Feature Selection For Medical Diagnosis Of Fever

War War Min
Computer University, Monywa, Myanmar
warwarmin84@gmail.com

Abstract

The system based on feature selection methods of artificial intelligence (AI). Artificial Intelligence becomes the major subject. Machine learning technology is one of the branches of the AI. The human brain provides inspiration for artificial intelligence researchers. It will react with the human behavior understanding and predicting the possible results. The symptoms will be accepted. By using the feature selection method of AI to the accepted symptom, the result will be displayed. The result will be the type of fever for the symptom. It is aim to develop a computerized diagnosis system for the types of fever. The aim of the paper is to develop a computerized diagnosis system for the types of fever. The facts and figures of the problem-domain expert are gained and converted to decision tree. The decision tree is used to determine the types of disease by feature selection.

1. Introduction

There are many types of fever. The symptoms of a fever are alike to the symptoms of another fever. To distinguish between the different types of fever the patient has to be diagnosed by the human expertise of that field area. Today Artificial Intelligence is widely used in every area. Expert system is one of the ways to substitute the requirement of human expertise with the machine with the help of the expert system. This system is suitable for the locations where human expertise is not available. By running the system and answering the system's predefined questions and the system can distinguish between the types of fever. The system is an expert system that will diagnose the patient's symptom. The proposed system uses the decision tree method to distinguish the symptoms and diagnosis the patient's disease by using feature selection method.

An expert system is software that attempts to reproduce the performance of one or more human experts, most commonly in a specific problem domain, and is a traditional application and/or subfield of artificial intelligence. A wide variety of methods can be used to simulate the performance of the expert however common to most or all are (1) the creation of a so-called "knowledge" which uses some knowledge representation formalism to capture the Subject Matter Expert (SME) knowledge and (2) a process of gathering that knowledge from the SME and codifying it according to the formalism, which is called knowledge Engineering. Expert systems may or may not have learning components but a third common element is that once the system is developed it is proven by being placed in the same real world problem solving situation as the human SME, typically as an aid to human workers or a supplement to some information system.

The system is the one of the decision support systems (DSS). Decision Support Systems (DSS) are a specific class of computerized information systems that supports business and organizational decision-making activities. A properly-designed DSS is an interactive software-based system intended to help decision makers compile useful information from raw data, documents, personal knowledge, and/or business models to identify and solve problems and make decisions.

Typical information that a decision support application might gather and present would be an inventory of all of your current information assets (including legacy and relational data sources, cubes, data warehouses, and data marts), comparative sales figures between one week and the next, projected revenue figures based on new product sales assumptions, the consequences of different decision alternatives, given past experience in a context that is described.

2. Related work

P.G.Neville et al provided decision analysis, a "decision tree" — and the closely-related influence diagram — is used as a visual and analytical decision support tool, where the expected values (or expected utility) of competing alternatives are calculated [4].

A decision tree (or tree diagram) is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal.

Another use of decision trees is as a descriptive means for calculating conditional probabilities [1].

V.Verma et al established data mining and machine learning, a decision tree is a predictive model; that is, a mapping from observations about an item to conclusions about its target value. More descriptive names for such tree models are classification tree (discrete outcome) or regression tree (continuous outcome). In these tree structures, leaves represent classifications and branches represent conjunctions of features that lead to those classifications. The machine learning technique for inducing a decision tree from data is called decision tree learning, or (colloquially) decision trees [5].

There are two steps to making productive use of decision trees

(a) building a decision tree model

(b) using the decision tree to draw inference and make prediction

The first step in building a decision trees is to collect a set of data values that the proposed can analyze. This data is called the "learning" or "training" dataset because it is used by the proposed system to learn how the value of a target variable is related to the values of predictor variables. This dataset must have instances for which you know the actual value of the target variable and the associated predictor variables. You might have to perform a controlled study to collect this data, or you might be able to obtain it from previously-collected historical records. These records can be gained from the human expertise [3].

F. Coenen et al expressed decision can be used for the following purposes to see the big picture of the whole process, to be used as a communication tool with the non-technical people, and to predict the highest possible end-nodes easily [2].

3. Methods

The following shows the associate symptoms of the fever. These symptoms can be used to distinguish the types of diseases. The no column given at the following table will be used instead of symptom-description. The different kind of symptoms, different kind of disease, different kind of fever and relationships among them come from the Davidson's Principle and Practice of Medicine (20th Edition).

Table 1. Fever symptoms of the system

No	Associated Symptom of fever
1	Chill
2	Rigor
3	Headache
4	Vomiting
5	Blurred Vision
6	Conversion
7	Loss of consciousness
8	Paraplegia
9	Cough
10	Tightness of chest
11	Expectoration
12	Chest Pain
13	Skin Rashes
14	Haemoptysis
15	Dyspnoea
16	Horsness of voice
17	Sorethroat
18	Pain in the mouth
19	Pain in the gum
20	Swelling of cheek
21	Yellow collouration of urine and sclera
22	Pain in right lower chest
23	nausea
24	pain in epigastrium
25	pain in right lower abdomin
26	cool and clamsy hand and feet
27	Joint pain
28	Dysuria
29	Frequency
30	Pain in suprapublic area
31	haematuria
32	pain in right or left loin
33	cloudy urine
34	pain in urethral
35	urethral discharge
36	pain in lower abdomin
37	white discharge from female organ
38	Testicular pain
39	skin ulcer
40	skin abscess
41	chronic pusdischarge skin ulcer
42	loose motion
43	blood and mucus stool
44	Increased respiration rate
45	muscular aches pain
46	sneezing

The names of the diseases that the system can predict are shown in the following table

Table 2. Possible diseases that the system can predict

No	Disease Name	Short number
1	Encephalitis	A1
2	Bronchitis	A2
3	Pulmonary Tuberculosis	A3
4	Urinary tract infection	A4
5	Urethritis	A5
6	Chronic Skin ulcer with osteomyelitis	A6

7	Pneumonia	A7
8	Common Cold	A8
9	Oralsepsis	A9
10	mumps	A10
11	Viral hepatitis	A11
12	Appendicitis	A12
13	Pancreatitis	A13
14	Arithritis	A14
15	Orchitis	A15
16	Skin ulcer with sepsis	A16
17	Meningitis	A17
18	Measles	A18
19	Skin abscess with sepsis	A19
20	Malaria	A20
21	Enteric Fever	A21

Table 3. Relationship of Symptom, Fever-temperature and Disease

Continuous fever	
Symptoms	Disease
1, 3, 23, 42	A21
Remittent fever	
Symptoms	Disease
1, 2, 3, 45	A20
Intermittent fever	
Symptoms	Disease
1, 2, 3, 45	A20
3, 4, 5, 6, 7, 8	A1
9, 10, 11, 12	A2
28, 29, 30	A4
Hyperpyrexia	
Symptoms	Disease
3, 4, 5, 6, 7, 8	A17
9, 13, 42, 46	A18
1, 2, 3, 45	A20
Low-temperature-fever	
Symptoms	Disease
3, 4, 5, 6, 7, 8	A1
6, 9, 10, 11, 12	A3
7, 9, 10, 11, 12	A2
28, 29, 30	A4
28, 29, 34, 36, 41	A5
41	A6
Moderate-temperature-fever	
Symptoms	Disease
3, 4, 5, 6, 7, 8	A1
9, 10, 11, 12,15, 44	A7
9, 17, 45, 46	A8
20	A10
18, 19, 20	A9
21, 22, 23	A11
4, 25	A12
4, 24	A13
4, 15	A14
38	A15
39, 45	A16

High-temperature-fever	
Symptoms	Disease
3, 4, 5, 6, 7, 8	A17
9, 10, 11, 12, 15, 44	A7
9, 17, 45, 46	A8
9, 13, 42, 46	A18
20	A10
4, 25	A12
4, 15, 27	A14
38	A15
39,45	A16
40, 45	A19
41	A6

The system include the following process

(a) Accepting user input for temperature
(b) Generating the decision tree for the user-selected temperature
(c) Ask the user about the symptoms of the fever.
(d) Predict and response backs the diseases.

There are seven different decision trees used in the system. Each decision tree is for different types of temperature. The different types of temperature will be used as a major differentiation of the fever symptoms and diseases.

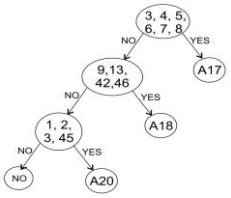


Figure 1. Decision Tree for Hyperpyrexia

The tree has totally seven nodes, including the root node. There are three possible diseases for the decision tree. There are four possible answers that the system can generate, including 'No Exact Match Found'.

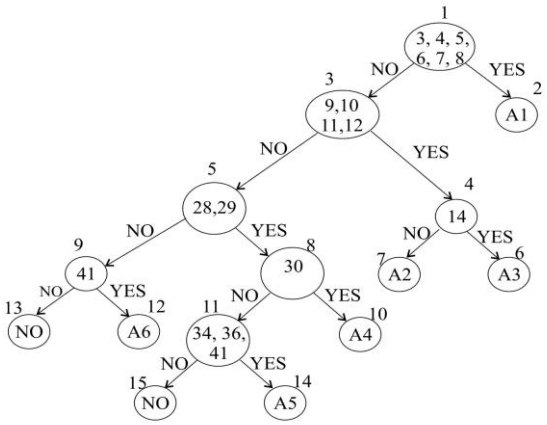


Figure 2. Decision Tree for low-temperature fever

The tree has totally fifteen nodes, including the root node. There are six possible diseases for the decision tree. There are eight possible answers that the system can generate, including 'No Exact Match Found'.

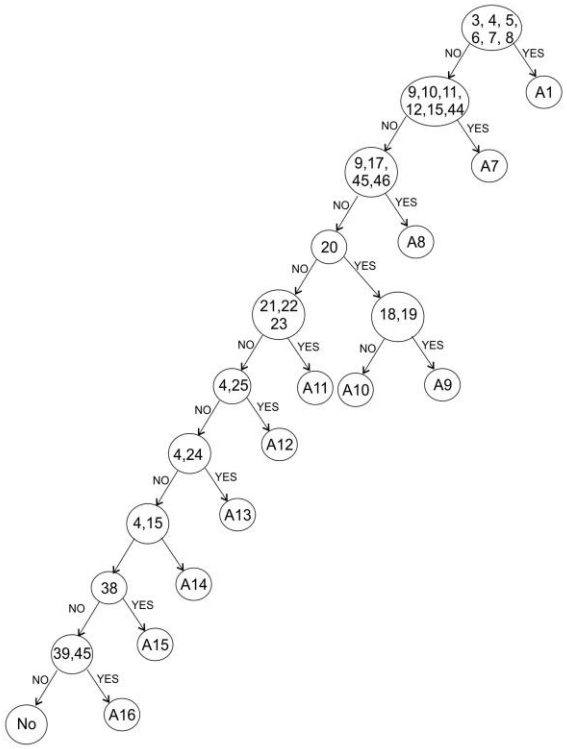


Figure 3. Decision Tree for moderate-temperature fever

The tree has totally seven nodes, including the root node. There are three possible diseases for the decision tree. There are four possible answers that the system can generate, including 'No Exact Match Found'.

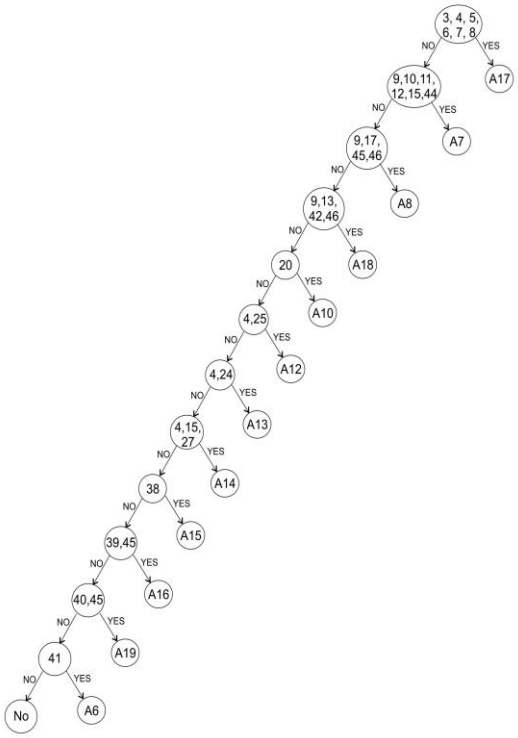


Figure 4. Decision Tree for high-temperature fever

The tree has totally twenty five nodes, including the root node. There are twelve possible diseases for the decision tree. There are thirteen possible answers that the system can generate, including 'No Exact Match Found'.

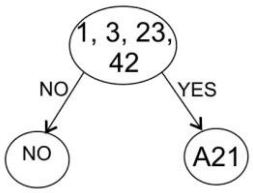


Figure 5. Decision Tree for Continuous fever

The tree has totally three nodes, including the root node. There is one possible disease for the decision tree. There are two possible answers that the system can generate, including 'No Exact Match Found'.

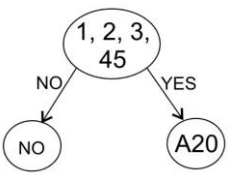


Figure 6. Decision Tree for Remittent fever

The tree has totally three nodes, including the root node. There is one possible disease for the decision tree. There are two possible answers that the system can generate, including 'No Exact Match Found'.

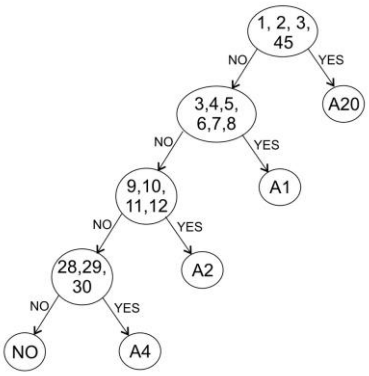
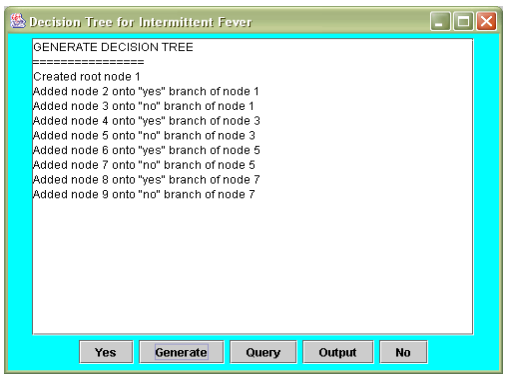
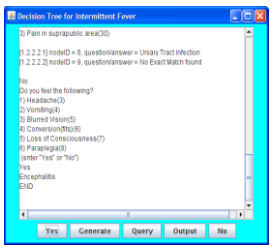


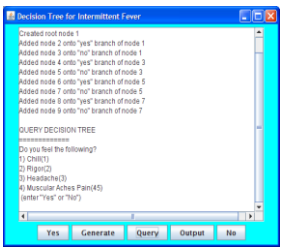
Figure 7. Decision Tree for Intermittent fever

The tree has totally nine nodes, including the root node. There are four possible diseases for the decision tree. There are five possible answers that the system can generate, including 'No Exact Match Found'.

4. Result







If patient inputs the symptoms which are related to Fever diseases and then the system's diagnosis process would generate the patients may be suffered Fever disease.

If patient inputs the symptoms which are not related to Fever diseases and then the system's diagnosis process would generate the patients may be suffered some diseases which have common symptoms with Fever disease.

If result has duplicate messages then there is more probabilities to suffer this disease described in the message.

The system can manipulate on the average of 90% accuracy. But for some diseases with distinct symptoms, the system can manipulate up to 100% accuracy such as Malaria, Pneumonia, Meningitis.

5. Conclusion

The system develops for a medical diagnose system. It distinguishes among the types of fever with associated fever. For the first, the user must enter the patient's temperature. After that the decision tree for the user-input-temperature-type is generated. The decision tree method is used. Then, the associated symptoms of fever are asked from the user via decision tree interactively and the result-disease is predicted and shown to the user.

The decision is created by using the decision tree which is in turn created by transforming form the recorded facts and figure of the problem domain expert. In this case, the problem domain expert is the Physician. The problem domain expert's pre-recorded facts and figures are used to transform a decision tree.

The system has the following advantages:

Provides is consistent answers for repetitive decisions, processes and tasks. Holds and maintains significant levels of information. Encourages organizations is to clarify the logic of their decision-making. Never "forgets" to ask a question, as a human might.

6. Further works

This system is based on the fever-distinguishing data. Other data set like this can also be added to this system. So the systems can diagnosis for other disease.

By adding the computer vision handling portion for the system, the system can now determine the physical appearance of the patient. Then, the system can make more accurate diagnose based on the visual appearance of the patient. The system can then determine the physical appearance of the patient as the human expert can see. Image processing for the computer vision is needed to be added.

7. References

[1] Decision Tree Lecture Notes
http://www.acm.org

[2] F. Coenen, *"A decision tree implementation paper"*

[3] Introduction to decision tree
http://www.dtreg.com

[4] P.G. Neville, *"Decision Trees for Predictive Modeling"*

[5] V.Verma *"Decision Tree Learning"*

[6] Stanley Davidson, *"Davison's Principle and Practice of Medicine"*