The CSI Journal on
Computer Science and Engineering
Vol. 19, No. 2, 2025
Pages 88-92
doi.org/10.22034/jcse.2025.182044
Regular Paper

Design of Fuzzy Cognitive Map for Breast Cancer

Sahar Sarikhani

Abstract

The prediction of cancer is very important and can help reduce this disease consequences. This study intended to designing fuzzy cognitive map for breast cancer risk factors and provide an efficient result for the predicting of breast cancer. In this study, fuzzy cognitive map is designed to analysis breast cancer based on real clinical data. The data were used in this study were collected from Mahdieh hospital in Tehran, Iran. In this study we collected two groups of datasets, one group related to blood test and the other one including familial and personal risk factors.

In addition to achieve benefit information about the effect of risk factors on breast cancer in Iranian women, risk factors of breast cancer were determined and investigated based on the native dataset. For collection of datasets, we tried to consider all of risk factors. An FCM[1] can be used in decision making problems and predicting the future.

Keywords: Design, Prediction, Breast cancer, Fuzzy Cognitive Map.

1. Introduction

Cancer is a disease in which cells are proliferated abnormally [1]. Cancer is the most preventable disease and early detection and treatment of cancer, can help saving lives and reducing this disease consequences. Whatever the time interval between diagnosis and initiation of treatment be more, the situation will be exacerbated. Therefore, early detection can increase a person's chance of survival. The prediction and prevention of cancer can reduce deaths [2] from cancer and it reduces the economic costs of treatment significantly[3].

Based on what was mentioned, in this research risk factors of the breast cancer were determined and measured to designing FCM. The concepts of FCM are the most important and effective risk factors and markers of breast cancer and include some clinical trials and some familial and personal risk factors, such as parity, duration of breastfeeding, age, family history, personal history in past years, age at menarche and age at menopause. The concepts of FCM have been described qualitatively by experts and have been converted into fuzzy sets with corresponding membership functions.

Literature review

Tyrer et al. [4] predicted breast cancer by model that is incorporating familial and personal risk factors. Some of woman's risk of breast cancer are genetic factors and related to family history, others are based on personal factors. Some of personal factors are reproductive history and medical history. The discovery of the BRCA1 and BRCA2 genes has verified some of the genetic specifications of breast cancer risk, but these genes alone do not explain all of the familial risk factors of breast cancer. In fact, for predicting breast cancer this model has been tried to incorporate the BRCA genes and other low penetrance gene and personal risk factors. For an individual woman her family history is used in conjunction with Bayes theorem to iteratively calculate the likelihood of her carrying any genes predisposing to breast cancer, which in turn affects her likelihood of breast cancer be inherited and developing breast cancer. After this step, the resulting model is updated based on personal history. For each woman, the risk of breast cancer between ages t1 and t2 can be calculated using the following is given by [4]:

$$\Pr(\text{cancer}) = 1 - \left(1 - \sum_{i=1}^{6} p_i F_i(t_1, t_2)\right)^{\alpha} \qquad (1)$$

Where pi is the probability of the woman having the relevant phenotype,

---

[1] Fuzzy Cognitive Map

$F_i(t_1, t_2)$ Is the probability of getting breast cancer between ages t1 and t2 given the woman's phenotype I and α is the relative risk due to personal factors.

The paper by Pharaoh et al. [5] describes a study in which using of identified six breast-cancer susceptibility alleles that are effective in breast cancer's prediction. In fact this is polygenic approach to the prevention of breast cancer. In this study, risk factors related (see [6]) to individual loci are less important than genetic risk factors and extensive research for detection them been done. In fact, with the discovery of genes in the chromosomes of cancer risk would allow accurate predictions of disease and facilitate prevention through measures directed toward persons at high risk.

The paper by Marilena Iorio et al. [7] for prediction of breast cancer used miRNAs gene expression. They described there are differences between RNAs in normal and cancer tissue. MicroRNAs (miRNAs) are a class of small noncoding RNAs[2] that control gene expression. So miRNA expression could clearly separate normal versus cancer tissues, with the most significantly deregulated miRNAs being mir-125b, mir-145 mir-21, and mir-155. Also in this paper for separation and prediction they used SVM[3] and PAM[4] [8] algorithm.

Sang-Hyun Hwang et al. [9] proposed a new methodology for medical diagnosis based on fuzzy clustering and parallel neural networks. This method composed of two parts. One part is classifying breast cancer data using fuzzy c-means clustering method (FCM). The other is designing the multiple neural networks using classified data by FCM.

Floyd et al. [10] Predicted breast cancer malignancy using an ANN[5]. An ANN can be trained to predict malignancy from mammographic findings with a high. Really this method can be replaced biopsy process.

Michael Phillips et al. [11] predicted breast cancer using volatile biomarkers in the breath. In this paper they used a fuzzy logic model that predicted breast cancer with more accuracy than previously reported findings. The fuzzy membership function of it defined as follows:

$$T_{\text{predictor}} = T_{\text{pos}} - T_{\text{neg}} \tag{2}$$

2. Dataset for prediction of breast cancer

The dataset of this research was collected from Mahdieh hospital in Tehran, Iran and for collecting the data we interviewed with patients and copied from their medicine document such as mammography, blood test. The concepts of FCM are fields of dataset which determined by consultation with the physicians. The collection of data was started since September 2014 to May 2015. Some of the fields are risk factors of breast cancer such as parity, duration of breastfeeding, age, family history, personal history in past years, age at menarche, age at menopause, the value of CA-125 [12]. The number of patients who responded to our questions and consented to say their information to us are about 122 people. Independent variables in dataset of this research are in table 1. Two group of risk factors are seen in figure1,2.

Table 1. Fields of Dataset and related concepts in FCM.

Concepts	Description of Variable
C1	Malignant variable is 1 for malignant tumor or that is 0 for benign tumor.
C2	The age that its amount is determined for each person according to his or her age.
C3	Menopause variable has ranges 1 for reached menopause and 0 for not menopause.
C4	Breast-feeding Variable has the range of values 0 to no history of breastfeeding and 1 for breastfeeding history.
C5	The duration of breastfeeding Variable according to the number of months of breastfeeding is determined.
C6	Date of the first appearance of the tumor variable is expressed based on the number of years from advent of mass.
C7	Lose weight Variable has range of values from 0 for did not lose weight to 1 for loose weight.
C8	Anorexia variable be set 0 for has anorexia and that is 1 for has not anorexia.
C9	Painful has the range of values from 0 for not painful mass to 1 for painful mass.
C10	Mobile variable has range of values from 0 for the fixed masses and 1 for the mobile mass.
C11	Solid or Cystic value has range of values from 0 for the Cystic masses and 1 for the Solid masses.
C12	Variable diagnosed of mass by the patient or doctor has range of values from 0 to the diagnosis by the patient and 1 for diagnosis by a doctor.

---

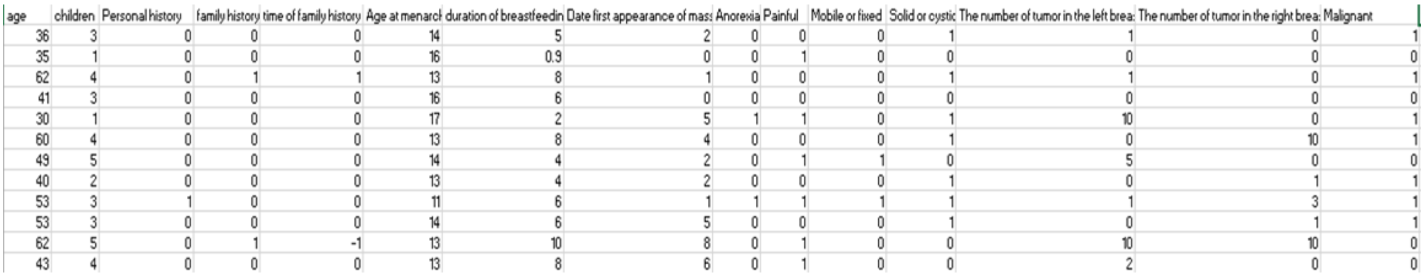
[2] Ribonucleic acid

[3] Support Vector Machine

[4] Prediction Analysis of Microarrays software

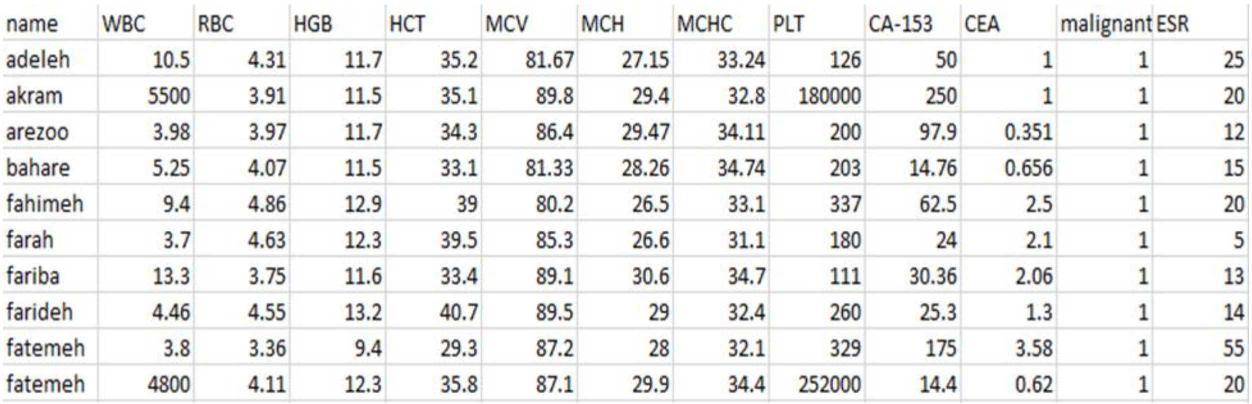
[5] Artificial Neural Network

C13	The number of masses in the left breast variable are set according to the number of masses existed in breast
C14	Metastasis variable is valued depending on the type of tumor 1 for metastatic tumor and 0 for non-metastatic.
C15	Age at first menarche variable is determined according to the age of each person at menarche.
C16	Family history is 1 for presence of family history and 0 for absence of family history.
C17	Personal history in the past years that the range of values for lack of personal history is 0 and 1 for exist of personal history.
C18	The number of children that the range of values starts from 0.
C19	Marital status that the range of values for married is 1, and 0 for is not married.



age	children	Personal history	family history	time of family history	Age at menarch	duration of breastfeedin	Date first appearance of mas:	Anorexia	Painful	Mobile or fixed	Solid or cystic	The number of tumor in the left brea:	The number of tumor in the right brea:	Malignant
36	3	0	0	0	14	5	2	0	0	0	1	1	0	1
35	1	0	0	0	16	0.9	0	0	1	0	0	0	0	0
62	4	0	1	1	13	8	1	0	0	0	1	1	0	1
41	3	0	0	0	16	6	0	0	0	0	0	0	0	0
30	1	0	0	0	17	2	5	1	1	0	1	10	0	1
60	4	0	0	0	13	8	4	0	0	0	1	0	10	1
49	5	0	0	0	14	4	2	0	1	1	0	5	0	0
40	2	0	0	0	13	4	2	0	0	0	1	0	1	1
53	3	1	0	0	11	6	1	1	1	1	1	1	3	1
53	3	0	0	0	14	6	5	0	0	0	1	0	1	1
62	5	0	1	-1	13	10	8	0	1	0	0	10	10	0
43	4	0	0	0	13	8	6	0	1	0	0	2	0	0

Figure 1. The part of dataset which related to familial and personal risk factors.



name	WBC	RBC	HGB	HCT	MCV	MCH	MCHC	PLT	CA-153	CEA	malignant	ESR
adeleh	10.5	4.31	11.7	35.2	81.67	27.15	33.24	126	50	1	1	25
akram	5500	3.91	11.5	35.1	89.8	29.4	32.8	180000	250	1	1	20
arezoo	3.98	3.97	11.7	34.3	86.4	29.47	34.11	200	97.9	0.351	1	12
bahare	5.25	4.07	11.5	33.1	81.33	28.26	34.74	203	14.76	0.656	1	15
fahimeh	9.4	4.86	12.9	39	80.2	26.5	33.1	337	62.5	2.5	1	20
farah	3.7	4.63	12.3	39.5	85.3	26.6	31.1	180	24	2.1	1	5
fariba	13.3	3.75	11.6	33.4	89.1	30.6	34.7	111	30.36	2.06	1	13
farideh	4.46	4.55	13.2	40.7	89.5	29	32.4	260	25.3	1.3	1	14
fatemeh	3.8	3.36	9.4	29.3	87.2	28	32.1	329	175	3.58	1	55
fatemeh	4800	4.11	12.3	35.8	87.1	29.9	34.4	252000	14.4	0.62	1	20

Figure 2. The part of dataset which related to blood test.

3. Methodology

In this paper have considered some cases, such as a history of breast cancer tumors in individual, family history, age, age at menopause, age at first birth and the size and shape of the tumor, whether fixed or moving mass is solid or cystic mass, single or multiple the masses and CA-125 [13]. Then proceeded to collection data on patients suspected or diagnosed with breast cancer. The fuzzy cognitive map was designed based on data and advice of expert.

4. Design Fuzzy Cognitive Map

To design a fuzzy cognitive map for breast cancer, a basic map based on the expert advice was implemented as figure 3.

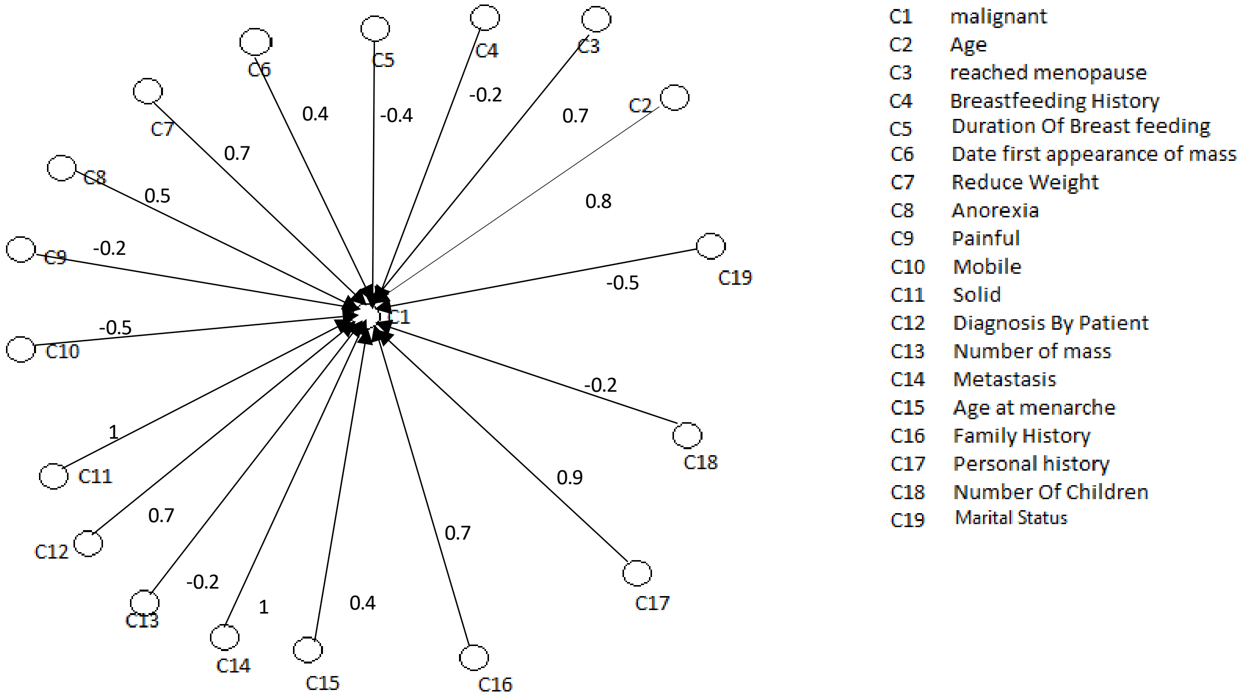


Figure 3. Initial FCM of components for to predicting breast cancer.

The final FCM was depicted based on expert advice, it was seen as figure 4.

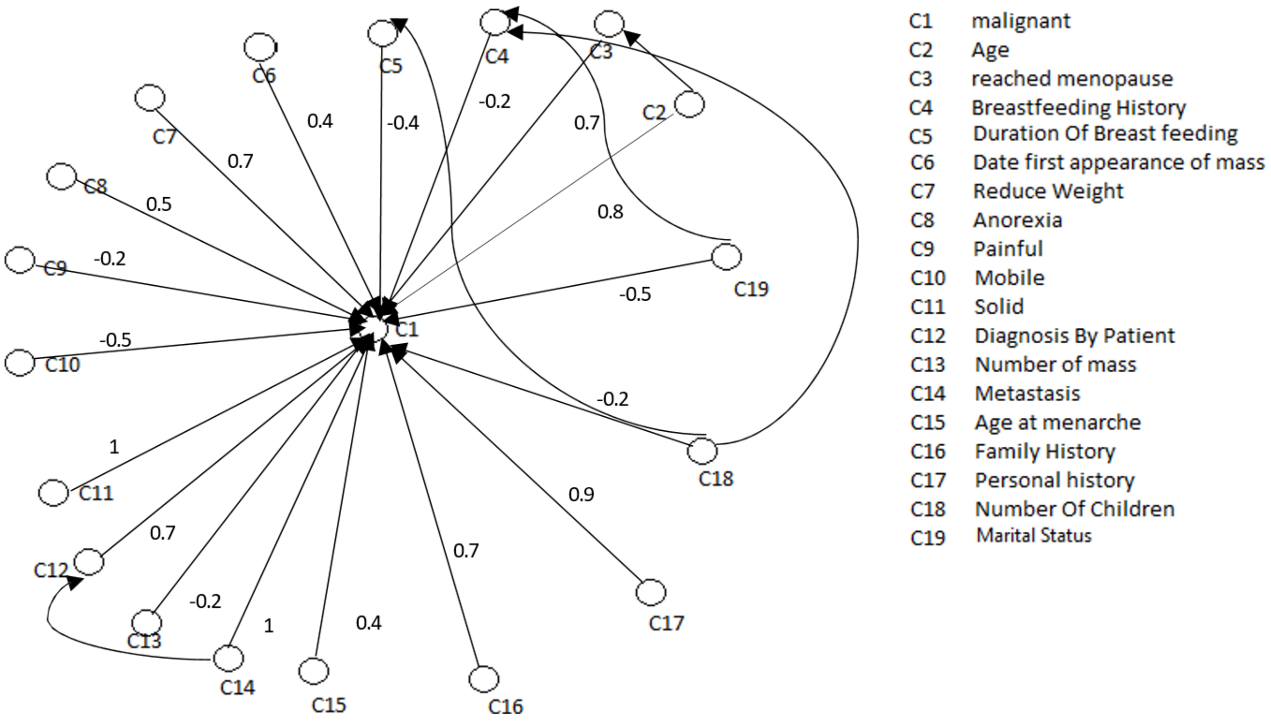


Figure 4. Final FCM of components to predicting breast cancer.

Also the FCM related to blood test concepts for prediction breast cancer was showed as figure 5. This model was constructed similar to previous method of constructing FCM.

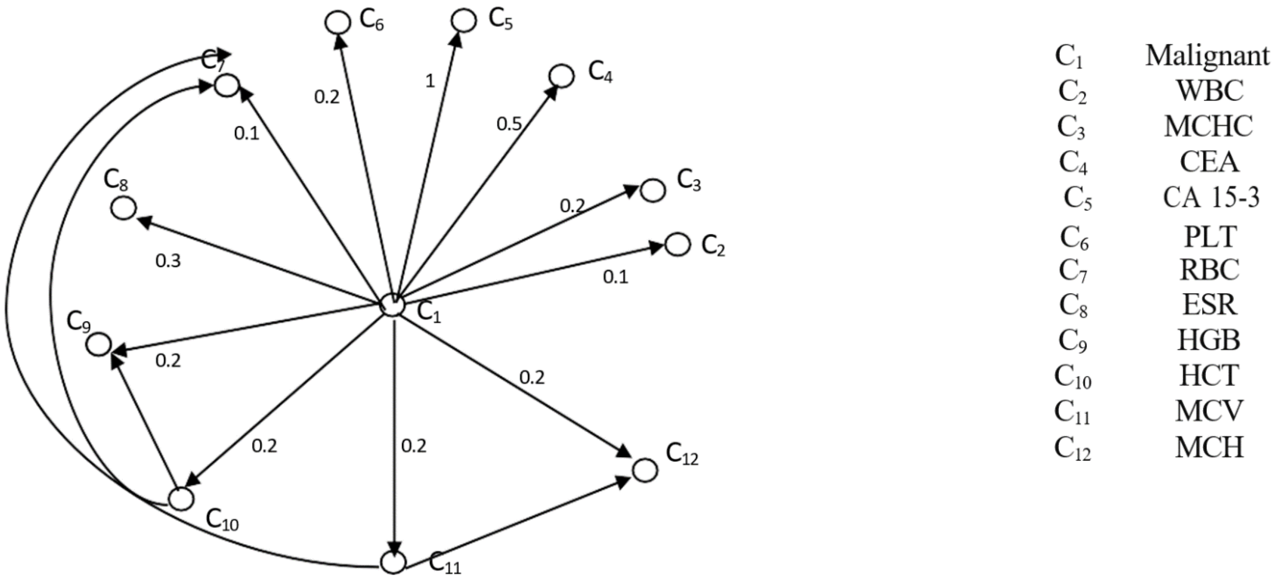


Figure 5. FCM for blood test to predicting breast cancer.

In fact proposed model is learned from data of patients which we collected them. At last fuzzy cognitive map was drawn based on these results.

5. Conclusions

In fact, one of the goals of this study was to evaluate the effect of each risk factor in breast cancer. Therefore, the main purpose of this study is recognizing the most important risk factors of breast cancer. Also, the dataset was collected in Iran and that is native dataset of breast cancer.

In conclusion, it is our belief these results have the capacity to provide an efficient document for the predicting of breast cancer.

Acknowledgement

The authors would like to thank patients who assisted us to constructing dataset and also, we gratefully acknowledge the Tehran's Mahdieh hospital.

References

[1] William, M.A., et al., Symptoms of Breast Cancer– An Analysis Using Induced Fuzzy Cognitive Maps (IFCMS). Internationl Journal of Computing Algorithm, 2014. **3**: p. 659- 663.

[2] Palmisano, W.A., et al., Predicting lung cancer by detecting aberrant promoter methylation in sputum. Cancer research, 2000. **60**(21): p. 5954-5958.

[3] Belda-Iniesta, C., et al., New screening method for lung cancer by detecting volatile organic compounds in breath. Clinical and Translational Oncology, 2007. **9**(6): p. 364-368.

[4] Tyrer, J., S.W. Duffy, and J. Cuzick, A breast cancer prediction model incorporating familial and personal risk factors. Statistics in medicine, 2004. **23**(7): p. 1111-1130.

[5] Pharoah, P.D., et al., Polygenes, risk prediction, and targeted prevention of breast cancer. New England Journal of Medicine, 2008. **358**(26): p. 2796-2803.

[6] Gail, M.H., et al., Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. Journal of the National Cancer Institute, 1989. **81**(24): p. 1879-1886.

[7] Iorio, M.V., et al., MicroRNA gene expression deregulation in human breast cancer. Cancer research, 2005. **65**(16): p. 7065-7070.

[8] Tibshirani, R., et al., Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proceedings of the National Academy of Sciences, 2002. **99**(10): p. 6567-6572.

[9] Hwang, S.-H., et al., Medical diagnosis system of breast cancer using FCM based parallel neural networks, in Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence. 2007, Springer. p. 712-719.

[10] Floyd, C.E., et al., Prediction of breast cancer malignancy using an artificial neural network. Cancer, 1994. **74**(11): p. 2944-2948.

[11] Phillips, M., et al., Prediction of breast cancer using volatile biomarkers in the breath. Breast cancer research and treatment, 2006. **99**(1): p. 19-21.

[12] Helmrich, S.P., et al., Risk factors for breast cancer. American journal of epidemiology, 1983. **117**(1): p. 35-45.

[13] Safi, F., et al., The value of the tumor marker CA 15-3 in diagnosing and monitoring breast cancer. A comparative study with carcinoembryonic antigen. Cancer, 1991. **68**(3): p. 574- 582.