Prostate Cancer Probability Prediction By Machine Learning Technique

Srđan Jović, Milica Miljković, Miljan Ivanović, Milena Šaranović & Milena Arsić

Published online: 15 Dec 2017.

Submit your article to this journal ⤤

View related articles ⤤

View Crossmark data ⤤

Taylor & Francis
Taylor & Francis Group

Prostate Cancer Probability Prediction By Machine Learning Technique

Srđan Jović[a], Milica Miljković[b], Miljan Ivanović[b], Milena Šaranović[b], and Milena Arsić[a]

[a]Faculty of Technical Sciences, University of Priština, Kosovska Mitrovica, Serbia; [b]Medical Faculty, University of Priština, Kosovska Mitrovica, Serbia

ABSTRACT

The main goal of the study was to explore possibility of prostate cancer prediction by machine learning techniques. In order to improve the survival probability of the prostate cancer patients it is essential to make suitable prediction models of the prostate cancer. If one make relevant prediction of the prostate cancer it is easy to create suitable treatment based on the prediction results. Machine learning techniques are the most common techniques for the creation of the predictive models. Therefore in this study several machine techniques were applied and compared. The obtained results were analyzed and discussed. It was concluded that the machine learning techniques could be used for the relevant prediction of prostate cancer.

Introduction

The function of the prostate is to produce fluid that protects and nourishes sperm cells in semen. The seminal vesicles behind the prostate gland produce most of the fluid for semen. The prostate wraps around the urethra, the tube that carries urine and semen out of the body. Most prostate cancers are slow growing; however, there are a significant number of cases of aggressive prostate cancers that can be very devastating. The decision making of prostate cancer treatments are particularly important for the patients diagnosed with prostate cancer. There are a number of prostate cancer treatments such as radiation therapy, surgery, and combination of radiation and surgery.

In order to find an appropriate model suitable for a multistate survival experiment, 634 patients with chronic myeloid leukaemia (CML) were selected to illustrate the method of analysis in article (1) and it was concluded that it was possible for an individual patient to predict the 4 possible probabilities at any time. In paper (2) was presented experiments with the application of rough set-based data mining methodology to discover predictive rules in small cell lung cancer patient data and the obtained results have been derived from patient data supplied by cancer researchers from the Allan Blair Cancer Center, Regina, Saskatchewan, Canada who also provided all the necessary background information and conducted medical evaluation of the results. Consumer acceptability of potato chips with different moisture contents (MC) was evaluated using survival analysis (3). Cancer survival forecasting may be attempted using models constructed through predictive techniques of various kinds, including statistical multivariate regression and machine learning (4). Prognostic models are becoming useful in assessing the severity of illness and survival analysis in medical domain (5). The aim of study (6) was to build a predictive survival process that takes into account the functional and nutritional evolution of the patients over time. Pre-therapeutic prediction of outcome is important for clinicians and patients in determining whether selective internal radiation therapy (SIRT) is indicated for hepatic metastases of colorectal cancer (CRC) (7). Survival statistics are fundamental in outcome evaluation of clinical studies for modern cancer science. The use of models to predict the outcome must follow adequate procedures for reliability evaluation and testing, in order to prevent the erroneous application on unsuitable patient populations (8). In paper (9) a dipolar neural network

designed for prediction of survival time was presented and the prediction capacity of neural network models was evaluated by accuracy, sensitivity, and specificity measures, which are compared to results of logistic regression. In paper (10) was discussed the survival probability for exclusive central diffractive production of a colorless small size system. Artificial neural network are nonlinear regression computational devices that have been used for over 45 years in classification and survival prediction in several biomedical systems, including colon cancer (11). Cancer survival studies are commonly analyzed using survival-time prediction models for cancer prognosis. A number of different performance metrics are used to ascertain the concordance between the predicted risks score of each patient and the actual survival time, but these metrics can sometimes conflict. Alternatively, patients are sometimes divided into two classes according to a survival-time threshold, and binary classifiers are applied to predict each patient's class. Although this approach has several drawbacks, it does provide natural performance metrics such as positive and negative predictive values to enable unambiguous assessments (12). In paper (13) was presented an empirical comparison between several paradigms coming from Statistics and Artificial Intelligence for solving a supervised classification problem. The use of computer based learning models in medical domain has become a significant area of research (14). Artificial Neural Networks have increasingly been used for building prognostic models in several medical domains (15). The paper (16) presented a computational-intelligence-based model to predict the survival rate of critically ill patients who were admitted to an intensive care unit (ICU) and the proposed model could achieve a 96% and 86% accuracy rate for the training data and testing data, respectively.

In order to improve predictive performances of the prostate cancer in this study machine learning techniques are applied (17–22). The machine learning techniques have high capability for prediction of nonlinear data (23–25). Also there is no need to explore internal knowledge of the prostate cancer. It is enough to acquire input/output data samples in order to train the machine learning techniques.

## Methodology

Statistical data

Before the applying of the machining learning techniques prostate cancer data was acquired by measurement. The data was divided in five inputs based on the time when it was acquired and recorded. For example the first input represents the prostate cancer data during the treatment. The second input represents the prostate cancer data six months before the treatment. The third input represents the prostate cancer data twelve months before the treatment. The third input represents the prostate cancer data eighteen months before the treatment. The third input represents the prostate cancer data twenty-four months before the treatment. The outputs should be survival prediction based on the given inputs of the prostate cancer.

The data set was divided into the training set (containing 70% of the observations) and the test set (containing 30% of the observations).

Machine learning techniques

As the machine learning techniques, three methods are applied. The first method is Extreme Learning Machine (ELM) which represents a learning algorithm for single hidden layer feed forward networks. This approach has easy application, fast training time, and good generalization of nonlinear data. ELM methodology is used for single hidden layer feed forward networks and the main structure of the network is shown in Figure 1.

The second machine learning technique is classical artificial neural network (ANN) and the third machine learning technique is genetic programming (GP).
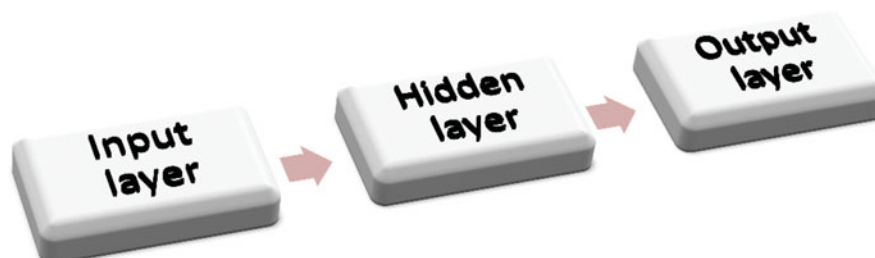
Figure 1. The topological structure of the network for ELM application.

ELM prediction of prostate cancer
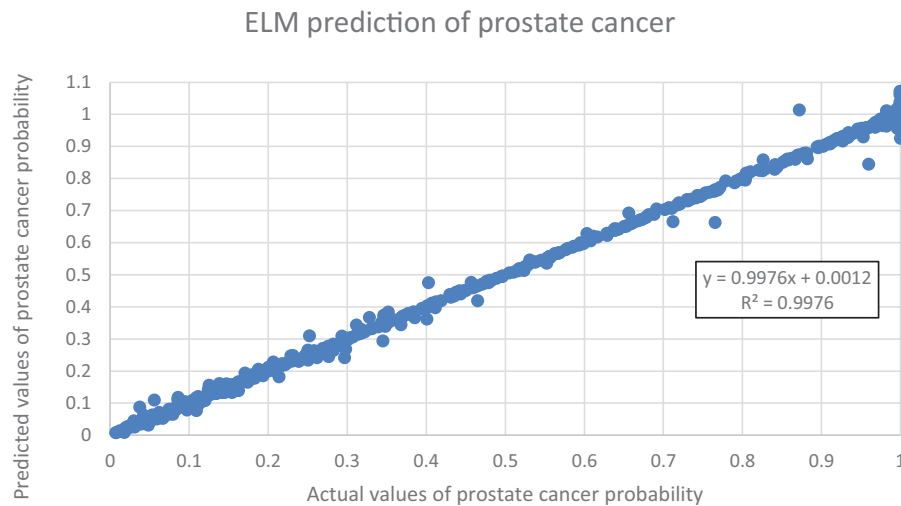
Figure 2. ELM forecasted values of probability of prostate cancer.

Predictive performances of the proposed models were analyzed by the root means square error (RMSE) and coefficient of determination (R^2).

Results and discussion

Before application of machine learning techniques, the optimal parameters of the techniques should be found. The optimal parameters are found based on trial and error procedure.

The network for ELM application has three layers with 5 inputs and 1 output. The hidden year has 6 neurons. The main learning rule of the network is extreme learning approach. The number of iteration is not defined strictly since it depends on the results. Also there is no need to define activation function for such type of the network.

The classical network ANN has also three layers with 5 inputs and 1 output. The hidden layer has 6 neurons. The main learning rule of the network is back propagation approach. The number of iteration is 1000 of the ANN network. The activation function of the ANN network is sigmoid function in this study.

GP methodology has 600 samples for population size and there is 5–9 head sized. The numbers of chromosomes are 20–30, the number of genes are 2–3, the mutation rate is 92, the crossover rate is 31, and the inversion rate is 109.

Prediction results

In this section, performance results of the ELM predictive model are reported for survival probability forecasting of prostate cancer patients. Figure 2

ANN prediction of prostate cancer
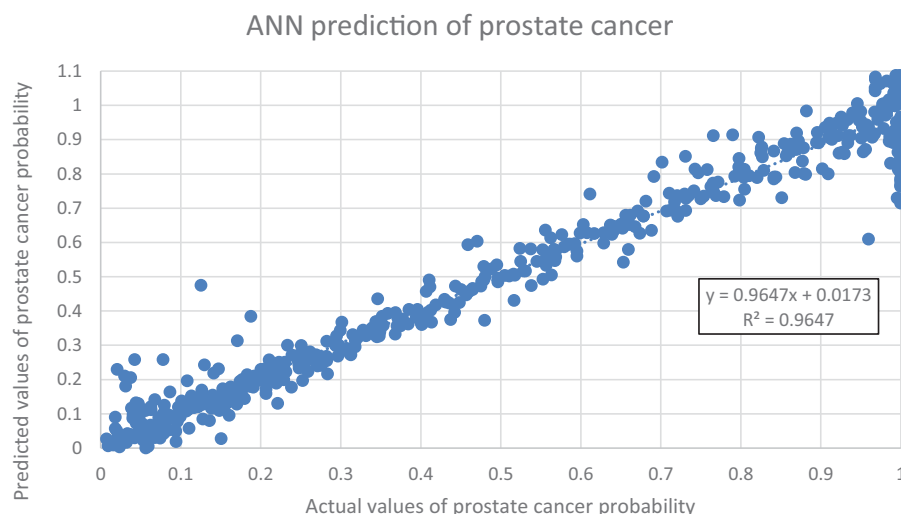
Figure 3. ANN forecasted values of probability of prostate cancer.

GP prediction of prostate cancer
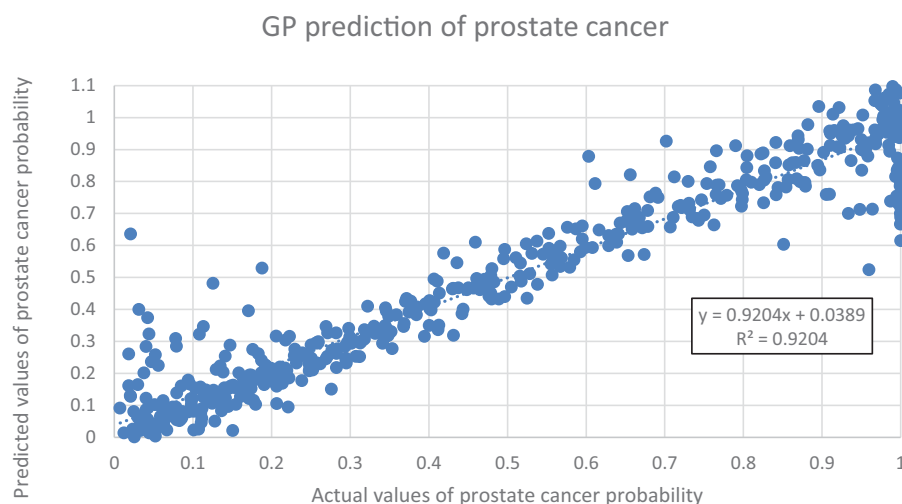
Figure 4. GP forecasted values of probability of prostate cancer.

presents the accuracy of the developed ELM predictive model for the survival probability. Figure 3 presents the accuracy of the developed ANN (21) predictive model for the survival probability. Figure 4 presents the accuracy of the developed GP (22) predictive model for the survival probability. The prediction accuracy of the ELM model is acceptable for this data. It can be seen that most of the points fall along the diagonal line. It follows that the prediction results are in very good agreement with the measured values for the ELM method. This observation can be confirmed with acceptable value for the coefficient of determination. The number of overestimated or underestimated values is limited. Consequently, it is obvious that the predicted values enjoy high level of precision.

Here one can see the prediction results of prostate cancer probability by the machine learning techniques (Figure 2–4). Based on the scatter plots one can see that the network with ELM methodology has the highest prediction accuracy. The coefficient of determination is very high for the ELM approach. On the other hand the GP prediction accuracy has the lowest coefficient of determination.

ELM approach has RMSE: 0.0167, ANN has RMSE: 0.0642 and GP has RMSE: 0.0964. Therefore based on the RMSE values also one can conclude that the ELM outperforms the ANN and GP methodologies.

Conclusion

Since the prediction of probability of prostate cancer could be of very high interest in this study was applied machine learning approach in order to find the best

technique for the prediction. Three methods were used for the same data and corresponding diagrams were reported and analyzed. Based on the results, coefficients of determination for ELM, ANN, and GP methods were 0.9976, 0.9647, and 0.9204, respectively. Root mean square errors for ELM, ANN, and GP were: 0.0167, 0.0642, and 0.0964, respectively. The simulation results revealed that ELM model is able to predict the prostate cancer probability favorably so that it provides the most accurate predictions.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

References

1. Fang Y, Hein P. Probability prediction in multistate survival models for patients with chronic myeloid leukaemia. J Huazhong Univ Sci Technolog Med Sci. February 2005;25(1):100–103. doi:10.1007/BF02831400.
2. Aryeetey K, Ziarko Tai, Ago C. Identification of Rules for Brain Metastases and Survival Time Prediction for Small Cell Lung Cancer Patients. Intelligent Information Systems 2001. Volume 10 of the series Advances in Intelligent and Soft Computing p. 1–12.
3. Jeong Y, Jang Y-A, Chang YH, Lee Y. Prediction of consumer acceptability of potato chips with various moisture contents using survival analysis. Food SciBiotechnol. December 2013;22(6):1621–1627. doi:10.1007/s10068-013-0259-6.
4. Floyd S, Ruiz C, Sergio AA, Tseng J, Whalen G. Prediction of Pancreatic Cancer Survival through Automated Selection of Predictive Models. Biomedical Engineering Systems

and Technologies Volume 127 of the series Communications in Computer and Information Science p. 29–43.

5. Raji CG, Vinod Chandra SS. Prediction and Survival Analysis of Patients After Liver Transplantation Using RBF Networks. Data Mining and Big Data, Volume 9714 of the series Lecture Notes in Computer Science p. 147–155.

6. Calle ML, Roura P, Arnau A, Yáñez A, Leiva A. Prognostic Factors and Prediction of Residual Survival for Hospitalized Elderly Patients. Advances in Statistical Methods for the Health Sciences, Part of the series Statistics for Industry and Technology p. 167–178.

7. Fendler WP, Ilhan H, Paprottka PM, Jakobs TF, Heineman V, Bartenstein Peter, Khalaf F, Ezziddin S, Hacker M, Haug A R. Nomogram including pretherapeutic parameters for prediction of survival after SIRT of hepatic metastases from colorectal cancer. Eur Radiol. September 2015;25(9):2693–2700. doi:10.1007/s00330-015-3658-7.

8. Valentini V, Damiani A, Dekker A, Dinapoli N. Statistics of Survival Prediction and Nomogram Development, Decision Tools for Radiation Oncology. Part of the series Medical Radiology. 7–28.

9. Krętowska M, Bobrowski L. Evaluation of Dipolar Neural Networks in Survival Time Prediction. Neural Networks and Soft Computing, Volume 19 of the series Advances in Soft Computing p. 230–235.

10. Gotsman E, Kowalski H, Levin E, Maor U, Prygarin A. Survival probability for exclusive central diffractive production of colorless states at the LHC. The European Physical Journal C – Particles and Fields. September 2006;47(3);655–669.

11. Ahmed, F. E. (2005). Artificial neural networks for diagnosis and survival prediction in colon cancer. Molecular cancer, 4(1); 29. PMID:16083507.

12. Chen, H. C., Chen, J. J., Cheng, K. F., & Kodell, R. L. (2012). Assessment of performance of survival prediction models for cancer prognosis. BMC medical research methodology, 12(1); 102. doi:10.1186/1471-2288-12-102.

13. Larrañaga P, Gallego MJ, Sierra B, Urkola L, Michelena MJ. Bayesian networks, rule induction and logistic regression in the prediction of women survival suffering from breast cancer. Progress in Artificial Intelligence, Volume 1323 of the series Lecture Notes in Computer Science p. 303–308.

14. Raji CG, Vinod Chandra SS. Graft survival prediction in liver transplantation using artificial neural network models. J Computational Sci. September 2016;16:72–78. doi:10.1016/j.jocs.2016.05.005.

15. Sapra RL, Mehrotra S, Nundy S. Artificial Neural Networks: Prediction of mortality/survival in gastroenterology. Current Medicine Research and Practice. May–June 2015;5(3):119–129. doi:10.1016/j.cmrp.2015.05.007.

16. Hsieh YZ, Su MC, Wang CH, Wang PC. Prediction of survival of ICU patients using computational intelligence. Comput Biol Med. 1 April 2014;47;13–19. doi:10.1016/j.compbiomed.2013.12.012.

17. Huang GB, Zhu QY, Siew CK. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. International Joint Conference on Neural Networks. 2004;2;985–990.

18. Wang D, Huang GB. Protein Sequence Classification Using Extreme Learning Machine. Proceedings of International Joint Conference on Neural Networks. 2005;3; 1406–1411.

19. Huang GB, Zhu QY, Siew CK. Real-time learning capability of neural networks. Technical Report ICIS/45/2003, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. 2003.

20. Nian R, He B, Zheng B, Heeswijk MV, Yu Q, Miche Y, et al. Extreme learning machine towards dynamic model hypothesis in fish ethology research. Neurocomputing. 2014;128:273–284. doi:10.1016/j.neucom.2013.03.054.

21. Liang N-Y, Huang G-B, Saratchandran P, Sundararajan N. A fast and accurate online sequential learning algorithm for feedforward networks. Neural Networks, IEEE Transactions on. 2006;17(6):1411–1423. doi:10.1109/TNN.2006.880583. Koza J. Genetic programming: On the programming of computers by natural selection. Cambridge, MA, USA: MIT Press, 1992.

23. Mohammadhassani M, Nezamabadi-Pour H, Suhatril M, Shariati M. Identification of a suitable ANN architecture in predicting strain in tie section of concrete deep beams. Struct Eng Mech. 2013;46(6):853–868. doi:10.12989/sem.2013.46.6.853.

24. Toghroli A, Mohammadhassani M, Suhatril M, Shariati M, Ibrahim Z. Prediction of shear capacity of channel shear connectors using the ANFIS model. Steel and Composite Structures.2014;17(5):623–639. doi:10.12989/scs.2014.17.5.623.

25. Safa M, Shariati M, Ibrahim Z, Toghroli A, Baharom SB, Nor N M, Petkovic D. Potential of adaptive neuro fuzzy inference system for evaluating the factors affecting steel-concrete composite beam's shear strength. Steel and Composite Structures. 2016;21(3):679–688. doi:10.12989/scs.2016.21.3.679.