

# Fever Classification and Remedial Recommendation System

San Hay Mar Shwe, Htet Ne Oo

**Abstract**—Healthiness is the most important part in the person's life. So, knowing of details about disease is essential for everyone. Nowadays, different fevers can be commonly found and every person including doctors wants to decide the disease correctly. So, in this paper fever classification and remedial recommendation system by using decision tree algorithm is presented and also basic characteristics of decision trees are mentioned. In medical decision making (classification, diagnosing, etc.) there are many situations where decision must be made effectively and reliably. Conceptual simple decision making models with the possibility of automatic learning are the most appropriate for performing such tasks. Decision trees are a reliable and effective decision making technique that provide high classification accuracy with a simple representation of gathered knowledge and they have been used in different areas of medical decision making.

**Keywords**— classification, decision tree, fever, ID3.

## I. INTRODUCTION

Nowadays, human being lives in IT age and more require latest technology. So, information technology is being used in almost every sector of life such as health care, transportation, economic, education and so on. Healthiness is the most important part in the person's life. So, knowing of disease is essential for everyone. Today, many countries around the world are changing the way of implementing health care for the patient through electronic health care system. In Myanmar, computerized health care systems appeared and they help the patients and the doctor to determine the disease with their symptoms.

Artificial Intelligence (AI) is the study of how to make computers do things at which, at the moment, people are better. AI has the definite goal of understanding intelligence and building intelligent systems. However, the methods and formalisms used on the way to this goal are not firmly set, which has resulted in AI course lies in conveying as many branches as possible without losing too much depth and precision [9].

Data classification is a two-step process. In the first step, a model is built describing a predetermined set of data classes or concepts. The model is constructed by analysing database

tuples described by attributes. Each tuple is assumed to belong to a predefined class, as determined by one of the attributes, called the class label attribute. In the context of classification, data tuples are also referred to as samples, examples, or objects. The data tuples analyzed to build the model collectively form the training data set. The individual tuples making up the training set are referred to as training samples. In the second step, the model is used for classification.

A classification technique is a systematic approach for building classification models from an input data set. There are many classification techniques. Among them, the system uses the Decision Tree Induction Algorithm. Making the right decision is becoming the key factor for the successful achievement of the goals in all areas of work. The ways of finding the right decision are as many as the number of people who have to make them. Nevertheless, the basic idea is the same for many of them: a decision is usually made as a combination of experiences from solving similar cases, the results of recent researches and personal judgment.

The number of solved cases and new researches is increasing rapidly. It could be expected that newly made decisions will become better and more reliable but for the individuals and groups who have to make decisions it is actually becoming more and more complicated, because they simply cannot process the huge amounts of data anymore. And there the need for a good decision support technique arises. It should be able to process those huge amounts of data and to help experts to make their decisions easier and more reliably. For this purpose it is equally or even more important as suggesting the possible decision, to provide also an explanation of how and why the suggested decision was chosen. In this manner an expert can decide whether the suggested solution is appropriate or not. As in many other areas, decisions play an important role also in medicine, especially in medical diagnostic processes. Decision support systems helping physicians are becoming a very important part in medical decision making. Since conceptual simple decision making models with the possibility of automatic learning should be considered for performing such tasks, decision trees are a very suitable candidate. They have been already successfully used for many decision making purposes [3].

In this paper an overview of decision trees is also presented with the emphasis on a variety of different induction methods available nowadays. Induction algorithms

San Hay Mar Shwe is with the University of Technology (Yatanarpon Cyber City) (e-mail: sanhaymarshwe@gmail.com).

Dr. Htet Ne Oo, is with the Faculty of Information and Communication Technology, University of Technology (Yatanarpon Cyber City), Myanmar. (e-mail: htetneoo.utycc@gmail.com).

ranging from the traditional heuristic based techniques to the most recent hybrids, such as evolutionary and neural network based approaches, are described. Basic features, advantages and drawbacks of each method are presented, biasing to the medical domain. For the readers not very familiar with the field of decision trees this paper can be a good introduction into this topic, whereas for more experienced readers it can broaden their perspective and knowledge.

## II. LITERATURE REVIEW

### A. Decision Tree Algorithm

A decision tree is formalism for expressing such mappings and consists of tests or attributes nodes linked to two or more sub-trees and leafs or decision nodes labelled with a class which means the decision. A test node computes some outcome based on the attribute values of an instance, where each possible outcome is associated with one of the subtrees. Inductive inference is the process of moving from concrete examples to general models, where the goal is to learn how to classify objects by analysing a set of instances (already solved cases) whose classes are known. Instances are typically represented as attribute-value vectors. Learning input consists of a set of such vectors, each belonging to a known class, and the output consists of a mapping from attribute values to classes. This mapping should accurately classify both the given instances and other unseen instances [3].

Data Mining is mainly used for specific set of six activities namely Classification, Estimation, Prediction, Affinity grouping or Association rules, Clustering, Description and Visualization. The first three tasks – Classification, estimation and prediction are all examples of data mining or supervised learning. Decision Tree is one of the most popular choices for learning from features. A decision tree is a flow-chart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent class or class distributions. The top most node in a tree is the root node. Decision tree is in the form of a top-down recursive divide-and-conquer manner. There are many specific decision-tree algorithms. Among them, this application is based on ID3 (Iterative Dichotomiser 3). It is a very simple decision tree learning algorithm which is used for the classification of the objects [7]. The following algorithm shows a version of ID3 algorithm.

**Algorithm:** Generate\_decision\_tree. Generate a decision tree from the given training data.

**Input:** The training samples, samples, represented by discrete-valued attributes; the set of candidate attributes, attribute\_list.

**Output:** A decision tree.

### Method:

- (1) Create a node N;
- (2) if samples are all of the same class, C then
- (3) return N as a leaf node labeled with the class C;
- (4) if attribute-list is empty then;
- (5) return N as a leaf node labeled with the mode common class in samples;
- (6) select test-attribute, the attribute among attribute-list with the highest information gain;
- (7) label node N with test-attribute;
- (8) for each known value  $a_i$  of test-attribute
- (9) grow a branch from node N for the condition test-attribute= $a_i$ ;
- (10) let  $s_i$  be the set of samples in samples for which test-attribute= $a_i$ ;
- (11) if  $s_i$  is empty then
- (12) attach a leaf labeled with the most common class samples;
- (13) else attach the node returned by Generate\_decision\_tree ( $s_i$ , attribute-list-test-attribute);

A decision tree can be built from a set of training objects with the “divide and conquer” principle. When all objects are of the same decision class (the value of the output attribute is the same) then a tree consists of a single node – a leaf with the appropriate decision. Otherwise an attribute is selected which value is of at least two different decision classes and a set of objects is divided according to the category of the selected attribute. The selected attribute builds an attribute (test) node in a growing decision tree, for each branch from that node the inducing procedure is repeated upon the remaining objects regarding the division until a leaf (a decision) is encountered [8]. This basic structure is shown in Fig. 1.

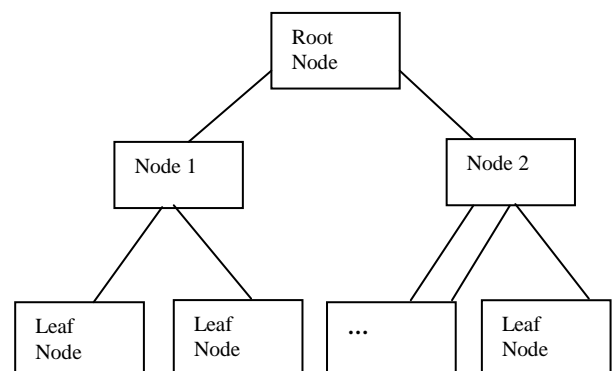


Figure 1. An example of a basic decision tree

### B. Fever Data

The fever problem is a data set which will be used to understand how a decision tree is built. It comes from Quinlan (1986), a paper which discusses the ID3 algorithm

introduced in Quinlan (1979). It is reproduced with slight modifications in Witten and Frank (1999), and concerns the conditions under what type of specific fever happens [5].

The different methods are shown in Table I. The data set for sample fifty data is shown in Table II.

TABLE I. VARIOUS DECISION TREE INDUCTION APPROACHES, EACH OF THE METHODS HAS SOME ADVANTAGES AND DISADVANTAGES SUMMARIZED [1][2][4][5][6]

Reference	Description/ Method	Induction approach	Discretization method	Space partitioning	Num. decision attributes
[Quinlan, 1993]	classical (ID3, C4.5, ...)	Heuristic	equidistant/ percentile/ dynamic	Orthogonal	One
[Babic, 2000]	fuzzy classical	heuristic	fuzzy	Orthogonal	One
[Zorman, 1999]	hybrid(MtDecit 2.0)	heuristic/ neural nets	dynamic	oblique	One
[Breiman, 1984]	classical (CART)	heuristic/ perturbations	equidistant/ percentile/ dynamic	oblique	One
[Heath, 1993a]	classical (SADT)	heuristic/ simulated annealing	equidistant/ percentile/ dynamic	Oblique	One
[Murthy, 1997]	classical (OC1)	heuristic/ random	equidistant/ percentile/ dynamic	Oblique	One
[Utgoff, 1989a]	classical incremental (ID5R)	heuristic/ incremental	equidistant/ percentile	Orthogonal	One
[Podgorelec, 2001a]	evolutionary (genTrees)	genetic algorithms	random	orthogonal	One
[Sprogar, 2000]	evolutionary vector (VEDEC)	genetic algorithms	random	Orthogonal	Any
[Podgorelec, 2001c]	Automatic programming (AREX)	gen. algorithms/ genetic programming	random	Oblique	One

TABLE II. SAMPLE DECISION RESULT FOR FEVER

Dysuria	Chill	Abdopain	Sore_throat	Headache	Cough	Pain	Chest_Pain	Inter_fev	Rashes	Class
Yes	No	no	No	No	No	no	No	yes	No	UTI
Yes	No	no	No	No	No	yes	No	yes	No	UTI
Yes	No	no	yes	No	No	yes	No	yes	No	UTI
Yes	No	no	yes	No	No	no	No	yes	No	UTI
Yes	normal	no	No	No	No	no	No	yes	No	UTI
Yes	normal	no	yes	No	No	yes	No	yes	No	UTI
Yes	normal	no	yes	No	No	yes	No	yes	No	UTI
Yes	normal	no	yes	No	No	yes	No	yes	No	UTI
Yes	normal	no	yes	No	No	yes	No	yes	No	UTI
Yes	normal	no	yes	No	No	yes	No	yes	No	UTI
No	normal	no	No	No	No	no	No	yes	No	Malaria
No	normal	no	yes	Yes	No	no	No	yes	No	Malaria
No	normal	no	yes	No	No	no	No	yes	No	Malaria
No	normal	no	No	Yes	No	no	No	yes	No	Malaria
No	High	no	yes	No	No	no	No	yes	No	Malaria

No	High	no	No	Yes	No	no	No	yes	No	Malaria
No	High	no	No	No	yes	no	No	yes	No	Malaria
No	High	no	yes	No	yes	no	No	yes	No	Malaria
No	High	no	No	Yes	yes	no	No	yes	No	Malaria
No	No	yes	No	No	No	no	No	no	No	Typhoid
No	No	yes	No	Yes	No	no	No	no	No	Typhoid
No	No	yes	No	No	No	yes	No	no	No	Typhoid
No	No	yes	No	Yes	No	yes	No	no	No	Typhoid
No	No	yes	No	No	No	no	No	no	Yes	Typhoid
No	No	yes	No	Yes	No	no	No	no	Yes	Typhoid
No	No	yes	No	No	No	yes	No	no	Yes	Typhoid
No	normal	yes	No	Yes	No	yes	No	no	Yes	Typhoid
No	normal	no	No	No	yes	no	No	no	Yes	Typhoid
No	normal	no	No	Yes	yes	no	No	no	Yes	Typhoid
No	normal	no	No	No	yes	yes	No	no	Yes	Typhoid
No	normal	yes	No	No	yes	no	No	no	Yes	Typhoid
No	normal	no	No	Yes	yes	yes	No	no	Yes	Typhoid
No	normal	no	yes	Yes	No	no	Yes	no	No	Influenza
No	normal	no	No	No	yes	no	Yes	no	No	Influenza
No	normal	no	yes	No	yes	no	Yes	no	No	Influenza
No	normal	no	No	Yes	yes	no	Yes	no	No	Influenza
No	normal	no	yes	Yes	yes	no	Yes	no	No	Influenza
No	normal	no	No	No	No	yes	Yes	no	No	Influenza
No	normal	no	yes	No	No	yes	Yes	no	No	Influenza
No	normal	no	No	Yes	No	yes	Yes	no	No	Influenza
No	High	no	No	Yes	yes	no	No	no	No	CC
No	High	no	yes	Yes	yes	no	No	no	No	CC
No	High	no	No	No	No	yes	No	no	No	CC
No	High	no	yes	Yes	yes	yes	No	no	No	CC
No	High	no	yes	No	No	yes	No	no	No	CC
No	High	no	No	Yes	No	yes	No	no	No	CC
No	High	no	No	No	yes	yes	No	no	No	CC
No	normal	no	No	No	No	no	No	no	No	CC
No	normal	no	yes	Yes	No	no	No	no	No	CC
No	No	no	No	No	yes	yes	No	no	No	CC

In this dataset, there are ten categorical attributes dysuria, chill, abdominal\_pain, sore\_throat, pain, headache, cough, chest\_pain, rashes, fever and class. We are interested in building a system which will enable us to decide what type of fever happens on the basis of the fever symptoms.

#### A. Entropy

In information theory, entropy is a measure of the uncertainty about a source of messages. The more uncertain a receiver is about a source of messages, the more information that receiver will need in order to know what message has been sent. For example, if a message source always sends exactly the same message, the receiver does not need any information to know what message has been

sent—it's always the same! The entropy of such a source is zero: there is no uncertainty at all. On the other hand, if a source can send  $n$  possible messages and each message occurs independently of the preceding message with equal probability, then the uncertainty of the receiver is maximized. The receiver will need to ask  $\log_2 n$  yes/no questions to determine which message has been sent, i.e. the receiver needs  $\log_2 n$  bits of information.

The average number of bits required to identify each message is a measure of the receiver's uncertainty about the source, and is known as the entropy of the source. Consider a source  $S$  which can produce  $n$  messages  $\{m_1, m_2, \dots, m_n\}$ . All messages are produced independently of each other and the probability of producing message  $m_i$  is  $p_i$ . For such a

source  $S$  with a message probability distribution  $P=(p_1, p_2, \dots, p_n)$ , the entropy  $H(P)$  is:

$$H(p) = H(P_1, P_2, \dots, P_n) = -\sum P_i \log_2 P_i \quad (1)$$

If a set  $T$  of records from a database (*i.e.* the training set for building the decision tree) is partitioned into  $k$  classes  $\{C_1, C_2, \dots, C_n\}$  on the basis of the output attribute, then the average amount of information (measured in bits) needed to identify the class of a record is  $H(P_T)$ , where  $P_T$  the probability distribution of the classes, estimated from the data as

$$P_T = (|C_1|/|T|, |C_2|/|T|, \dots, |C_k|/|T|) \quad (2)$$

The notation  $|C_i|$  means the number of elements in set  $C_i$ . Note that here we identify the entropy of the set  $H(T)$  with the entropy of the probability distribution of the members of the set,  $H(P_T)$  [8]. For the fever data

$$\begin{aligned} \text{Entropy} &= -\text{UTI}/50 \log_2 \text{UTI}/50 - \text{Malaria}/50 \log_2 \text{Malaria}/50 - \text{Typhoid}/50 \log_2 \text{Typhoid}/50 - \text{Influenza}/50 \log_2 \text{Influenza}/50 - \text{CC}/50 \log_2 \text{CC}/50 \\ &= -10/50 \log_2 10/50 - 9/50 \log_2 9/50 - 13/50 \log_2 13/50 - 8/50 \log_2 8/50 - 10/50 \log_2 10/50 \\ \text{Entropy} &= 0.464 + 0.445 + 0.505 + 0.423 + 0.464 \\ &= 2.301 \end{aligned}$$

**B. Information Gain**

Information gain  $IG(A)$  is the measure of the difference in entropy from before to after the set  $S$  is split on an attribute  $A$ . In other words, it is how much uncertainty in  $S$  was reduced after splitting set  $S$  on attribute  $A$ .

$$IG(A, S) = H(S) - \sum_{t \in T} P(t) H(t) \quad (3)$$

Where,  $H(S)$  = Entropy of set  $S$

$T$  = The subsets created from splitting set  $S$  by attribute  $A$  such that  $S = \cup_{t \in T} t$

$P(T)$  = The proportion of the number of elements in  $t$  to the number of elements in set  $S$

$H(t)$  = Entropy of subset  $t$

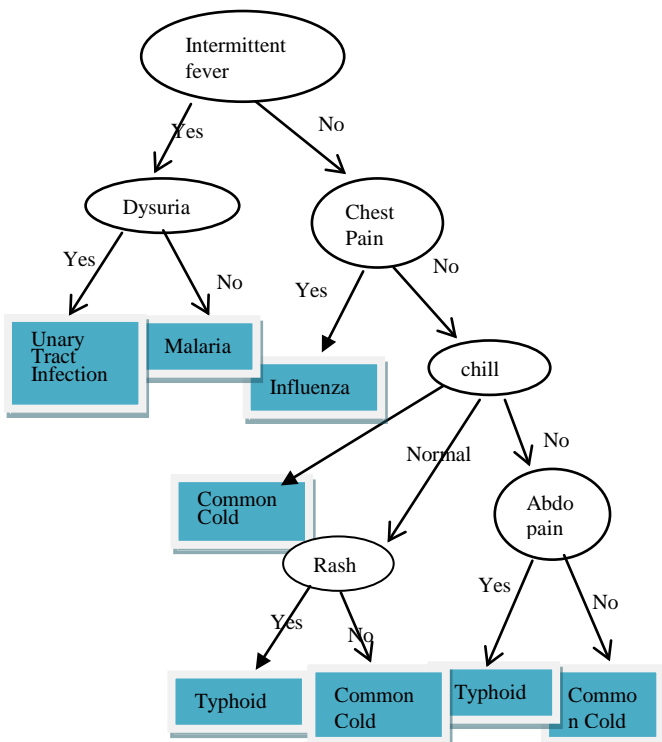


Figure 2. Decision tree of proposed system

In ID3, information gain can be calculated (instead of entropy) for each remaining attribute. The attribute with the largest information gain is used to split the set on this iteration [8].

$$\begin{aligned} G(S, \text{Dysuria}) &= 2.301 - \{10/50 (0) + 40/50 (-9/40 \log_2 9/40 - 13/40 \log_2 13/40 - 8/40 \log_2 8/40 - 10/40 \log_2 10/40)\} \\ G(S, \text{Dysuria}) &= 2.301 - (0 + 1.9755) = 0.3255 \end{aligned}$$

With the entropy and information gain, by calculating above the data set step by step, finally decision tree as in Fig. 2 is obtained.

**III. PROPOSED SYSTEM**

The proposed system has two main steps: classification and recommendation. The system classifies various kinds of fevers and suggests home remedies by using Decision tree system. Fevers include Malaria, Urinary Tract Infection, Typhoid, Influenza and common cold. Totally there are 10 attributes (symptoms) as input and five class attribute as output. The symptoms of each patient are Dysuria, Chill, Abdominal pain, Sore throat, Pain, Headache, Cough, Chest pain, Rashes and intermittent fever. In this step, the system can be divided into two phases. These are Data Training and Data Testing. In Data Training phase, attributes which are stored in Database can be calculated using ID3 Algorithm to appear Decision Tree rules. In Data Testing phase, user can input the attribute and then classify the fever group using Decision Tree Rules from Data Training Phase.

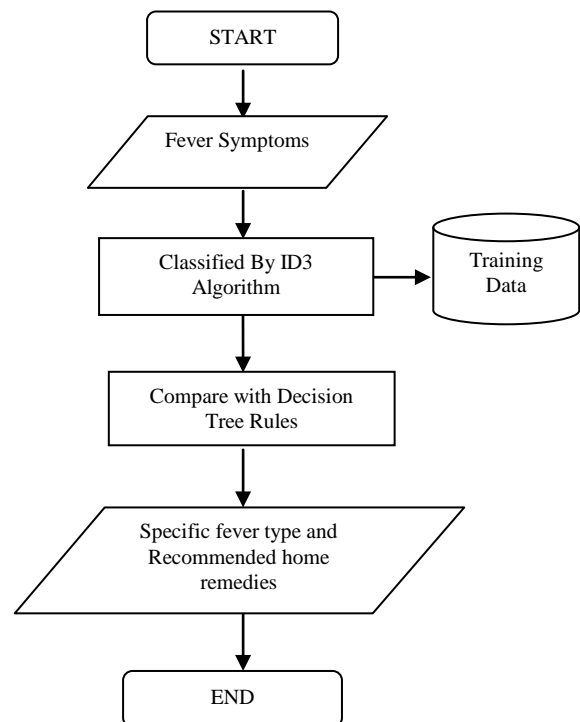


Figure 3. System Design for the Fever Classification system

In this way, the system can be classified into five fever group. Based on the classification, in second step, the system will guide various kinds of home remedies that are easily obtained everywhere. So, the user will cure the disease with natural way according to the specific fever. Besides, it will minimize the side effects by using home remedies and allows medical person to improve their diagnostic skill.

IV. IMPLEMENTATION OF THE SYSTEM

This system produces the specific fever which is occurred in patient. It explains the detail facts about each disease. Based on the disease, the user will also cure the disease with natural way. Therefore, it will give effect in the easiest, fastest, safest, and cheapest way. Moreover, the various kinds of home remedies can be known widely.



Figure 4. User interface for fever classification

By clicking classify button, it will obtain the corresponding class as in Fig. 5. If user wants to cure home remedies, user must be select yes button. And the result is shown in Fig. 6.



Figure 5. Result for fever type



Figure 6. Treatment for resulting fever

From this page, the user can also know the detail information about disease by clicking about typhoid button as follows.



Figure 7. Explanation about result disease

In Fig. 8, comparison of two decision tree resulting from weka application and ID3 Algorithm is shown.

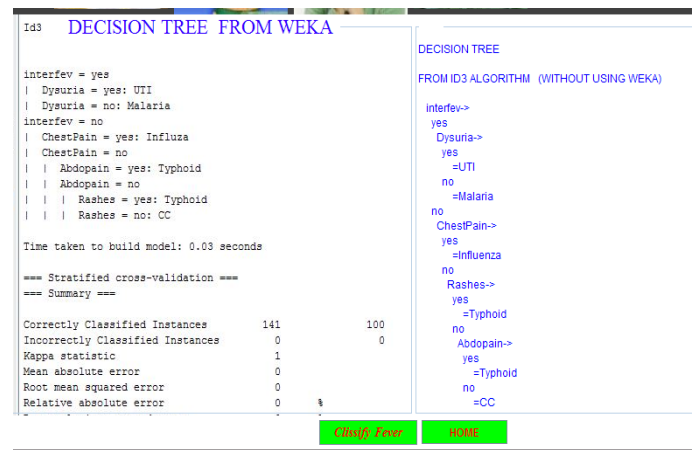


Figure 8. Decision tree of the system

V. CONCLUSION

This paper describes fever classification and remedial recommendation system by using decision tree algorithm and also represents basic characteristics of decision trees. In medical decision making (classification, diagnosing, etc.) there are many situations where decision must be made effectively and reliably. Conceptual simple decision making

models with the possibility of automatic learning are the most appropriate for performing such tasks.

This classification system can help the user give exact fever types by analyzing the symptoms. And then, it also gives appropriate home remedies. Moreover, it allows medical person to improve their diagnostic skill. According to the facts mentioned above, hope the system will be effective and useful for everyone. In addition to those facts, various kinds of decision tree approaches can also be known.

#### ACKNOWLEDGMENT

The author would like to place on record her deep sense of gratitude to Professor Dr. Aung Win, Rector, University of Technology (Yatanarpon Cyber City) for his generous guidance, help and useful suggestions. The author is extremely thankful to her supervisor, Dr. Htet Ne Oo, for her invaluable guidance and kind supervisions. The author would also thank to her parents and all the teachers who taught her throughout the whole life.

#### REFERENCES

[1] Bonner, G., Decision making for health care professionals: use of decision trees within the community mental health setting.

[2] Cremilleux, B., Robert, C., A theoretical framework for decision trees in uncertain domains: Application to medical data sets, *Lecture Notes in Artificial Intelligence*, vol. 1211, pp. 145-156, 1997.

[3] Han J., Pei J., Kamber M. *Data mining: concepts and techniques*. – Elsevier, 2011.

[4] Dietterich, T.G., Kong, E.B., *Machine learning bias, statistical bias and statistical variance of decision tree algorithms*, *Machine Learning*, 1995.

[5] Gambhir, S.S., *Decision analysis in nuclear medicine*, *Journal Of Nuclear Medicine*, vol. 40, num. 9, pp. 1570-1581, 1999.

[6] Murthy, K.V.S., *On Growing Better Decision Trees from Data*, PhD dissertation, Johns Hopkins University, Baltimore, MD, 1997.

[7] Witten, I. H. and Frank, E. (1999). *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann, San Francisco, CA, USA. <http://www.cs.waikato.ac.nz/ml/weka/book.html>

[8] CSE5230 Tutorial: The ID3 Decision Tree Algorithm [http://web.info.uvt.ro/~dzaharie/dm2016/projects/DecisionTrees/DecisionTrees\\_ID3Tutorial.pdf](http://web.info.uvt.ro/~dzaharie/dm2016/projects/DecisionTrees/DecisionTrees_ID3Tutorial.pdf)

[9] Black N. T., Ertel W. *Introduction to artificial intelligence*. – Springer Science & Business Media, 2011.

**San Hay Mar Shwe** is a B.E student at Faculty of Information and Communication Technology, University of Technology (Yatanarpon Cyber City), Myanmar. Her area of interest is artificial intelligence and machine learning.

**Dr. Htet Ne Oo** is an Assistant Lecturer at Faculty of Information and Communication Technology, University of Technology (Yatanarpon Cyber City), Myanmar. Her specialization includes cryptography and network security and machine learning.