society. We used hypothyroid disease which is related to the mechanism of thyroid-stimulating hormone (TSH) as a clinical example for the study of irrelevant values problem in the decision tree. The attributes of many medical data are continuous values; therefore, the new algorithm may be used to solve irrelevant values problem of a decision tree

An algorithm for identifying irrelevant values problem of the decision tree with discrete values is introduced in section 2. The new algorithm for removing irrelevant values of the decision tree with continuous values is given in section 3. Since the discrete values can be viewed as a special case of the continuous values; as a result, the new algorithm can deal with both discrete and continuous values. The irrelevant values problem for medical examination is discussed in section 4. The conclusions are stated in section 5.

2. An Algorithm to Identify Irrelevant Values to Discrete Values

In this section, we introduced some definitions and theorems from our previous work with respect to identify irrelevant values of the decision tree with discrete values [12]. Let $A = \{A_1, ..., A_n\}$ be a set of attributes, $C = \{C_1, ..., C_s\}$ be a set of classes, and $\mathbf{a} = (a_{ij}, ..., a_{nm})$ be a branch's values of Br. To represent a decision tree by a set of branches, the branch Br of the decision tree can be represented as the form, $Br[A_1] \wedge ... \wedge Br[A_n] \to C_k$ or $a_{1r} \wedge ... \wedge a_{nm} \to C_k$, where $Br[A_i]$ is the branch value out of an attribute A_i in the branch Br, $a_{ij} \in Br[A_i]$, $i = 1 ... n$ and $1 \le k \le s$.

For the decision tree, not all attributes A_i will be the nodes of branches Br in the decision tree. When the corresponding values of $Br[A_i]$ are missing, these attributes are irrelevant with respect to Br. If there are many rules implied by Br, the irrelevant attributes will be considered. This observation can be explained by the following definitions.

Definition 1.

Let $Br[A_1] \wedge ... \wedge Br[A_{j-1}] \to C_k$ be a branch in the de-

cision tree, then the rules with respect to attributes $A_1, ..., A_n$ implied by Br are:

$\{Br[A_1] \wedge ... \wedge Br[A_{j-1}] \wedge a_j \wedge ... \wedge a_{ns} \to C_k \mid a_{jr}, ..., a_{ns} \in$ domain $(A_j, ..., A_n)\}$,

where domain $(A_j, ..., A_n) = $ domain $(A_j) \times ... \times$ domain (A_n).

To easily identify irrelevant values of a branch in a decision tree, we further define the following definition.

Definition 2.

Let Br and Br' be two different branches in a decision tree, where $Br = Br[A_1] \wedge ... \wedge Br[A_j] \to C_{k1}$, then Br is in conflict with Br' with respect to attributes $A_1 ... A_j$ if and only if $Br[A_1] \wedge ... \wedge Br[A_j] \to C_{k2}$ is a part of rule implied by Br' and $C_{k1} \neq C_{k2}$.

According to the semantics of irrelevant values; a value, $Br[A_j]$, is an irrelevance of a rule. This value can be deleted or replaced by any value from the same domain value without affecting the correctness of the rule. Therefore, based on definitions 1 and 2, a combinatorial explosion can be applied to the number of comparisons to all the branches for identification of irrelevant values of a branch. The process of identifying irrelevant values by definition 1 and 2 is very time-consuming. To enable users to focus on only relevant conditions of the rules, the following theorems are able to work out the irrelevant values problem for a complex decision tree. These theorems were proven in [13,14,16].

Theorem 1.

Let $Br[A_1] \wedge ... \wedge Br[A_{j-1}] \wedge Br[A_j] \wedge Br[A_{j+1}] \wedge ... \wedge Br[A_n] \to C_k$ be a branch through a non-leaf node P in a decision tree, and the branching attribute with respect to P be A_j. For all branches through P of the decision tree, if Br is not in conflict with these branches with respect to attributes $A_{j+1} ... A_n$, then $Br[Aj]$ is an irrelevant value in Br.

Theorem 2.

Let Br be a branch through a non-leaf node P of the decision tree. When the branch value $Br[P]$ has been identified by theorem 1. Other branches through P are useless for the following process to identify the irrelevant values of Br.

Theorem 3.

Let $Br[A_1] \wedge ... \wedge Br[A_j] \wedge Br[A_{j'}] \wedge ... \wedge Br[A_{n1}] \rightarrow C_{k1}$ and $Br'[A_1] \wedge ... \wedge Br'[A_j] \wedge ... \wedge Br'[A_{n2}] \rightarrow C_{k2}$ be two branches through a non-leaf node P in the tree, where the branching attribute with respect to P is A_j. Let $A = \{A_{j'}, ..., A_{n1}\}$ and A_1 be the same attributes in these two branches, where $A_1 \subseteq A$ then, Br is in conflict with Br' with respect to A if and only if $Br[A_1] = Br'[A_1]$ and $C_{k1} \neq C_{k2}$.

According to theorem 1, for each selected node P, the branch value of node P can identify whether the branch value of node P is an irrelevant value of a branch, Br, or not. By theorem 2, if the branch value of node P is an irrelevant value, then branches, which are through node P, can be ignored for the following process to identify irrelevant values of Br. In other words, by theorem 1 and theorem 2, to identify all the irrelevant values of a branch, we need to check all the branches in the decision tree only once. Moreover, since we do not have to consider the rules implied by each branch in the decision tree by theorem 3, the computation time of identifying whether two branches are conflict between each other can be reduced greatly. Actually, without losing generality, since the number of common nodes of two branches is always small, we can assume that the time complexity of identifying whether two branches are in conflict with each other is constant. Therefore, the time complexity of identifying all irrelevant values of a branch by these theorems is reduced to $O(m)$ at worst case, where m is the number of branches of the tree.

3. An Algorithm to Identify Irrelevant Values to Continuous Values

For many applications, some attributes may contain continuous values. Therefore, we provided a new algorithm, which is represented by the following theorem, to solve the irrelevant values problem of the decision tree with continuous values.

Theorem 4.

Let $Br[A_1] \wedge ... \wedge Br[A_j] \wedge Br[A_{j'}] \wedge ... \wedge Br[A_{n1}] \rightarrow C_{k1}$ and $Br'[A_1] \wedge ... \wedge Br'[A_j] \wedge ... \wedge Br'[A_{n2}] \rightarrow C_{k2}$ be two branches through a non-leaf node P in the tree, where the branching attribute with respect to P is A_j. Let

$A = \{A_{j'}, ..., A_{n1}\}$, A_1 be the same attributes in these two branches and a_1 be a branch's values of $Br[A_1]$, where $A_1 \subseteq A$. Then, Br is in conflict with Br' with respect to A if and only if

when $A_1 \neq \varnothing$, $\exists a_1$, $a_1 \in Br[A_1]$, $a_1 \in Br'[A_1]$ and $C_{k1} \neq C_{k2}$ (1) or

when $A_1 = \varnothing$, $C_{k1} \neq C_{k2}$ (2).

Proof.

Let $A = \{A_{j'}, ..., A_{n1}\}$ and A_1 be the same attributes in these two branches.

Let $A_1 \neq \varnothing$, a_1 be the branch's values, and $a_1 \in Br[A_1]$ (1). When $\forall a_1$, $a_1 \notin Br'[A_1]$, it implies that these two branches will never be in conflict with each other with respect to A_1 by definition 2. Therefore, we need only to consider the case $\exists a_1$, $a_1 \in Br[A_1]$, $a_1 \in Br'[A_1]$ and $C_{k1} \neq C_{k2}$. According to definition 1, when $a_1 \in Br'[A_1]$, $a_1 \rightarrow C_{k2}$ must be a part of rule implied by Br'. Therefore, Br must be in conflict with Br' with respect to A if and only if $\exists a_1$, $a_1 \in Br[A_1]$, $a_1 \in Br'[A_1]$ and $C_{k1} \neq C_{k2}$ by definition 2.

Let $A_1 = \varnothing$ and $C_{k1} \neq C_{k2}$ (2). According to definition 1, when $A_1 = \varnothing$, it implies that $\forall a$, $a \rightarrow C_{k2}$ must be a part of rule which is implied by Br', where $a \in Br[A]$. Since $C_{k1} \neq C_{k2}$, Br must be in conflict with Br' with respect to A by definition 2.

According to theorem 4, when $C_{k1} = C_{k2}$; branches $Br \rightarrow C_{k1}$ and $Br' \rightarrow C_{k2}$, are never in conflict with each other. To identify all irrelevant values of a branch Br, we need only to consider those branches; Br', whose leaves are different from C_{k1}. When attribute has only discrete and finite values, it can check easily whether it has branch values in Br and Br' at the same time or not based on theorem 3. On the other hand, when some attributes contain continuous values, the situation becomes more complicated. When the branching attribute, $Br[A]$, contains continuous values, the branch value of A is computed from the training data and it may or may not appear in the training data. After the branch's value of A is determined, the values of A in Br is ranged by one of $\leq$, $>$, $<$ or $\geq$ function. Consequently, we have to check the training data; otherwise, it is very hard to known whether or not Br is in conflict with Br' by theorem 4.

For example, let $Br = A > 5.8 \rightarrow c_1$, $Br' = A \leq 6 \rightarrow c_2$, and $c_1 \neq c_2$. According to theorem 4 and without consideration of the training data, we will conclude that Br

is in conflict with Br' because $Br[A] \cup Br'[A] \neq \varnothing$. However, since the braches values of continuous values is computed from training data, the value 5.8 may not be in the range of $Br[A]$. In other words, if the smallest value in the range of $Br[A]$ is larger than 6, the new conclusion is that Br is not in conflict with Br'. Therefore, we have to check the training data in the process of removing all irrelevant conditions. Otherwise, some irrelevant values can not be removed. Even that, our algorithm is still useful to integrate into any existing decision tree system without training data of example 3. In addition, we can not modify the existing systems and can not use example cases to guide the conversion process. To solve this problem, we may scan the database one time for the continuous value attributes to find out the largest and smallest values from the corresponding training data. Consequently, all irrelevant values were removed from the resultant rules without usage of training data. Moreover, to make the process of removing all irrelevant values from a branch more efficient, we recursively applied the above theorems until the node P was at the root of decision tree. The corresponding algorithm is shown below.

Input: A decision tree

Output: A set of rules without irrelevant conditions;

Let $\boldsymbol{Br} = \{Br_1, ..., Br_m\}$; /* the branches of the decision tree */

For each branch Br' in Br **Do**

$\{$Let $Br' = Br'[A_1] \wedge ... \wedge Br'[A_k] \rightarrow C_i$

For $j = k$ down to 1 Do

$\{$Apply theorem 1, 2 and 4, and check whether $Br'[A_j]$ is an irrelevant value;

If $Br'[A_j]$ is an irrelevant value **Then** remove $Br'[A_j]$ from Br';$\}$

Represented Br' by a rule;$\}$

Let us consider the decision tree depicted in Figure 2. This decision tree was proposed by Quinlan for hypothyroid disease examinations. Without consideration of training data in the database, the original rules and the resultant rules are shown in Table 1. Comparing these two sets of rules, we found out that our rules are more concise and comprehensible than the original rules.

4. Our Algorithm for Medical Examinations

There are many medical examinations available for

our decision tree study. One of most interesting cases is hypothyroid disease which is related to the mechanism of thyroid-stimulating hormone (TSH). TSH is secreted by thyrotrope cells in the pituitary gland, which regulates the endocrine function of the thyroid gland. TSH stimulates the thyroid gland to secrete the triiodothyronine (T3) and thyroxine (T4). The level of T3 and T4 in the blood has an effect on the pituitary release of TSH; when the levels of T3 and T4 are low, the production of TSH is increased, and on the converse, when levels of T3 and T4 are high, TSH production is decreased. This effect creates a regulatory negative feedback loop. Additionally, the level of T3 and T4 in the blood is the index for diagnosis of thyroid-related disease.

As a result, we use the decision tree depicted in Figure 2 to discuss the irrelevant values problem regarding to medical examinations of hypothyroid disease. The rules found in Table 1 have two conditions; one is irrelevant and the other is no irrelevant. When two conditions are compared to each other, some characters can be found.

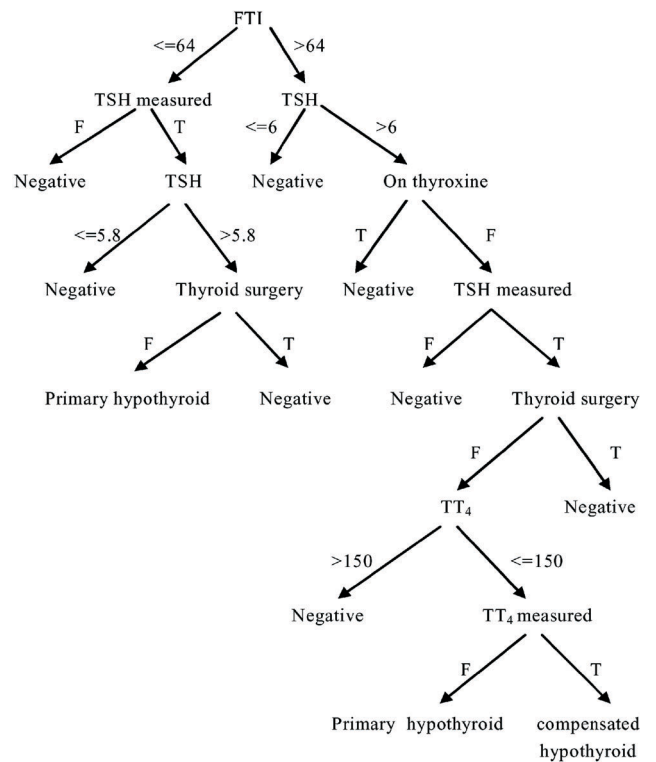First, the decision tree model recommends the doctors involve many unnecessary investigations for the

Figure 2. The ID3 decision tree for the diagnosis of hypothyroid disease (unit in mIU/L).

Table 1. The recurred rules of the decision tree (unit in mIU/L)

The original rules without considering irrelevant values problem	The resultant rules without irrelevant conditions
FTI ≤ 64* ∧ TSH measured = f → Negative	(1) TSH measured = f → Negative
FTI ≤ 64* ∧ TSH measured = t* ∧ TSH ≤ 5.8 → Negative	(2) TSH ≤ 5.8 → Negative
FTI ≤ 64 ∧ TSH measured = t ∧ TSH > 5.8 ∧ Thyroid surgery = f → Primary hypothyroid	(3) FTI ≤ 64 ∧ THS measured = t ∧ TSH > 5.8 ∧ Thyroid surgery = f → Primary hypothyroid
FTI ≤ 64 ∧ TSH measured = t ∧ TSH > 5.8* ∧ Thyroid surgery = t → Negative	(4) Thyroid surgery = t → Negative
FTI > 64 ∧ TSH ≤ 6 → Negative	(5) FTI > 64 ∧ TSH ≤ 6 → Negative
FTI > 64 ∧ TSH > 6* ∧ On thyroxin = t → Negative	(6) FTI > 64 ∧ On thyroxin = t → Negative
FTI > 64 ∧ TSH > 6 ∧ On thyroxin = f* ∧ TSH measured = f → Negative	The resultant rule is equal to (1)
FTI > 64 ∧ TSH > 6 ∧ On thyroxin = f ∧ TSH measured = t ∧ Thyroid surgery → Negative	The resultant rule is equal to (4)
FTI > 64 ∧ TSH > 6 ∧ TSH measured = t ∧ Thyroid surgery = f* ∧ TT4 > 150 → Negative	(7) FTI > 64 ∧ TT4 > 150 → Negative
FTI > 64 ∧ TSH > 6 ∧ TSH measured = t ∧ thyroid surgery = f ∧ TT4 ≤ 150 ∧ TT$_4$ measured = f → Primary hypothyroid	(8) TSH > 6 ∧ THS measured = t ∧ Thyroid surgery = f ∧ TT4 ≤ 150 ∧ TT$_4$ measured = f → Primary hypothyroid
FTI > 64 ∧ TSH > 6 ∧ THS measured = t ∧ Thyroid surgery = f ∧ TT4 ≤ 150 ∧ TT$_4$ measured = t → Compensated hypothyroid	(9) FTI > 64 ∧ TSH > 6 ∧ THS measured = t ∧ Thyroid surgery = f ∧ TT4 ≤ 150 ∧ TT$_4$ measured = t → Compensated hypothyroid

FTI ≤ 64*: Free thyroxin (FT4) can be measured or calculated directly as the free thyroxin index (FTI). The FTI is a level of T4 in relation to the amount of thyroxin-binding globulin present. The FTI is calculated from the T4 and T3 uptake, or T3U. The FTI value can indicate when an abnormal level of thyroxin-binding globulin in the blood causes an abnormal level of T4. Also the FTI can help tell if abnormal amounts of T4 are present because of abnormal amounts of thyroxin-binding globulin. **TSH > 5.8* and TSH > 6:** The interaction between feedback mechanism and thyroid relies on TSH value. Normal values range from 0.4 to 5.8 mIU/L for people with no symptoms of an under-active thyroid. Values below 0.4 are considered hyperthyroid; values 5.8 mIU/L or slightly higher are considered sub clinical hypothyroid, and values generally above 10 mIU/L indicate full hypothyroid condition. **TSH measured = t*:** TSH measured is a qualitative measurement of TSH. True indicates an abnormal plasma TSH level and False refers to a normal level of TSH. **Thyroid surgery = f*:** This means a patient without thyroid surgery. Thyroid surgery could result in the low level of thyroxin which mimics a manifestation of hypothyroid. **On thyroxin = f*:** Patient who is under medication of thyroxin will change the level of serum thyroxin. False result indicates the patients are not under thyroxin medication. The Table 2 is a comparison of different types of investigation diagnostic methods.

patients. For example, the rules of 1, 2, and 4 show that they are much simpler for physicians to make diagnosis from the investigation results. In our algorithm, a history with negative TSH, thyroid surgery, or TSH less than 5.8 leads the diagnosis to a negative result. This means that if a patient has a history of thyroid surgery or thyroxin medication, his diagnosis will always be negative and doe not need to have the other test like TT4 or FT4. For a patient with a measurement of TSH below 5.8, the afore-mentioned tests are not required. Only when TSH is above 5.8; patient will need the examination of FTI or TT4 to make a correct diagnosis.

Second, the proposed algorithm shows a better way for a physician to make a correct diagnosis without prob-

Table 2. Features of the types of investigation in diagnosis of hypothyroid disease

Types of investigation	Cost	Convenience	Harmful
TSH	Medium	Yes	No
TSH measured	Low	Yes	No
Thyroid surgery	Low	Yes	No
TT$_4$	Medium	Yes	No
On thyroxin	Low	Yes	No
FTI	High	Yes	No

ing every available investigation. Only the patients without a thyroid surgery history and thyroxin medication need to have their FTI, TT4 and TSH examined to make

the diagnosis. On the contrary the decision tree model shows that all tests are needed. Therefore, we can optimize lab investigations for diagnosing a possible hypothyroid patient with a new algorithm than the original decision model. Less investigation is both good for patients and the society. For each patient, he/she might not spend time and money to the medical examination which is not needed for diagnosis. For our society, it is a waste if we spend our fiscal expenditure for any medical procedure that is not quite needed. The doctors' time is limited, the patients are suffering from the disease both physically and mentally, the ever-increasing cost of medical expenditure is a huge burden to the governments. So it is a great benefit for all parts if we can find alternatives to cut down the medical expenditure, such as the lab investigations for the diagnosis.

Third, the decision tree model shows that different values in TSH, FTI, T4, and TT4 could lead to different diagnosis which depends on the value of other investigations. When combination of numerous investigations is involved, physicians could make mistakes easily with a tight schedule. The new algorithm reduces the irrelevant rules and makes the diagnosis much simpler and; therefore, it is less likely to make mistakes in clinical work. Therefore, the new algorithm model cuts down the rules and helps clinicians to find the easiest and fastest way to apply suitable investigations for a particular patient correctly.

5. Conclusion

The solution of removing irrelevant values problem depends only on the semantics of the decision. Consequently, our new algorithm can overcome the irrelevant values problem by integration of existing tree-construction algorithm without increase of computational cost for the construction of decision tree. Medical problem is a field where "decision tree" algorithm is usually applied. As the cost of healthcare system is rising again and again, it is urgent for government to reduce the fast increase of the medical expenditure while maintaining the quality. Using the decision tree model could be helpful for us to retrieve the relevant medical tests needed in clinical cases under different condition. However, the new rules we adopted here and testified in a hypothyroid model showed more efficient and more useful to exclude

the irrelevant conditions. Therefore, the new rules could be a good technique for physicians to find the easiest and fastest way to make a diagnosis and cut down the indispensable expenditures of lab investigations. Moreover, as discussed in section 3, the rules may be further reduced by considering the training data. We pause to collect real data to do the cases. Therefore, we plan to combine these topics into our algorithm in the near future.

References

[1] Jerez-Aragonés, J. M., et al., "A Combined Neural Network and Decision Trees Model for Prognosis of Breast Cancer Relapse," *Artificial Intelligence in Medicine*, Vol. 27, pp. 45–63 (2003).

[2] Kononenko, I., "Machine Learning for Medical Diagnosis: History, State of the Art and Perspective," *Artificial Intelligence in Medicine*, Vol. 23, pp. 89–109 (2001).

[3] Neumann, A., et al., "Machine Learning for Medical Diagnosis: History, State of the Art and Perspective," *Artificial Intelligence in Medicine*, Vol. 32, pp. 97–113 (2004).

[4] Cheng, J., et al., "Improved Decision Trees: A Generalized Version of ID3," in *Proceedings of the Fifth International Conference on Machine Learning*, Ann Arbor, Michigan, USA, pp. 100–108 (1988).

[5] Bohanec, M. and Rajkovic, V., "DEX: An Expert System Shell for Decision Support," *Sistemica*, Vol. 1, pp. 145–157 (1990).

[6] Delen, D., et al., "Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods," *Artificial Intelligence in Medicine*, Vol. 34, pp. 113–127 (2005).

[7] Fayyad, U. M., "Branching on Attribute Values in Decision Tree Generalization," in *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, Seattle, Washington, pp. 601–606 (1994).

[8] Fayyad, U. M. and Irani, K. B., "A Machine Learning Algorithm (GID3*) for Automated Knowledge Acquisition: Improvements and Extensions," MI: GM Research Labs, Warren (1991).

[9] John, G. H., et al., "Irrelevant Features and the Subset Selection Problem," in *Machine Learning: Proceedings of the Eleventh International Conference*, W. W. Cohen and H. Hirsh, Eds., ed San Fransisco, CA: Mor-

gan Kaufmann Publisher, pp. 121−129 (1994).

[10] Maher, P. E. and St. Clair, D., "Uncertain Reasoning in an ID3 Machine Learning Framework," in *Proceedings of the Second IEEE International Conference on Fuzzy Systems*, pp. 7−12 (1993).

[11] Breiman, L., et al., *Classification and regression trees*, Belmont, California: Wadsworth (1984).

[12] Chiang, D.-A., et al., "The Irrelevant Values Problem in the ID3 Tree," *Computers and Artificial Intelligence*, Vol. 19, pp. 169−182 (2000).

[13] Cai, Y., et al., "An Attribute-Oriented Approach for Learning Classification Rules from Relational Databases," in *Proceedings of the Sixth International Conference on Data Engineering*, Los Angeles, California, USA, pp. 281−288 (1990).

[14] Chi, Z. and Yan, H., "ID3-Derived Fuzzy Rules and Optimized Defuzzification for Handwritten Numeral Recognition," *Fuzzy Systems, IEEE Transactions on*, Vol. 4, pp. 24−31 (1996).

[15] Kamber, M., et al., "Generalization and Decision Tree Induction: Efficient Classification in Data Mining," in *Proceedings of the Seventh International Workshop on Research Issues in Data Engineering*, pp. 111−120 (1997).

[16] Ou, M. H., et al., "Dynamic Knowledge Validation and Verification for CBR Teledermatology System," *Artificial Intelligence in Medicine*, Vol. 39, pp. 79−96 (2007).