A fuzzy inference model for cancer risk prediction

Mara Camila Vsquez Correa
Mathematics department
Universidad EAFIT
Medelln, Colombia
mvasqu49@eafit.edu.co

Abstract—In this article is presented a model of fuzzy inference in cancer risk, with three variables of support: habits, heritage and age. The model is tested in various scenarios with different methods for logical operations and defuzzification.

Index Terms—Fuzzy, Inference, FIS, cancer, risk.

I. INTRODUCTION

Cancer is the leading life-threatening disease for people in todays world. Although cancer formation is different and doctors often cannot explain why one person develops cancer and another does not, research shows that certain risk factors increase the chance that a person will develop cancer [1]. Cancer prevention through consequent screening programs, early discovery and timely, improved and diversified means of treatment are usually the most successful ways to reduce mortality.

Where uncertainty exists such as in the medical field, fuzzy logic could play an important role in making decisions. Fuzzy logic is the science of reasoning, thinking, and inference that recognizes and uses the real world phenomenon that everything is a matter of degree. In the simplest terms, fuzzy logic theory is an extension of binary theory that does not use crisp definitions and distinctions. Instead of assuming everything must be defined crisply into black and white (binary view), fuzzy logic is a method that captures and uses the concept of fuzziness in a computationally effective manner [2].

The objective of this work is to design and implement a fuzzy inference model and to analyze its results in predicting the risk of cancer in several cases. We suppose then that three variables are somehow enough to describe a possibility of developing cancer and we test the model in different scenarios.

The rest of the paper is organized as follows: first, there is a review of the current work regarding the topic. Second, there is an explanation of the methodology that is going to be used. Then the results of the model are presented and finally some conclusions are derived from the work.

II. LITERATURE REVIEW

Many authors have addressed the problem of an early diagnose of cancer using fuzzy models that incorporate characteristics of the patients. And although these models share a common core, they are very different and approach the problem from different perspectives.

Most of the existing models try to model the risk of developing cancer in a specific organ. Such models include: breast cancer ([3], [4], [5], [6], [7], [8], [9], [10], [11], [12]), prostate cancer ([13], [14]), oral cancer ([15]), gastric cancer ([16]), colon cancer ([17]) and lung cancer ([18]). However, there exists some other models that try to asses the risk of developing any kind of cancer, such as [19] and [20].

They vary also in the type of fuzzy model that is being used. Most of them are a fuzzy inference model of type Mandani ([3], [20], [17], [5], [6], [7], [16], [12]) or Sugeno ([19], [9]). Though neural fuzzy models ([13], [15], [14], [18]), fuzzy cognitive maps ([4], [8], [10]) and fuzzy probabilistic models ([11]) have been used in this field.

The way of modelling the characteristics of the patients is very similar in some models. They include hereditary characteristics, like the genetic history; lifestyle characteristics, such as the nutrition, the smoking habits, the body mass index and physical activity; biological factors like hormones, pre-existing tumors and cell conditions. Finally, they consider socio-demographic factors, such as age and gender.

In this work we aim to use three of the main groups of characteristics, socio-demographic, lifestyle and hereditary, to propose a Mandani fuzzy inference model for predicting the risk of developing any type of cancer.

III. METHODOLOGY

The model used to carry out this work is described in detail below:

A. Variables

- u_1: Quality of the habits (lifestyle).
- u_2: Quality of the genetic heritage. A good genetic heritage implies rare or non history of cancer.
- u_3: Age of the person.
- y: Risk of developing cancer.

B. Speech Universe

- $u_1 \in \mathbb{Z} \cap [0, 10]$: Being 0 the worst habits and 10 the best.
- $u_2 \in \mathbb{Z} \cap [0, 10]$: Being 0 the worst heritage and 10 the best.
- $u_3 \in \mathbb{Z} \cap [0, 90]$.
- $y \in \mathbb{R} \cap [0, 100]$.

C. Linguistic categories

- u_1: Poor, medium and high quality.
- u_2: Poor, medium and high quality.
- u_3: Young, medium and old.
- y: High, medium and low.

D. Membership functions

In the Figure 1 are shown the respective membership functions for each one of the linguistic categories. They are defined as:

- u_1:
    - *Poor*: $trapezoid(x; 0, 0, 2, 4)$
    - *Medium*: $gauss(x; 5, 1)$
    - *High*: $trapezoid(x; 6, 8, 10, 10)$
- u_2:
    - *Poor*: $sigmoid(x; 3, -2.64)$
    - *Medium*: $gauss(x; 5, 1)$
    - *Low*: $sigmoid(x; 7, 2.64)$
- u_3:
    - *Poor*: $sigmoid(x; 30, -0.53)$
    - *Medium*: $gauss(x; 50, 15.3\bar{3})$
    - *Low*: $sigmoid(x; 70, 0.264)$
- y:
    - *Poor*: $sigmoid(x; 30, -0.264)$
    - *Medium*: $gauss(x; 50, 10)$
    - *Low*: $sigmoid(x; 70, 0.264)$

E. Rule base

Habits Poor and [Heritage Poor] and [Age Low] then [Risk Medium]

Habits Poor and [Heritage Poor] and [Age Medium] then [Risk Medium]

Habits Poor and [Heritage Poor] and [Age High] then [Risk High]

Habits Poor and [Heritage Medium] and [Age Low] then [Risk Low]

Habits Poor and [Heritage Medium] and [Age Medium] then [Risk Medium]

Habits Poor and [Heritage Medium] and [Age High] then [Risk High]

Habits Poor and [Heritage High] and [Age Low] then [Risk Low]

Habits Poor and [Heritage High] and [Age Medium] then [Risk Low]

Habits Poor and [Heritage High] and [Age High] then [Risk Medium]

Habits Medium and [Heritage Poor] and [Age Low] then [Risk Medium]

Habits Medium and [Heritage Poor] and [Age Medium] then [Risk High]

Habits Medium and [Heritage Poor] and [Age High] then [Risk High]

Habits Medium and [Heritage Medium] and [Age Low] then [Risk Low]

Habits Medium and [Heritage Medium] and [Age Medium] then [Risk Medium]

Habits Medium and [Heritage Medium] and [Age High] then [Risk High]

Habits Medium and [Heritage High] and [Age Low] then [Risk Low]

Habits Medium and [Heritage High] and [Age Medium] then [Risk Low]

Habits Medium and [Heritage High] and [Age High] then [Risk Medium]

Habits High and [Heritage Poor] and [Age Low] then [Risk Medium]

Habits High and [Heritage Poor] and [Age Medium] then [Risk Medium]

Habits High and [Heritage Poor] and [Age High] then [Risk Medium]

Habits High and [Heritage Medium] and [Age Low] then [Risk Low]

Habits High and [Heritage Medium] and [Age Medium] then [Risk Low]

Habits High and [Heritage Medium] and [Age High] then [Risk Medium]

Habits High and [Heritage High] and [Age Low] then [Risk Low]

Habits High and [Heritage High] and [Age Medium] then [Risk Low]

Habits High and [Heritage High] and [Age High] then [Risk Low]

IV. RESULTS

V. CONCLUSIONS

REFERENCES

REFERENCES

[1] D. Coggon and C. Martyn, "Time and chance: the stochastic nature of disease causation," *The Lancet*, vol. 365, no. 9468, pp. 1434 – 1437, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140673605663805

[2] L. A. Zadeh, "Fuzzy sets and information granularity," *Advances in Fuzzy Systems-Applications and Theory*, vol. 26, no. 4, pp. 551–556, 1996.

[3] V. Balanic, I. Dumitrache, M. Caramihai, W. Rae, and C. Herbst, "Evaluation of breast cancer risk by using fuzzy logic," *UPB Scientific Bulletin, Series C: Electrical Engineering*, vol. 73, no. 1, pp. 53–64, 2011.

[4] A. Büyükavcu, Y. E. Albayrak, and N. Göker, "A fuzzy information-based approach for breast cancer risk factors assessment," *Applied Soft Computing Journal*, vol. 38, pp. 437–452, 2016.

[5] B. M. Gayathri and C. P. Sumathi, "Mamdani fuzzy inference system for breast cancer risk detection," *2015 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2015*, 2016.

[6] R. Khezri, R. Hosseini, and M. Mazinani, "A fuzzy rule-based expert system for the prognosis of the risk of development of the breast cancer," *International Journal of Engineering, Transactions A: Basics*, vol. 27, no. 10, pp. 1557–1564, 2014.

[7] K. Latha, "Visualization of Risk in Breast Cancer Using Fuzzy Logic in MATLAB Environment," *International Journal of Computational Intelligence Techniques*, vol. 4, no. 1, pp. 114–117, 2013.

[8] E. I. Papageorgiou, J. Subramanian, A. Karmegam, and N. Papandrianos, "A risk management model for familial breast cancer: A new application using Fuzzy Cognitive Map method," *Computer Methods and Programs in Biomedicine*, vol. 122, no. 2, pp. 123–135, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.cmpb.2015.07.003

[9] A. A. Shleeg and I. M. Ellabib, "Comparison of Mamdani and Sugeno Fuzzy Interference Systems for the Breast Cancer Risk," *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, vol. 7, no. 10, pp. 1343–1347, 2013.

[10] J. Subramanian, A. Karmegam, E. Papageorgiou, N. Papandrianos, and A. Vasukie, "An integrated breast cancer risk assessment and management model based on fuzzy cognitive maps," *Computer Methods and Programs in Biomedicine*, vol. 118, no. 3, pp. 280–297, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.cmpb.2015.01.001

[11] F. Tatari, M. R. Akbarzadeh-T, and A. Sabahi, "Fuzzy-probabilistic multi agent system for breast cancer risk assessment and insurance premium assignment," *Journal of Biomedical Informatics*, vol. 45, no. 6, pp. 1021–1034, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.jbi.2012.05.004

[12] A. Yilmaz and K. Ayan, "Risk analysis in breast cancer disease by using fuzzy logic and effects of stress level on cancer risk," *Scientific Research and Essays*, vol. 6, no. 24, pp. 5179–5191, 2011.

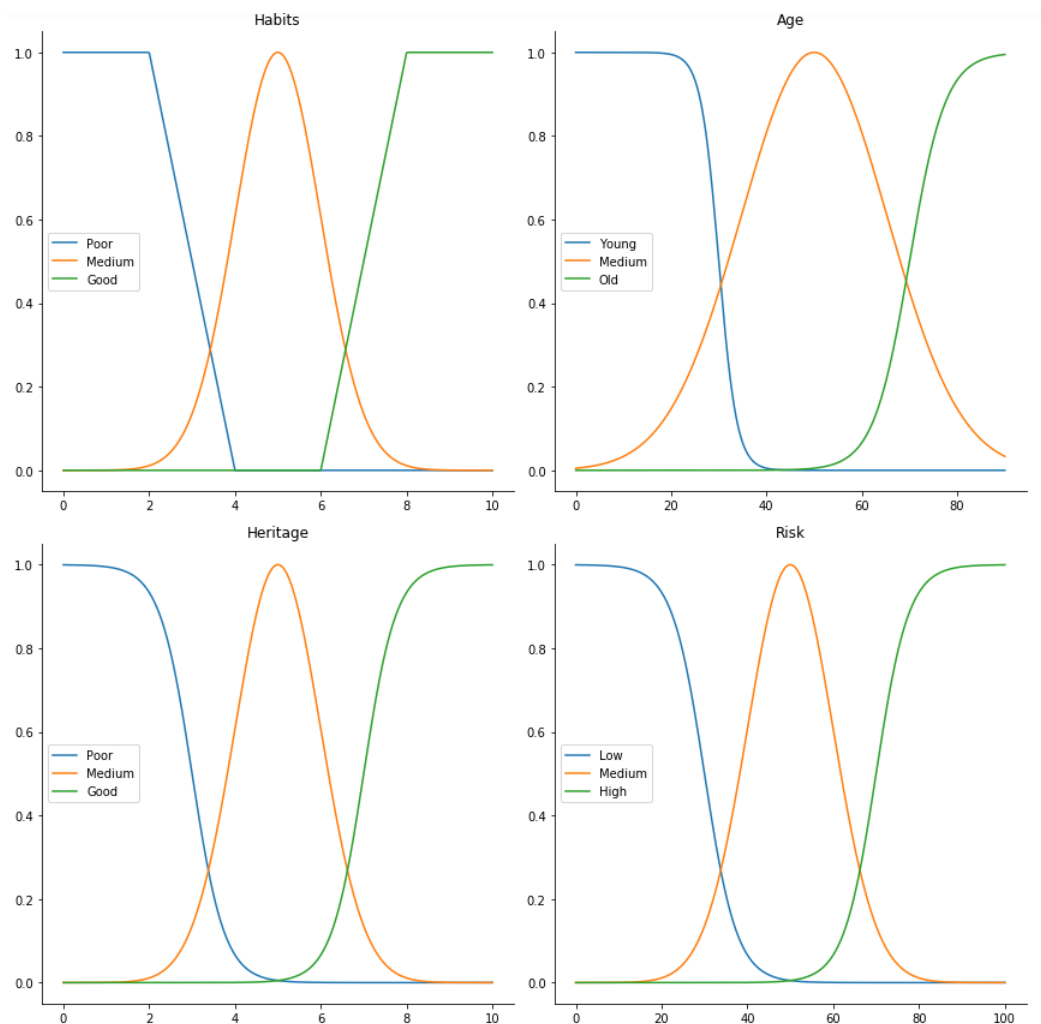[13] L. Benecchi, "Neuro-fuzzy system for prostate cancer diagnosis," *Urology*, vol. 68, no. 2, pp. 357–361, 2006.

Fig. 1. Membership functions

[14] G. Cosma, G. Acampora, D. Brown, R. C. Rees, M. Khan, and A. G. Pockley, "Prediction of pathological stage in patients with prostate cancer: A neuro-fuzzy model," *PLoS ONE*, vol. 11, no. 6, pp. 1–27, 2016.

[15] R. M. Dom, B. Abidin, S. A. Kareem, S. M. Ismail, and N. M. Daud, "Determining the critical success factors of oral cancer susceptibility prediction in Malaysia using fuzzy models," *Sains Malaysiana*, vol. 41, no. 5, pp. 633–640, 2012.

[16] R. Safdari, H. K. Arpanahi, M. Langarizadeh, M. Ghazisaiedi, H. Dargahi, and K. Zendehdel, "Design a fuzzy rule-based expert system to aid earlier diagnosis of gastric cancer," *Acta Informatica Medica*, vol. 26, no. 1, pp. 19–23, 2018.

[17] R. M. Brand, D. D. Jones, H. T. Lynch, R. E. Brand, P. Watson, R. Ashwathnayaran, and H. K. Roy, "Risk colon cancer in hereditary non-polyposis colorectal cancer patients as predicted by fuzzy modeling: Influence of smoking," *World Journal of Gastroenterology*, vol. 12, no. 28, pp. 4485–4491, 2006.

[18] A. Ylmaz, S. Ar, and Ü. Kocabçak, "Risk analysis of lung cancer and effects of stress level on cancer risk through neuro-fuzzy model," *Computer Methods and Programs in Biomedicine*, vol. 137, pp. 35–46, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.cmpb.2016.09.002

[19] Y. Atnccedil and A. Kuuml rat, "Comparison with Sugeno Model and Measurement Of Cancer Risk Analysis By New Fuzzy Logic Approach," *African Journal of Biotechnology*, vol. 11, no. 4, pp. 979–991, 2012.

[20] G. Dudek, A. Strzelewicz, M. Krasowska, A. Rybak, and R. Turczyn, "Fuzzy analysis of the cancer risk factor," *Acta Physica Polonica B*, vol. 43, no. 5, pp. 947–959, 2012.

[21] L. A. Zadeh, "Probability measures of fuzzy events," pp. 421–427, 1968.

[22] L. A. Zadeh, "Fuzzy sets," p. 4, 1965.

[23] O. L. Quintero, *Machine Intelligence for Human Decision Making*. Universidad EAFIT, November 2018.

[24] E. Kentel and M. M. Aral, "Probabilistic-fuzzy health risk modeling," *Stochastic Environmental Research and Risk Assessment*, vol. 18, no. 5, pp. 324–338, 2004.