IJCSI International Journal of Computer Science Issues, Volume 17, Issue 3, May 2020
ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784
www.IJCSI.org https://doi.org/10.5281/zenodo.3987155

A Machine Learning for an Expert System on Behalf of Decision Making: Application to Lumbago

Judith Mosingalete Lumbu [1] Gabriel Nguma Maphana[2] Tshims Tshimanga Mutombo [3] Simon Ntumba badibanga [4]

[1] Mathématiques et Informatique, Université de Kinshasa, Kinshasa, DRC

[2] Mathématiques et Informatique, Université de Kinshasa, Kinshasa, DRC

[3] Mathématiques et Informatique, Université de Kinshasa, Kinshasa, DRC

[4]Mathématiques et Informatique, Université de Kinshasa, Kinshasa, DRC

Abstrat

In this articles, we intend to set an expert system to contribute to the making of good decisions a solution based on the resolution to separate the base of knowledge part involving the expertise under the form If … Cond … Action.

This study is on lumbago or backache, which is not an illness but a symptom. This is about pain in the back, which is acute if it lasts less than a month and chronic beyond three months. This pain is common, almost everyone suffers from i tat one stage of his/her life. The study covers a period going 2015 to 2018 at the hospital of Aenemia at Yolo

Machine learning is based on the idea that system scan learn about data, identify trends and take decisions with less human intervention. It automates the creation of analytical models.

Key words*: machine learning, aid to decision, lumbago, expert system.*

1. Introduction

The degree and the growing complexity of medical knowledge such as diagnostic knowledge means enable the physiotherapist to manage more information to treat a patient, he must take a number of decisions resulting in the medical act, that occurs according to logical reasoning. The development of the aid system to decision making consists of analyzing the medical decision that means a study of the process and the conditions of reasoning.

The physiotherapist collects signs and symptoms, while listening and examining the patient, and then he figures out diagnostic hypotheses likely to accounts for the signs observed.

Regarding the nature of the medical decision, data are often imprecise, ambiguous and incomplete: a given sign can be present or absent in the same illness; the same sign can be present in several illness. The results of additional examinations do not only have an imperfect reliability. This compels the physiotherapist to take his/her decisions in an uncertain climate according to a particular reasoning. The problem is to know how to recognize that from recurring symptoms in the patient, it can be established that it is lumbago.

Therefore we are going to set up an expert system likely to help decision making thanks to machine learning which brings support to and in the area of decision making concerning tasks to accomplish to reach a stable state from the physiotherapist's point of view. Machine learning is about studying a statistical model throngh the computer thanks to driving data.

2. Expert system [2, 10]

An expert system is an intelligent computer program which uses a body of knowledge and inference procedures witch the aim of solving problems of such difficulty that thry require an adapted human expertise. It is a system of assistance to decision based on an inference engine and a data base.

Furthermore an expert system is a program having multiple data in a specialized field, there data generally coming from a human expert and likely to reach the performance of an expert in that field.

2.1 Components and architecture of an expert system [1, 2, 8, 9]

✓**Components**

The heart of an expert system has three wain components:

1. Database: it has all that we know on data to be managed or all the data resulting from the expert which allow him to diagnose.
Expert data are represented in the form of a definite production rule: if … Condition then … Action
2. Datamart: it has what the expert system knows on the case being studied.
3. Inference engine: it manages the data base relying on the questions asked the facts verified.

✓Simple architecture of expert system

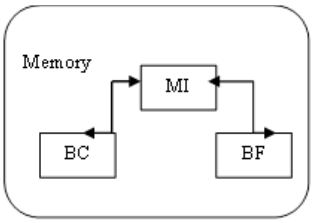

Fig. 1: Simple architecture of expert system

2.2 Management of an expert system [2, 3, 11]

The expert system is managed from the following elements: the database made of the data mart, rules and inference engine. The Datamart contains the different facts weful to the application stage; facts are variables describing the world. The rules base has expert data, that means they represent the reasoning process done by an expert. The inference engine allow to proced with logical reasoning using in a parallel fashion facts and rules.

We have used the production line, and the inference engine is run in that mode when the Datamart represents information that proved to be true, that means this mode goes from data to the goals.

✓**Production line algorithm [9, 10, 11]**

At the start: Datamart, base of rules, fact to be demonstrated
Beginning
As long as
The fact to be demonstrated is not in the Datamart and there is in the base of rules a rule which is applicable
Do
Choosing Base of rules=BR-R (deactivation of R)
Datamart=BR ∪ Conclusion (triggering off of the rule R its conclusion is added to the Datamart)
End of as long as.
If FAD belongs database then FAD is established
Otherwise FAD is not established.

3. Machine learning [3, 4, 5]

Machine learning is a technique which consists of teaching computers what humans are naturally able to do: to draw lessons from their experiences. Otherwise, machine learning enables computers to learn without being programmed. Explicitly to that end to learn and get developed computers nevertheless need to analyze and to train with.
Machine learning permits value extraction from the source of various data without banking on a human being, contrary to traditional analytical tools, it can also be injected into the machine learning system, the more that system can learn and attach results to samples of a superior quality.
Thus machine learning permits the discovery of models hidden in the data more efficiently than human intelligence can. Cognitive computer systems constantly learn in the health area and intelligently predict their trends. The level of cognitive applications is defined by four main characteristics: the understanding of no structured data, the capacity of reasoning and extracting and extracting ideas, the capacity to refine expertise at each interaction, and the ability to see, speak, and hear, to interact with humans in a natural fashion. To that purpose, it is useful to natural language.

3.1 How machine learning works [5, 12]

Machine learning is based on two types of techniques: supervised learning which consists in basing a model on input and output data, so that it can predict future results, and unstructured or unsupervised learning which identifies hidden models or intrinsic structures in the input data.

Supervised machine learning sets a model which makes prediction basing itself on evidence in case of uncertainty, a applies a given number of know input data and know answers to data (results) and

results in a model likely to provide reasonable predictions as answers to new data.

3.2 Iterative process of machine learning [5, 6]

In the iterative process, we have the feature engineering phase which consists of creating a new feature from what we have. To separate data, in three phases or steps:

- Train: which allows to build a model and the hyper parameter;
- Validation: make predictions, this phase allows to recover results and see if they can be generalized when the result is correct, we more to the last phase.
- Test: we keep the data and compare the result

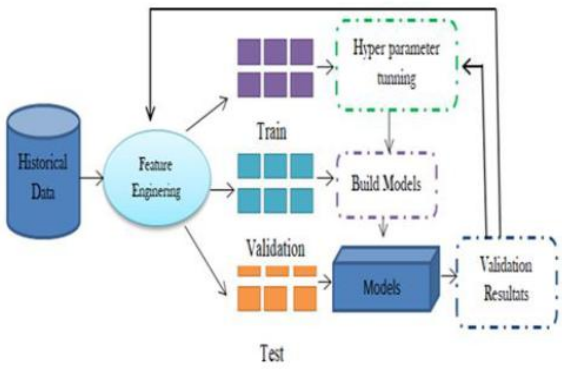

Fig. 2: Iterative process of machine learning

3.3 Bayesian channel [1, 5, 6, 7]

The bayesien channel is one trend of machine learning, it is a probalistic graphic model representing a number of random variables in the form of an acyclic oriented graph. A bayesian channel is at the same time:

- A model of representation of knowledge;
- A computing machine for conditional probabilities;
- A basis for systems impacting decision.

In this medical field, we describe the causal relations between interest variables through graph. In this graph, causes to effect relations between variables are not deterministic but probabilistic. Therefore we will have the conditional probability which is as follows:

$$P(A/B)=P(A\cap B)/P(B) \qquad (1)$$

This is transformed into Bayesian probability:

Hypotheses

$$\begin{cases} Ai \neq \emptyset, \forall = 1,2,\dots,n \\ Ai \cap Aj = \emptyset, i \neq j \\ \cup Ai = \Omega \quad \text{Where i=1to n} \end{cases} \qquad (2)$$

Given the events A1, A2,… An forming a complete system (that means n event entirely exclusive), that means:
Suppose another event B sould be achieved only in combination with one of the events Ai (i=1,2,…,n) that means.

$$B=\cup (A \cap B)$$

$$P(Ak\backslash B) = \frac{P(Ak).P(B\backslash Ak)}{\Sigma P(Ai)P(B\backslash Ai)} \text{ Where k } \in 1,2,\dots, \text{ n and } B\neq\emptyset \qquad (3)$$

Demonstration:
Through a definition of conditional probabilities:

$$P(A_k\backslash B) = \frac{P(Ak\cap B)/P(B)}{P(B)} = \frac{P(Ak).P(B\backslash Ak)}{P(B)} = \frac{P(Ak).P(B\backslash Ak)}{\Sigma P(Ai)P(B\backslash Ai)} \qquad (4)$$

4. Application

This consists of representing the inference engine based on the bases of the data and the base of rules in the following manner.

4.1 Knowledge representation

Base of facts

- Bad sitting or laying position
- Pain in the back area
- Tense muscles very long
- Depression
- Extended stress
- Muscular tensions
- Lumbago
- Back pain for less than a month
- Acute lumbago
- Back pain for more than 3 months
- Chronic lumbago
- Acute lumbago after lifting weight
- Chronic lumbago presenting geodes
- Nervous compression
- Strainde muscle
- Mechanic lumbago

- Radiationg sciatic pain through feet
- Severe cough and sneezing
- Slipped disk
- Articular pain
- Muscle pain
- Rhumatism
- Articulat rhumatism
- Arthritis
- Stiffness
- Lack of mobility
- Radiating pain
- Pain with headache and dizziness
- Cervical arthritis
- Back pain in the morning when you get up
- Lombarthritis
- Nervous compression
- Sciatic lumbago
- Pain even when resting
- Inflammatory lumbago
- Bad pôsition
- Minor lumbago
- Bad weight lifting
- Arthritis of the spine
- Lombarthritis
- Cervis arthritis
- Inter apophysis arthritis
- Knee arthretis
- Gonarthritis
- Shoulder arthritis
- Omarthritis
- Spine arthritis
- Back arthitis

Base of rules

R1 If:

- ✓ Bad position
- ✓ Back pain
- ✓ Tense muscle for very long
- ✓ Depression
- ✓ Extended stress

Then lumbago

R2 If:

- ✓ Back pain for less than a month

Then acute lumbago

R3 If:

- ✓ Back pain for more than 3 months

Then chronic lumbago

R4 If:

- ✓ Acute lumbago after lifting weight

Then lumbago

R5 If:

- ✓ Bad weight lifting

Then nervous compression

R6 If:

- ✓ Nervous compression

Then sciatic lumbago

R7 If:

- ✓ Nervous compression
- ✓ Muscle stretching

Then mechanic lumbago

R8 If:

- ✓ Pain even when resting

Then inflammatory lumbago

R9 If:

- ✓ Bad position

Then minor lumbago

R10 If :

- ✓ Articular ^pain

Then articular rhumatism

R11 If:

- ✓ articular rhumatism

Then arthritis

R12 if:

- ✓ articular pain in the neck
- ✓ stiffness
- ✓ lack of mobility
- ✓ radiating pain
- ✓ pain with headache and dizziness

Then cervic arthritis

R13 If:

- ✓ cervic arthritis

Then inter apophysis arthritis

R14 If:

- ✓ sciatic pain
- ✓ severe cough and sneezing

Then slipped disc

R15 If:

- ✓ knee arthritis
- ✓ unilateral or bilateral pain

Then gonarthritis

R16 If:

- ✓ shoulder arthritis

Then omarthritis

R17 If:

- ✓ spine arthritis

Then backarthritis

R18 if:

- ✓ back pain
- ✓ chronic lumbago presenting geodes

Then lombarthritis

4.2 The representation of the rule of production can be summed up as follows: [7, 9, 11]

R1: F1, F2, F3, F4, F5 —> F6
R2: F7 —> F8
R3: F9 —>F10
R4: F11 —>F12
R5: F13 —> F14
R6: F14 —> F15
R7: F14, F15 —> F16
R8: F17 —> F18
R9: F1 —> F19
R10: F20 —> F21
R11: F21 —> F22
R12: F23, F24, F25, F26, F27 —> F28
R13: F28 —> F29
R14: F30, F31 —> F32
R15: F33, F34 —> F35
R16: F36 —> F37
R17: F38 —> F39
R18: F2, F10 —> F40

4.3 Result

In the application of the inference engine, we based ourselves on the second process:
The solution corresponds to a selected rule among the applicable rules, the one which has the most mumerous conditions, in case of equality; it is the rule of the weakest number which will be used.

BF: {F1, F2}
But: F6
1st cycle:
S= {R1, R9, R18}
F= {R9}
R= R9
E=F19
R9
BF= {F1, F2, F19}
2nd cycle:
S= {R1, R18}
F= {R18}
R= R18
E=F40
R18 BF= {F1, F2, F19,F40}

3rd cycle:
S= {R1}
F= {R1}
R= R1
E=F6
R1
BF= {F1, F2, F19, F40, F6}

Interpretation
In using the bayesian network, events representd by facts are repetitive symptoms appearing in a patient. Thus, we will have:

F1: a bad sitting or bying position, 9 persons with 0,025%
F2: pain in the back, 15 persons with 0,018%
F19: pain even while resting, 13 persons with 0,012%
F40: cervic arthritis, 8 persons with 0,014%
F6: lumbago, 5 persons with 0,038%
R: the probability of having one of the symptoms during 3 years.
According to Bayes,

$$P(F6/R) = \frac{P(F6).P(R/F6)}{P(R)}$$

$$P(R) = \Sigma\, P(Fi)x\, P(R/Fi) \quad \text{where } i=1,...5$$

The event R can be realized only in combination with one of events F1, F2, F19, F40 or F6.
P (R) = P (F1) x P(R/F1) + P (F2) x P (R/F2) + P (F19) x P (R/F19) + P (F40) x P(R/F40) + P (F6) x P (R/F6)
= (2, 5 x 1/9) + (1, 8 x 1/15) + (1, 2 x 1/13) + (1, 4 x 1/8) + (3, 8 x 1/5)
=0, 825+ 0, 36+ 0, 27+0, 518+2, 28
=4, 259

Definitely
P(F6/R)= (P(F6).P(R/F6))/(P(R)) =(3,8x1/5)/4,259
P(F6\R)=0,535=53%

So, 53% of patients have suffered from lumbago in the period going from 2015 to 2018.

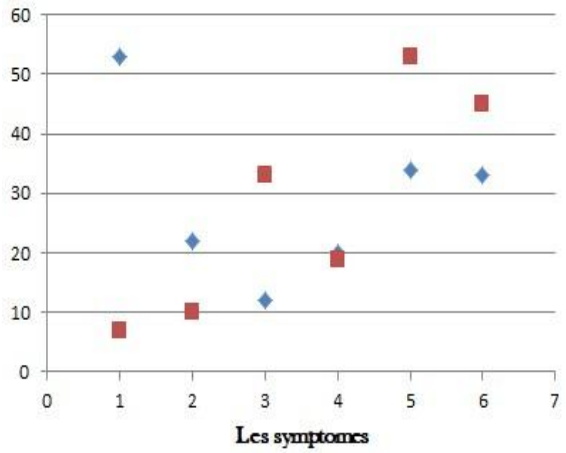

Fig. 3: representation of result

The results show that all the cloud points which adjust the straight line reveals that 53% of patients have suffered from back achs.

5 Implementation

The implementation is conceived in Java which is an interpreted language, which means that a compiled program can not be directly executed by the exploitation system, but it must be interpreted by another program.

5.1 Some codes on application

```
Package application_aia ;

Import java.sql.resultSet;
Public class base_fact extends javax.swing.JFrame
{
   int line = 0;
String tabcol[][];
public base_fact() {
initComponents();
}     private void display() {         String
req = "select * from T_fact";
       try {
           connexion b = new
connexion();            ResultSet rep =
b.resultat(req);            line = 0;
while (rep.next()) {                line +=
1;
          }
tabcol = new String[line][2];
          b.close();            connexion bbe =
new connexion();            ResultSet repbe
= bbe.result(req);
          int j = 0;            while (repbe.next()) {
tabcol[j][0] = repbe.getString("id_fact");
tabcol[j][1] = repbe.getString("libelle");
j += 1;
```

Conclusion

The study in this article deals with the application of an expert system useful to decision making in the case of lumbago or backache; the main purpose was to help a specialist or expert (a physiotherapist) to recognize from prevailing symptoms that the illness is lumbago.

The result revealed by the inference engine which aims at detecting facts to reach the aim applying the production rule. Thus 53% of patients have suffered from backache; that was revealed following the application of machine learning.

The expert system thus conceived resulted in an efficient solution, quick and reliable for the physiologist for the patient's diagnostic and good decision making. Furthermore, the system enabled a deep analysis with a database from the experience acquired by the expert in the medical area, it also enabled to assess in real time the state of the patient's health. The technology has also been useful to analyze data to free out trends or alert indicators allowing or resulting in the improvement of diagnostics and treatments.

Considering the growing volume of data, store medical there must be a confrontation with new challenges. The integration of data for better analytic results, and the accessibility of data are at stake. Machine learning proved to be very useful and has met these different challenges.

References

[1]. Boyne, S., D. Hall, and F. Williams, Policy, support and promotion for food-related tourism initiatives: A marketing approach to regional development. Journal of Travel & Tourism Marketing, 2003. 14(3-4): p. 131-154.

[2]. Emmanuel Grolleau, introduction to machine learning, observatory Paris Lesia, August 2017.

[3]. Esma Aimeur, introduction to artificial intelligence, University of Montreal Canade, 2015.

[4] Eugène Mbuyi Mukendi, Simon Ntumba badibanga, Data analysis and data mining neural network simulation by main component analysis supervised learning cases, University of Kinshasa, 2018.

[5]. Fabien Moutardeartificial learning in machine learning, School of Mines to the Paris Tech, mai 2011.

[6]. Jean Philippe, machine learning with matlab, Villaréal, 2014.

[7] Jean pepe M. Buanga, Simon Ntumba B and Richard Kabamba I, Enhanced parallel skyline on multi-Core Architecture with low memory space cost, IJCSI, vol 13, Issue 5, september 2016

[8]. Jean Christopher Routeir, Expert System and prolog: Order 0 and Order 1, Universuty of Lille 1, 2018

[9]. Peter Habermehl, Artificiel Intelligence master first Data processing year University of Paris, 2005-2006.

[10]. Serge Soudo Plaatoff, Artificial intelligence expertise everywhere accessible to all, FoundaPol Paris, Février 2018

[11]. Van Pinxteren, Y., G. Geleijnse, and P. Kamsteeg. Deriving a recipe similarity measure for recommending healthful meals. in Proceedings of the 16th international conference on Intelligent user interfaces. 2011. ACM

[12]. Witten, I.H., et al., Data Mining: Practical machine learning tools and techniques. 2016: Morgan Kaufmann.

First Author is Assistant of Mathematic and informatic department of the University of Kinshasa, DRC. Research area : Data analysis. As publications, optimization of on

line power connection a building, Avril 2014. The installation of an architecture customer waiter for management of patient case of CUK 2015. I worked in a seminar on the replication of data of the bank RAW in 2017.

Second Author is Assistant of Mathematic and informatics department of the University of Kinshasa, DRC. As publication, installation of base of date for detection and the assumption of responsability of the defavorized social layers, 2014 Research area : data base.

Third Authoris Assistant of Mathematic and informatics department of the University of Kinshasa, DRC. As publication, infrastructue of the telephonie IP migration of the telephonie analogy with tne IP memory of and of study 2014. Research area : the network.

Fourthauthor is professor and head of Mathematic and informatics department of the University of Kinshasa. As publications, Author of many publications, such as: "Enhanced Parallel Skyline on multi-core architecture with lax Memory space Cost", IJCSI, volume 13, Issue 5, September 2016, Data mart approach for stock management model with a calendar under budgetary constraint, IJCSI, volume 15, Issue 5, September 2018, Poster et the 2nd International conference on Big Data Analysis and Data Mining, San Antonio, USA, 30 November -30 December 01 2015; "Data Mart Approach for Stock Management Model with a calendar Uner Budgetary contraint" IJCSI, volume 15, Issue 5, September 201