

# Supervised Algorithms of Machine Learning for the Prediction of Cervical Cancer

Asadi F.<sup>1\*</sup>, Salehnasab C.<sup>2</sup>, Ajori L.<sup>3</sup>

## ABSTRACT

**Background:** Compared to other genital cancers, cervical cancer is the most prevalent and the main cause of mortality in females in third-world countries, affected by different factors, including smoking, poor nutritional status, immune-deficiency, long-term use of contraceptives and so on.

**Objective:** The present study was conducted to predict cervical cancer and identify its important predictors using machine learning classification algorithms.

**Material and Methods:** In a cross-sectional study, the data of 145 patients with 23 attributes, which referred to Shohada Hospital Tehran, Iran during 2017–2018, were analyzed by machine learning classification algorithms which included SVM, QUEST, C&R tree, MLP and RBF. The criteria measurement used to evaluate these algorithms included accuracy, sensitivity, specificity and area under the curve (AUC).

**Results:** The accuracy, sensitivity, specificity and AUC of Quest and C&R tree were, respectively 95.55, 90.48, 100, and 95.20, 95.55, 90.48, 100, and 95.20, those of RBF 95.45, 90.00, 100 and 91.50, those of SVM 93.33, 90.48, 95.83 and 95.80 and those of MLP 90.90, 90.00, 91.67 and 91.50 percentage. The important predictors in all the algorithms were found to comprise personal health level, marital status, social status, the dose of contraceptives used, level of education and number of caesarean deliveries.

**Conclusion:** This investigation confirmed that ML can enhance the prediction of cervical cancer. The results of this study showed that Decision Tree algorithms can be applied to identify the most relevant predictors. Moreover, it seems that improving personal health and socio-cultural level of patients can be causing cervical cancer prevention.

**Citation:** Asadi F, Salehnasab C, Ajori L. Supervised Algorithms of Machine Learning for the Prediction of Cervical Cancer. *J Biomed Phys Eng.* 2020;10(4):513-522. doi: 10.31661/jbpe.v0i0.1912-1027.

## Keywords

Cervical Cancer; Machine Learning; Prediction; Neural networks; Decision Trees; Support Vector Machine

## Introduction

Compared to other genital cancers, cervical cancer is the most prevalent and the main cause of mortality in females in third-world countries. Based on global estimates, over 57000 new cases of this cancer are annually identified, 80% of those emerge in developing countries. Moreover, 77% of deaths in women are caused by this cancer

<sup>1</sup>PhD, Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>2</sup>PhD Candidate, Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>3</sup>MD, PhD, Department of obstetrics and gynaecology, preventative gynaecology Research centre, Shahid Beheshti University of Medical Sciences Tehran, Iran

\*Corresponding author:  
F. Asadi  
Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran  
E-mail: asadifar@sbmu.ac.ir

Received: 16 December 2019  
Accepted: 5 January 2020

[1-4]. The prevalence of cervical cancer has been reported to be lower in Iran compared to in some other countries. According to a 2018 report by the Iran National Cancer Registry of the Ministry of Health and Medical Education, the five-year prevalence of cervical cancer in Iran has been 2613 cases in a total cancer cases of 248392 in all age groups, and its ranking the 22<sup>nd</sup> compared to all types of cancer in both genders [5]. Research suggests that human papillomavirus (HPV) significantly contributes to developing cervical cancer [6] and that infection with this virus can cause cervical cancer over a 10-15-year period [7]. Given its prolonged pre-invasive period, accessibility of the infected organ for sampling and the opportunity to administer Pap smear, this cancer appears preventable and diagnosable in early stages [8]. Moreover, the cytological factors in Pap-smear that are considered as prognostic risk factors for cervical cancer include the shape of gland cells, squamous epithelial tissue, the presence of metaplastic cells, abnormal polymorphic cells and dysplasia cells, different epithelial shapes and the presence of blood, bacteria and fungi in the patients sample [9]. Research suggests that merely 5% of women in developing countries participate in Pap smear screening programs [10] and mainly use surgery or radiotherapy to treat this cancer, which exerts different harmful effects on women's reproductive organs [11-13].

Many factors are associated with cervical cancer, including smoking (of the person or their spouse), poor nutrition status, immunodeficiency, using immunosuppressive medications, long-term use of contraceptives [14], age, race [15], deficiencies of vitamins A and C and folic acid [16], a history of several marriages (having several sex partners), successive pregnancies, childbirth at young ages, certain sexually-transmitted genital infections, a poor socioeconomic status, inhaling the smoke of burning wood and coal and low education levels [17, 18]. A collection of these variables and risk factors are required to be concurrently

evaluated in order to predict the probability of developing cervical cancer faster and more accurately. Using non-invasive methods such as supervised machine learning (ML), classification algorithms are crucial for predicting cervical cancer. These models include artificial neural networks [19-22], decision trees [9, 23-26] and support vector machine (SVM) [9, 23, 27-30]. Neural networks are highly-complex analytical techniques that predict new observations from other observations after running the so-called process of "learning" from available data [31]. The most popular neural network-based algorithms, which are used as powerful estimating functions in prediction problems, include multi-layer perceptron artificial neural network (MLP-ANN) and radial basis function (RBF-ANNs) [32]. Decision trees maximize the accuracy of prediction results using a tree structure and recursively putting data into branches according to predetermined criteria [33]. A decision tree is a tree structure such as a flowchart in which each internal node represents the test of a feature or attribute, each branch the result of the test and leaf nodes the classes or class distributions [34]. Compared to other machine learning classification algorithms, the rules inferred from decision tree algorithms can be properly and easily interpreted [35]. SVM is also a popular machine learning algorithm differentiating between different results by designing data points in a multidimensional linear or nonlinear space and plotting a super-plan separator [36, 37]. According to the discussed points, the present study was conducted to predict cervical cancer using the cited algorithms.

## Material and Methods

In a cross-sectional study, the methodology consisted of six phases, as follows:

- 1- An applied cross-sectional study used library resources and the latest studies to determine the important variables and risk factors affecting cervical cancer. The most popular predictive machine learning models used for

the subject was also identified.

2- A researcher-made questionnaire was designed, and its validity was confirmed using content validity based on a review of the literature and expert opinions about the study subject, and its reliability confirmed by calculating a Cronbach's alpha of 0.87.

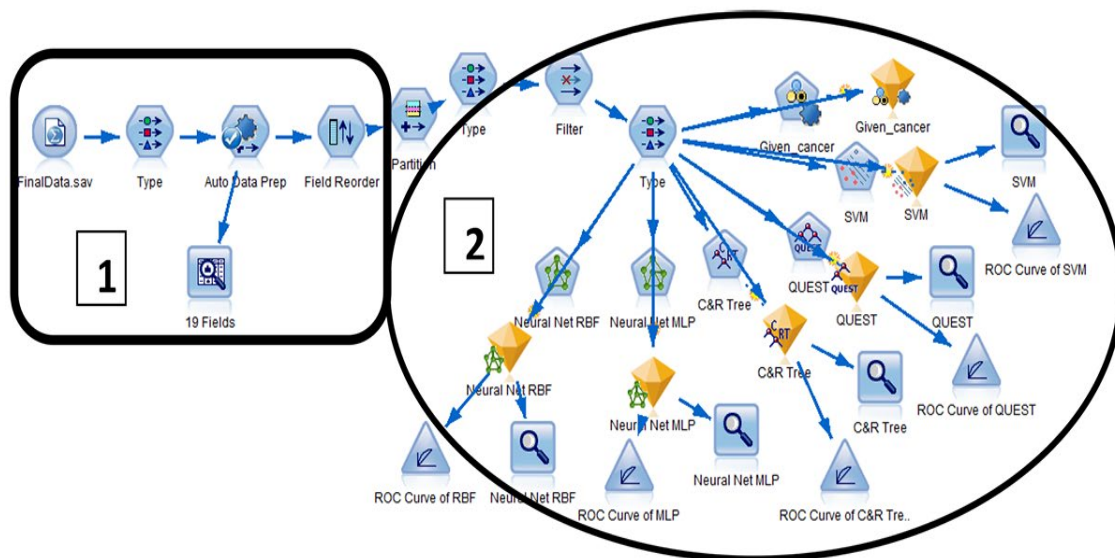
3- The authors presented to a teaching hospital affiliated to Shahid Beheshti University of Medical Sciences to obtain the necessary permissions and investigate the data available in the patients' medical records.

4- The data of all the patients presenting in 2017-18 were collected in a cross-sectional retrospective manner by reviewing their medical records and interviewing them after obtaining their informed consent. A total of 145 out of 219 patients receiving treatment were selected after excluding the incomplete records.

5- The data collected were pre-processed to prepare them for modelling (Figure 1-part 1). Given the dependency of the results obtained from machine learning classification algorithms on the quality of raw data, pre-processing is essential for improving the data. There-

fore, the variables with over 50% loss of data were eliminated from the study, and the other data attribute lost replacement by the mean for the continuous data, with the mode for the nominal data and the median for the ordinal data. The continuous data were also normalized.

6- After conducting the pre-processing, the data were modelled using machine learning algorithms (Figure 1, part 2). A total of 70% of the pre-processed data randomly underwent the training process and 30% were tested. After the classification, modelling was performed in two stages using SVM, QUEST, C&R Tree, MLP-ANNs and RBF-ANNs. In the first stage, all the study variables were included in the algorithms. In the next stage, the entrance variables, the independent predictors were identified by at least two models based on AUC, accuracy, sensitivity, and specificity of the models in phase 1, the expert's opinion, and clinical findings, were used in the selected models as input. The most appropriate models and the most significant predictors for predicting cervical cancer were ultimately identi-



**Figure 1:** Mechanism of pre-processing the data and developing machine learning classification algorithms

fied through modelling and re-evaluating the models in IBM SPSS Modeler 18. The models were evaluated using sensitivity, specificity, area under the ROC curve and accuracy.

## Results

Table 1 presents the most significant predictors for predicting cervical cancer obtained by reviewing library resources and recent studies.

Therefore, twenty-two variables were measured for each patient. The numerical value of the target variable, i.e. developing cervical cancer, was either one or zero.

The mean age of the patients was 47 years, 54% were married, 41.4% were illiterate, 38.6% had a high school diploma and the rest had higher education levels. Social status was poor in 46.2% and moderate in 39.3%. Economic status was poor in 33.1% and moderate in 66.2%. Moreover, personal health level was poor in 42.1% and moderate in 37.2%. In addition, 1.4% had a family history of cervical cancer, 54.5% a history of using contraceptives and 53.7% a history of early pregnancy (younger than 21 years). A total of 25% had more than four children, 97.2% had no history of smoking, 96.6% no history of alcohol consumption and 17.2% had a history of immunodeficiency. Moreover, 10.3% had HPV, 100% had no HSV2, 48.3% had one sex partner, 56.6% had a marriage age below 21 years, 99.3% had no history of chlamydia, 98.6% no history of sexually-transmitted diseases, 40.7% had a history of chronic diseases, including diabetes and hypertension, and 44.1% were ultimately found to have developed cervical cancer. The three variables of the duration of alcohol consumption and the presence of HIV and HSV2 were excluded during pre-processing, and modelling was performed using the cited algorithms.

After performing the modelling in the first stage, nine variables were excluded based on the two principles cited (Table 2), i.e. none of them was presented as a predictor in the two models.

**Table 1:** Important variables obtained from library studies.

Row	Variable	Type	Role
1	Age	Continuous	Input
2	Marital status	Nominal	Input
3	Education level	Nominal	Input
4	Social status	Nominal	Input
5	Economic status	Nominal	Input
6	Personal health level	Nominal	Input
7	Family history of cervical cancer	Nominal	Input
8	The dose of contraceptives used	Continuous	Input
9	Age at the first childbirth	Continuous	Input
10	Number of childbirths by caesarean	Nominal	Input
11	Number of pregnancies	Continuous	Input
12	Period of smoking consumption	Continuous	Input
13	Period of alcohol consumption	Continuous	Input
14	Immunodeficiency	Nominal	Input
15	HPV	Nominal	Input
16	*HSV2	Nominal	Input
17	Number of sex partners	Nominal	Input
18	Marriage Age	Continuous	Input
19	*HIV	Nominal	Input
20	Chlamydia	Nominal	Input
21	Number of sexually-transmitted diseases	Nominal	Input
22	History of chronic diseases	Nominal	Input
23	Given/Not Given cervical cancer	Flag	Target

\* Excluded in the pre-processing stage

**Table 2:** Predictors excluded in the first stage

Predictor	SVM	C&R Tree	QUEST	RBF	MLP	Occurrence
Number of sexually-transmitted diseases					√	1
Number of sex partners	√					1
Marriage Age				√		1
HPV					√	1
History of chronic diseases	√					1
Economic status	√					1
Family history of cervical cancer						0
Duration of smoking						0
Chlamydia						0

The second stage of modelling was carried out with the remaining variables, and the evaluative indicators were separately calculated for each model (Table 3). Based on the evaluative criteria of accuracy, sensitivity, specificity and area under the ROC curve, decision trees, decision trees and support vector machine, decision trees and RBF and support vector machine algorithms respectively performed the best.

Evaluating the ROC curve (Figure 2) and the area under (Table 3) for the algorithms run in the second stage of modelling found the highest area under the ROC curve to be associated with the support vector machine for the test data, whereas all the algorithms except for RBF neural network performed the same for

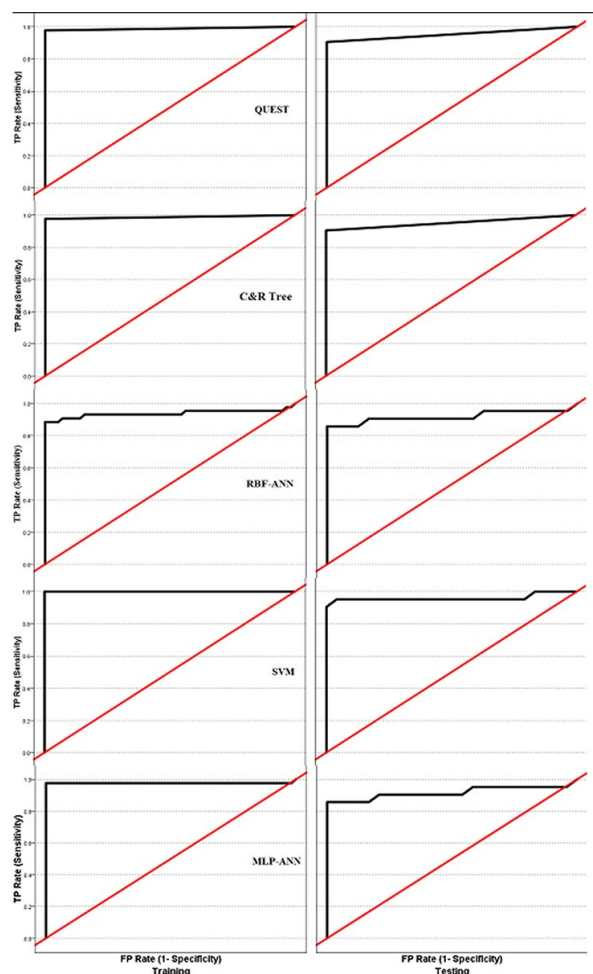
the training data.

Table 4 shows the most important predictors as the final predictors of the present study. Although ten variables were confirmed in the second stage of modelling, personal health level, marital status, social status, dose of contraceptives used, education level and the number of caesarean deliveries were ultimately considered essential in all the algorithms, and the decision tree algorithms rejected age, age at the first pregnancy, number of pregnancies and immunodeficiency.

Personal health levels were the most important predictor, and other variables were equally crucial for the development of cervical cancer in all the algorithms except for RBF neural network (Figure 3) in which the dose of con-

**Table 3:** Evaluating the algorithms in the second modelling stage arranged by the accuracy of the test data

Row	ML algorithms	%Accuracy	%Sensitivity	%Specificity	%AUC
1	QUEST Tree	95.55	90.48	100.00	95.20
2	C&R Tree	95.55	90.48	100.00	95.20
3	RBF-ANNs	95.45	90.00	100.00	91.50
4	SVM	93.33	90.48	95.83	95.80
5	MLP-ANNs	90.90	90.00	91.67	91.50



**Figure 2:** The ROC curve for classification algorithms

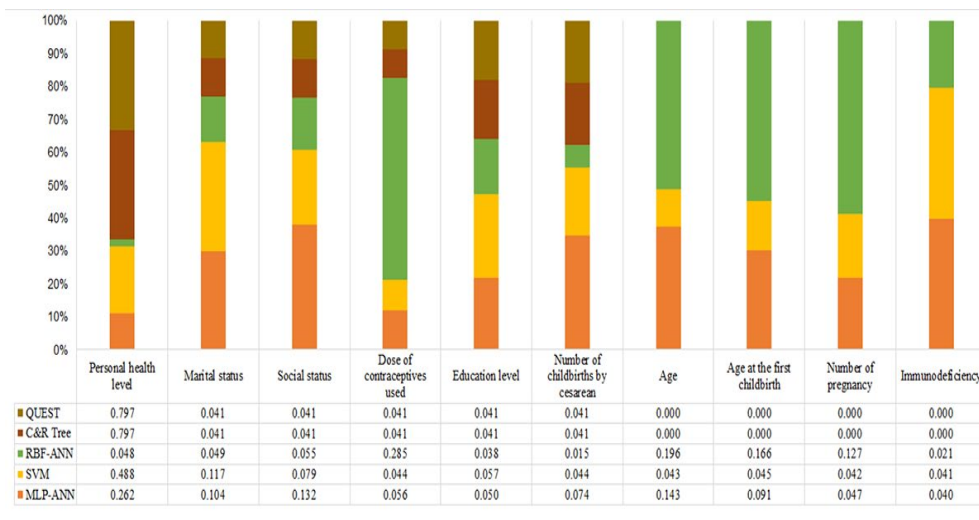
traceptives was the most important predictor.

### Discussion

The present study developed a model for predicting the probability of developing cervical cancer. Diagnosing this disease in the early stages is crucial, as it does not exhibit specific early symptoms. The majority of women seek medication at advanced stages of this cancer, which further complicate the treatment and impose a huge financial and psychological burden on the patient [38]. Therefore, the present study investigated the important predictors and the most popular algorithms for predicting cervical cancer. Excluding the variables of the duration of alcohol consumption, infections with HIV and HSV2 in the pre-processing stage showed that variables with minor changes to the patients' samples cannot be considered effective predictors. Given the potential importance of a variable for the community or the patient's social status, fewer predictors may be required to be considered in future studies; for instance, 96.6% of the subjects having no history of alcohol consumption is natural in Iran. On average, over 97% of the study subjects had no history of smoking or cervical cancer in their family or chlamydia. Given the discussed point, these three variables were

**Table 4:** Significant predictors in the second stage of modelling

Predictor	SVM	C&R Tree	QUEST	RBF	MLP	Occurrence
Personal health level	√	√	√	√	√	5
Marital status	√	√	√	√	√	5
Social status	√	√	√	√	√	5
Dose of contraceptives used	√	√	√	√	√	5
Education level	√	√	√	√	√	5
Number of childbirths by caesarean	√	√	√	√	√	5
Age	√			√	√	3
Age at the first childbirth	√			√	√	3
Number of pregnancies	√			√	√	3
Immunodeficiency	√			√	√	3



**Figure 3:** The importance of the important predictors in the algorithms

normally disregarded in all the algorithms in the first stage. In contrast, infection with HIV, history of smoking and HPV were essential and influential variables in other studies [22, 24]. The present study found personal health levels, marital status, social status, the dose of contraceptives used and education level to be respectively the most important predictors, which is inconsistent with the results of other studies [22, 24]. Therefore, the sociocultural context of a community can play a critical role in obtaining patient data. Moreover, certain risk factors for cervical cancer, including the history of smoking and alcohol consumption, reveal their effect over time.

Five algorithms were examined after preparing the data. The results obtained in terms of an evaluative index of over 90% confirmed all the algorithms. A study by Vidya et al., [39] divided the data with five features into 500 training data and 100 test data, and compared to other algorithms, found the best performance to be associated with MLP with an accuracy of 98%, sensitivity of 98% and area under the ROC curve of 99%. With a slightly weaker performance, SMO was identified as the best algorithm followed by the J48 deci-

sion tree. The SMO algorithm had an 88% accuracy, 91% sensitivity and an 89% area under the ROC curve, while the J48 algorithm had a 58% accuracy, 62% sensitivity and 65% area under the curve [39]. The present study found the best performance to be related to the decision tree and the poorest to MLP-ANNs. In [39], MLP-ANN and SVM had the best results in terms of all the indices and the area under the ROC curve compared to the present study. This discrepancy of results can be explained using larger sample size, i.e. 500 training and 100 test data. Although the MLP algorithm was reported to be the best in studies by Hemalatha and Usha Rani [40] with an 85.5% accuracy, a 78.94% sensitivity and a 60.72% precision, and by Kusy et al., [19] with 107 samples, and a 72% accuracy, a 69% sensitivity, a 74% specificity and a 67% area under the ROC curve, it poorly performed in the present study (with a larger sample size). In the study by Kusy et al., [19] the RBF neural network algorithm showed a poorer performance with a 55% accuracy, a 42% sensitivity, a 67% specificity and a 48% area under the ROC curve compared to in the present study. In a survey by Kurniawati et al., [9] the SVM algorithm

with a 79% accuracy, a 67% precision and an 85% area under the ROC curve also showed poorer performance compared to in the present study. The present study found Decision Tree algorithms to have the best performance.

## Conclusion

This investigation confirmed that ML can enhance the prediction of cervical cancer. The results of this study showed that Decision Tree algorithms can be applied to identify the most relevant predictors. The proposed models reduce the computational cost as the number of important predictors for analysis reduced. With the aid of machine learning, the disease can be predicted with greater accuracy. Moreover, it seems that improving personal health and socio-cultural level of patients can be causing cervical cancer prevention.

## Acknowledgment

The authors would like to thank the reviewers for their valuable comments, which helped to improve our paper.

## Conflict of Interest

None

## References

1. Jaime G, Morales-Vásquez F, Meneses-Garcia A, editors. Cervical cancer. Springer; 2017. doi: 10.1007/978-3-319-45231-9.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *A Cancer Journal for Clinicians*. 2018;**68**(6):394-424. doi: 10.3322/caac.21492.
3. De Martel C, Plummer M, Vignat J, Franceschi S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer*. 2017;**141**(4):664-70. doi: 10.1002/ijc.30716.
4. Momenimovahed Z, Salehiniya H. Cervical cancer in Iran: integrative insights of epidemiological analysis. *BioMedicine*. 2018;**8**(3):18. doi: 10.1051/bmdcn/2018080318. PubMed PMID: 30141405. PubMed PMCID: PMC6108227.
5. Globocan W. Estimated cancer incidence and

mortality and prevalence worldwide in 2018. Globocan; 2018 [cited 11 March 2019]. Available from: <http://gco.iarc.fr/today/data/factsheets/populations/364-iran-islamic-republic-of-factsheets.pdf>.

6. Burchell AN, Winer RL, De Sanjosé S, Franco EL. Epidemiology and transmission dynamics of genital HPV infection. *Vaccine*. 2006;**24**:S52-61. doi: 10.1016/j.vaccine.2006.05.031.
7. Aminisani N, Armstrong BK, Canfell K. Cervical cancer screening in Middle Eastern and Asian migrants to Australia: a record linkage study. *Cancer epidemiology*. 2012;**36**(6):e394-400. doi: 10.1016/j.canep.2012.08.009.
8. Mohaghegh F, Ahmadlou M. A Study of the Prevalence of Cervical Cancer among Married Women in Arak, 2013. *Journal of Arak university of Medical Sciences*. 2015;**18**(4):65-70.
9. Kurniawati YE, Permanasari AE, Fauziati S. Comparative study on data mining classification methods for cervical cancer prediction using pap smear results. 1st International Conference on Biomedical Engineering (IBIOMED); Yogyakarta, Indonesia: IEEE; 2016. p. 1-5. doi: 10.1109/IBIOMED.2016.7869827.
10. Do HH, Taylor VM, Yasui Y, Jackson JC, Tu SP. Cervical cancer screening among Chinese immigrants in Seattle, Washington. *Journal of immigrant health*. 2001;**3**(1):15-21. doi: 10.1023/A:1026606401164.
11. Ramondetta L. What is the appropriate approach to treating women with incurable cervical cancer?. *Journal of the National Comprehensive Cancer Network*. 2013;**11**(3):348-55.
12. Esmati E, Kalaghchi B. Uterine cervix carcinoma: pathologic characteristics, treatment and follow-up evaluation. *TUMJ*. 2008;**65**(11):55-9.
13. Le Borgne G, Mercier M, Woronoff AS, Guizard AV, Abeillard E, Caravati-Jouvenaux A, et al. Quality of life in long-term cervical cancer survivors: a population-based study. *Gynecologic oncology*. 2013;**129**(1):222-8. doi: 10.1016/j.ygyno.2012.12.033.
14. Latha DS, Lakshmi PV, Fathima S. Staging Prediction in Cervical Cancer Patients—A Machine Learning Approach. *International Journal of Innovative Research and Practices*. 2014;**2**(2):14-23.
15. Mandelblatt J, Andrews H, Kerner J, Zauber A, Burnett W. Determinants of late stage diagnosis of breast and cervical cancer: the impact of



- age, race, social class, and hospital type. *Am J Public Health*. 1991;**81**(5):646-9. doi: 10.2105/ajph.81.5.646. PubMed PMID: 2014871. PubMed PMID: PMC1405079.
16. Workowski KA, Bolan GA. Sexually Transmitted Diseases Treatment Guidelines, 2015. Morbidity and Mortality Weekly Report (MMWR); Atlant: CDC; 2015. p. 1-137.
  17. Schiffman M, Wentzensen N. A suggested approach to simplify and improve cervical screening in the United States. *Journal of lower genital tract disease*. 2016;**20**(1):1. doi: 10.1097/LGT.000000000000170. PubMed PMID: 26704326. PubMed PMID: PMC4692178.
  18. Dietrich AJ, Tobin JN, Cassells A, Robinson CM, Greene MA, Sox CH, Beach ML, DuHamel KN, Younge RG. Telephone care management to improve cancer screening among low-income women: a randomized, controlled trial. *Annals of internal medicine*. 2006;**144**(8):563-71. doi: 10.7326/0003-4819-144-8-200604180-00006.
  19. Kusy M, Obrzut B, Kluska J. Application of gene expression programming and neural networks to predict adverse events of radical hysterectomy in cervical cancer patients. *Medical & biological engineering & computing*. 2013;**51**(12):1357-65. doi: 10.1007/s11517-013-1108-8.
  20. Sokouti B, Haghypour S, Tabrizi AD. A framework for diagnosing cervical cancer disease based on feedforward MLP neural network and ThinPrep histopathological cell image features. *Neural Computing and Applications*. 2014;**24**(1):221-32. doi: 10.1007/s00521-012-1220-y.
  21. Qiu X, Tao N, Tan Y, Wu X. Constructing of the risk classification model of cervical cancer by artificial neural network. *Expert Systems with Applications*. 2007;**32**(4):1094-9. doi: 10.1016/j.eswa.2006.02.020.
  22. Benazir B, Nagarajan A. An Expert System for Predicting the Cervical Cancer using Data Mining Techniques. *International Journal of Pure and Applied Mathematics*. 2018;**118**(20):1971-87.
  23. Tseng CJ, Lu CJ, Chang CC, Chen GD. Application of machine learning to predict the recurrence-proneness for cervical cancer. *Neural Computing and Applications*. 2014;**24**(6):1311-6. doi: 10.1007/s00521-013-1359-1.
  24. Fatlawi HK. Enhanced classification model for cervical cancer dataset based on cost sensitive classifier. *International Journal of Computer Techniques*. 2017;**4**(4):115-20.
  25. Chang CC, Cheng SL, Lu CJ, Liao KH. Prediction of recurrence in patients with cervical cancer using mars and classification. *International Journal of Machine Learning and Computing*. 2013;**3**(1):75.
  26. Alam TM, Khan MM, Iqbal MA, Abdul W, Mush-taq M. Cervical cancer prediction through different screening methods using data mining. *International Journal of Advanced Computer Science and Applications*. 2019;**10**(2):388-96. doi: 10.14569/IJACSA.2019.0100251.
  27. Zhang J, Liu Y. Cervical cancer detection using SVM based feature screening. International Conference on Medical Image Computing and Computer-Assisted Intervention; Berlin, Heidelberg: Springer; 2004. p. 873-880.
  28. Mukhopadhyay S, Kurmi I, Dey R, Das NK, Pradhan S, Pradhan A, Ghosh N, et al. Optical diagnosis of colon and cervical cancer by support vector machine. Biophotonics: Photonic Solutions for Better Health Care V, 98870U; Brussels, Belgium: International Society for Optics and Photonics; 2016. doi: 10.1117/12.2227316.
  29. Nematollahi M, Akbari R, Nikeghbalian S, Salehnasab C. Classification models to predict survival of kidney transplant recipients using two intelligent techniques of data mining and logistic regression. *International journal of organ transplantation medicine*. 2017;**8**(2):119-122. PubMed PMID: 28959387. PubMed PMID: PMC5611541.
  30. Rezaianzadeh A, Dastoorpoor M, Sanaei M, Salehnasab C, Mohammadi MJ, Mousavizadeh A. Predictors of length of stay in the coronary care unit in patient with acute coronary syndrome based on data mining methods. *Clinical Epidemiology and Global Health*. 2019. doi: 10.1016/j.cegh.2019.09.007.
  31. Haykin S. Neural Networks: A comprehensive Foundation. New Jersey: Prentice Hall PTR Upper Saddle River; 1999;7458:161-75.
  32. West D. Neural network credit scoring models. *Computers & Operations Research*. 2000;**27**(11-12):1131-52. doi: 10.1016/S0305-0548(99)00149-5.
  33. Quinlan JR. Induction of decision trees. *Machine learning*. 1986;**1**(1):81-106. doi: 10.1007/BF00116251.
  34. Han J, Pei J, Kamber M. Data mining: concepts and techniques. Elsevier; 2011.
  35. Kim JW, Lee BH, Shaw MJ, Chang HL, Nelson M.

- Application of decision-tree induction techniques to personalized advertisements on internet storefronts. *International Journal of Electronic Commerce*. 2001;5(3):45-62. doi: 10.1080/10864415.2001.11044215.
36. Burges CJ. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*. 1998;2(2):121-67. doi: 10.1023/A:1009715923555.
37. Burges CJ, Smola AJ, Scholkopf B. Advances in kernel methods. Support Vector Learning. London, England: MIT press; 1999.
38. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *The Lancet*. 2007;370(9590):890-907. doi: 10.1016/S0140-6736(07)61416-0.
39. Vidya R, Nasira GM. Knowledge extraction in medical data mining: a case based reasoning for gynecological cancer an expert diagnostic method. *ARPJ Journal of Engineering and Applied Sciences*. 2006;10(9):3997-4001.
40. Hemalatha K, Rani DU. Improvement of multi-layer perceptron classification on cervical pap smear data with feature extraction. *International Journal of Innovative Research in Science, Engineering, and Technology*. 2016:20419-24.