RESEARCH ARTICLE OPEN ACCESS

Descriptive Analysis of Machine Learning and Its Application in Healthcare

Sai Sruthi Gadde [1], Venkata Dinesh Reddy Kalli [2]
Software Developer [1], Research Scientist [2]
Cardiac and Vascular Group, Medtronic, Bangalore
India

ABSTRACT

The dynamic world of big data in the healthcare sector characterized by huge numbers, complexity, and speeds is also not suited to conventional research methods. Methods are especially required that can efficiently estimate models across comprehensive datasets of medical usage data, clinical data, personal computer data, and many other sources. While these data sets are quite large, they may also be very sparse (e.g., system data may only be accessible for a small subset of people), creating difficulties with conventional regression models. Most machine learning approaches successfully overcome these limitations but still are subject to the standard triggers of partiality that are typical in observatory studies. The models should be tested by standard design tests for researchers using machine learning techniques like a lasso or ridge regression.

Keywords:- ML

I. INTRODUCTION

The term "machine learning" refers to a large family of mathematical and statistical methods that have historically been focused on prediction [1]. We are also involved in forecasting healthcare. Which form of flu is likely to occur in the next influenza season? How many influenza vials are available to fulfill the care demand? Nevertheless, predictions are not necessarily the same as predicting drug outcomes. The job of a doctor is to isolate the effect of an operation on the outcome of a patient in order to select the correct drug [2]. The same methodological problems face policy assessments. Some methods of machine learning can predict therapy results, and some do not. However, in the literature for machine learning, the gap between prediction and treatment-effect estimates is almost entirely absent. In brief, in order to generate highly accurate classification algorithms, a key focus of any machine learning is to segment data into training and validation data sets. After the algorithms are built, the full data to make the prediction is applied [3].

It is no wonder that medicine is overwhelmed by groundbreaking claims from machine learning to large-scale healthcare data. Recent examples show that big data and machine learning can build algorithms equivalent to human doctors. While computer education and big data can at first look mysterious, they are, in fact, closely related to traditional statistical models that most clinicians recognize. Initially, machine learning was defined as being a system where the work or decisions are made automatically from the data instead of the actions being explicitly programmed [4]. This concept is, however, comprehensive and can cover almost any sort of data-driven approach.

II. DEFINITION OF MACHINE LEARNING

Perhaps it is easier to consider the life of an algorithm in a continuum between absolutely human-directed and fully machine-led analysis. It is crucial that you see how much of the structure or parameter of a predictive or diagnostic algorithm is said to be an example of computer education [5]. The trade-off between human characteristics of predicational algorithms against the processing of data is called the machine learning continuum.

This is what an algorithm is trying to do. Since people place fewer expectations on the algorithm, the learning range of the computer is further increased [6]. However, a model does not immediately become "machine learning," but rather, all such methods exist in a continuum, based on the number of human constraints put on the algorithm.

An example of a high-level machine-learning method in the form of so-called deep learning models has recently emerged. Deep learning models are astonishingly complex neuron networks that have been developed to construct accurate models from raw data directly [7]. Recently researchers have demonstrated an in-depth learning algorithm that can detect diabetic retinopathy with a sensitivity that is equal or greater than that of ophthalmologists. This model got the diagnosis from the raw pixels of the images without any human interference outside a team of ophthalmologists who annotated the correct diagnosis on each image [8]. Since the task is mastered with little to no human experience, these profound learning algorithms are fundamental in the Master Spectrum of Learning.

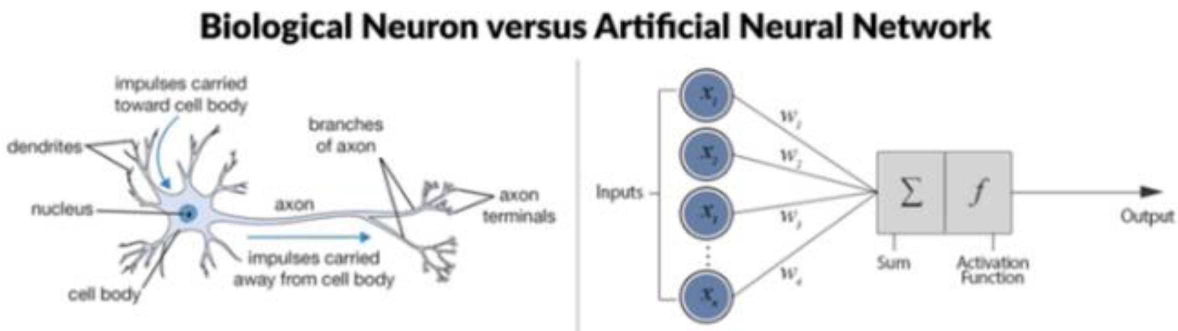


Figure 1. Biological neurons VS artificial neurons network

While less personal guidance is needed, deep learning algorithms to recognize images require enormous data amounts to capture the full complexity, variety, and nuance of real-world models. Such algorithms, therefore, often demand the extraction of the outstanding image features that are connected to the result in hundreds of thousands of instances [9]. Higher placement in the continuum of master learning does not mean superiority since various tasks need different levels of human involvement. Although spectrum algorithms are also very versatile and can learn several jobs, they often are not interpretable; there is also the velocity aspect – the speed at which users can communicate. EMR data is also almost in real-time available. In addition, data diversity is increasing. Claims and EMR data are increasingly associated with broad-based health risk assessments, socio-demographic information, and vital signs. Some algorithms are used in the optimization of antiurolithiatic activities [10]. More recently, new data on human genetic traits as well as data from devices such as Fitbit and biometric sensors are available. Such knowledge is vibrant but sparse.

That creates challenges for conventional multivariate methods like standard smaller-square regression analysis since many observations are lost due to the lack of data. For the study of empirical evidence, there are several successful statistical methods [11]. Nevertheless, the sheer quantity of data along with its features, including unequal data completeness, raises concerns about the potential for new methods of addressing issues of treatment efficacy, patient benefit, strengths of alternative care system models and limitations, policy behavior, etc.

Some approaches to machine learning use prediction methods based on regression. Lasso approaches, for instance, use a correction factor to reduce the overfit chance. Since Lasso may reduce those variables' coefficients to zero, it is useful for selecting variables. Most notably, since the regression of Lasso requires estimating coefficients in a multivariate model, the use of machine learning in determining the treatment effects is a short step forward [12]. Many researchers believe that computers do not easily select the final model specification. This can be known. The final model for the theoretical or clinical plausibility of researchers will, however, definitely be tested and subject to the standard battery of design checks. Furthermore, a set of starting variables from which the model is built will expertly handle the risk of an unforeseeable scenario [13]. Machine learning approaches make the initial variables much more substantial than standard health

care research practice, but the idea of a theoretical or clinical model must not be thrown out entirely.
Unfortunately, machine learning protects from the typical problems faced by observational data analysis is simply nothing magical [14]. In fact, it does not defend against prejudice by merely running machine learning methods on Big Data. Increased-sample size, for example, is not going to address the bias issue if the data collection lacks essential clinical seriousness indicators such as cancer stage in a breast cancer model.

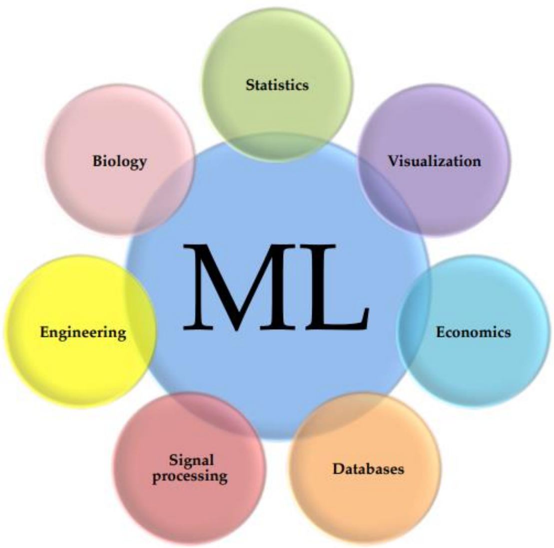


Figure 2. Machine learning is interdisciplinary

Artificial Intelligence (AI)

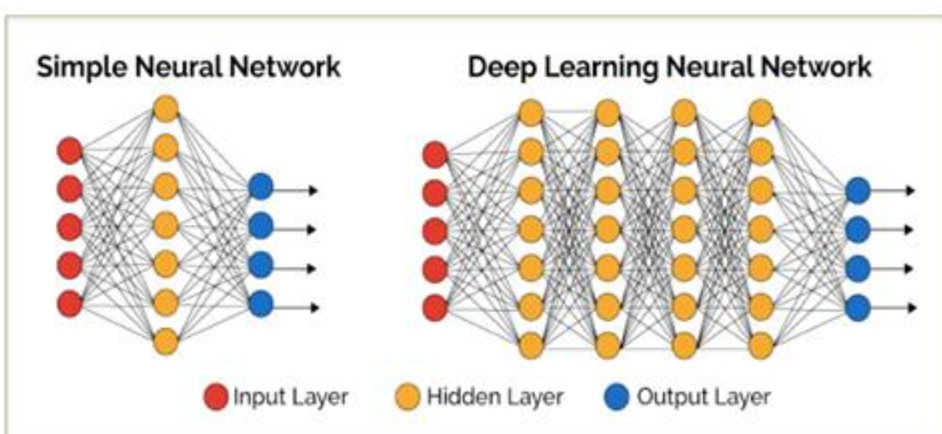
A specific science discipline focused on philosophy, mathematics, and computer science that aims at understanding and creating structures that exhibit intelligence properties [15].

Machine Learning

A sub-discipline of AI in which computer programs (algorithms) learn predictive power correlations from data examples. The implementation of mathematical models on computers is most obviously machine learning. Machine learning uses a wider variety of mathematical methods that are popular in medicine. New techniques like Deep Learning are based on models where the underlying information is less expected and thus capable of processing more complex data [16].

Deep Learning

Deep learning methods allow a computer to supply large quantities of raw data to detect or classifying the necessary representations. Detailed methods for learning are focused on multiple data layers with successive transformations that amplify input aspects of discrimination, which are essential to remove irrelevant variations. Profound schooling can be regulated or unregulated. Deep learning approaches are responsible for many of the new machine learning advancement [17].



Supervised Learning

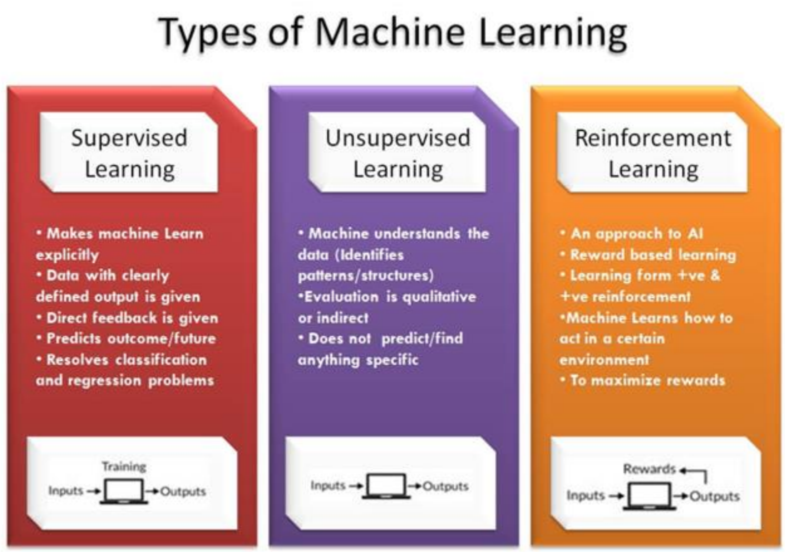
Computer training programs to learn links between data inputs and outputs through analysis of interest outputs identified by a (typically human) supervisor. After the understanding of correlations, they may predict future instances based on current evidence. This is one of the best-known fields in machine learning, with many cases in and outside of healthcare [18].

Unsupervised Learning

Computer programs which learn associations with data without external association concept. In comparison to simply building upon existing connections, unsupervised learning may classify previously unknown forecasters [19].

Reinforcement Learning

Computer programs that learn behavior by maximizing a given reward. This strategy is inspired by conduct psychology and was widely used in games where knowledge is ideal, with several potential choices and no specific worldwide fault costs [20].



III. AI AND DECISION MAKING IN HEALTH SYSTEMS

In fact, efficient health system management is a set of activities of information processing; for example, the provision of public health or health care.

Policymakers adjust organizational and governance health system structures, funding, and resource management to achieve health system efficiency and program objectives [21].

The healthcare sector itself requires two main processing tasks: the first is to scan for and diagnose the historical, review, and investigative, and the second is to prepare, execute and follow up a multi-stage mechanism to achieve a potential outcome [22]. Hypothesis development, hypothesis testing, and intervention constitute the basic form of these processes in the fields of health system governance and treatment. Machine learning can increase the development and testing of hypotheses within a health system by exposing previously concealed patterns in data and thus has the potential for major implications both at the individual patient level and at the system level [23].

Machine learning builds on current statistical techniques and uses approaches that are not based on prior data distribution assumptions, which can be used for the formulation of hypotheses and hypothesis testing by using patterns in the data. There are also several more variables to be implemented, generalizable to a much wider variety of data types, whereas machine-learning models are more difficult to understand and can result in more complicated circumstances [24]. Such techniques have been used to test and detect and forecast future incidents in the study context. Such implementations are situated in different settings, usually hospitals rather than urban environments, with consequences for reproducibility and universality in the vast majority of cases based on data from single centers. Furthermore, both within health care and in all information processing activities in society, the exponential growth of machine learning continues [25-29].

IV. POTENTIAL EFFECT OF AI ON CLINICAL CARE AND HEALTH WORKFORCE

Machine learning has become a "general technology," which is all-encompassing, can be refined over time, and has the potential to produce additional innovations. The use of such innovations appears to result in "a large economic revolution, with resulting winners and losers." Economists Acemoglu and Restrepo have studied the historical impact of automation – mechanization replacement – and claim that automation has been replaced by machines in places where machines have a differing advantage. automation is a relocation effect [26-28]. However, countervailing forces, which increase labor demand, compensate for the effect of displacement: a growth effect, which increases production and costs. This allows savings in effect for existing non-automated tasks and for the development of new non-automated tasks, in part involving direct automation technology. It is worth reviewing the clinical field best currently described in machine literature, diagnostic radiography, and seeing if this general trend might relate to health workers [230-33].

V. APPLICATIONS OF MACHINE LEARNING AND DEEP LEARNING

Since in-depth learning algorithms developed new diagnostic image analysis performing norms, some commentators predicted the eventual retirement of radiologists and challenged the need for training of new radiologists. It is possible for machine learners to manage more cases and shift responsibility for diagnostic diagnosis to non-radiologists assisted by machine learning systems as machine learning systems function more independently [34-37]. This reorientation of duties will give the healthcare sector an opportunity to reassess the mixture of expertise and deployment of radiology teams, with more primary care research and less-automatic research and unusual cases being treated by less secondary and tertiary radiologists [38].

The investigators behind a pneumonia-diagnostic machinery learning system have established a mechanism whereby the technology system first "reads" the image and points to a target for the human radiologist, thereby allowing a human decision-maker to concentrate on workflow efficiency in a way that enables them to use the picture most effectively and solve several additional cases. The same technologies are also required to turn pathology and other specialties based on image processing [39-41].

This means that machine learning produces human and computer hybrid systems. Such instances provide an ideal combination for the capacity of human beings to produce expectation, cooperate and supervise AI systems in order to manipulate AI's ability to evaluate vast quantities of data in order to identify predictive power correlations or optimize against a successive criterion [31].

VI. CONCLUSION

In this article, we addressed the direct influence of machine learning on healthcare systems but did not examine the indirect impacts of machine learning on healthcare systems, the discovery of drugs, and others. The prediction is fundamentally difficult: technology changes its environment, and the world produces new possibilities and new technology constraints. Basically, general intelligence, since a variant of it already exists in human brains, would be feasible. However, it seems impossible in the 5-10 years to systematically extrapolate current techniques for re-creating general intelligence. However, a Federation of "narrow" and "targeted" machine learning systems capable of solving central health system issues by improving decision-makers' skills and thereby setting up new standards in clinical and management operations can be and therefore should be prepared for immediately possible. This is a tremendous opportunity for the improvement of the health system, as the costs of growing decision-making capability are unlikely to be substantial across the health sector. There is no other method that can have such an effect without a corresponding cost scaling. The fixed cost of implementing machine learning technologies is considerable: the expense of research and development and re-tooling a health system is significant, but the potential scalability means that the investment rationale is straightforward. There is an opportunity to grow in machine learning by creating clinical data sets of high resolution and the appropriate data sharing frameworks and collaborative work to create both productivity and health.

REFERENCES

[1] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs.JAMA. 2016;316(22):2402-2410.

[2] Brand RJ, Rosenman RH, Sholtz RI, et al. Multivariate prediction of coronary heart disease in the Western Collaborative Group Study compared to the findings of the Framingham study. Circulation. 1976;53(2):348-355.

[3] Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data.JAMA. 2014;311 (24):2479-2480.

[4] Atun R. Transitioning health systems for multimorbidity. Lancet. 2015;386:721-2. Medline:26063473 doi:10.1016/ S0140-6736(14)62254-6

[5] Kocher R, Sahni NR. Rethinking health care labor. N Engl J Med. 2011;365:1370-2. Medline:21995383 doi:10.1056/ NEJMp1109649

[6] Badawi O, Brennan T, Celi LA, Feng M, Ghassemi M, Ippolito A, et al. Making big data useful for health care: a summary of the inaugural mit critical data conference. JMIR Med Inform. 2014;2:e22. Medline:25600172 doi:10.2196/medinform.3447

[7] Jones SS, Heaton PS, Rudin RS, Schneider EC. Unraveling the IT productivity paradox—lessons for health care. N Engl J Med. 2012;366:2243-5. Medline:22693996 doi:10.1056/NEJMp1204980

[8] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436. Medline:26017442 doi:10.1038/nature14539

[9] Beam A, Kohane I. Big data and machine learning in health care. JAMA. 2018;319:1317-8. Medline:29532063 doi:10.1001/jama.2017.18391

[10] Gul, Muhammad Tayyab, Ali Sami Dheyab, Ekremah Kheun Shaker, Norhayati Muhammad, and Aslia Natasha Pauzi. "In vitro evaluation of anti-urolithiatic properties of Strobilanthes crispus extracted using different solvents." *Research Journal of Chemistry and Environment. Vol* 24 (2020): 1.

[11] Marcus G. Deep learning: A critical appraisal. arXiv:1801.00631. 2018.

[12] Atun R, Aydın S, Chakraborty S, Sümer S, Aran M, Gürol I, et al. Universal health coverage in Turkey: enhancement of equity. Lancet. 2013;382:65-99. Medline:23810020 doi:10.1016/S0140-6736(13)61051-X

[13] Henglin M, Stein G, Hushcha PV, Snoek J, Wiltschko AB, Cheng S. Machine learning approaches in cardiovascular imaging. Circ Cardiovasc Imaging. 2017;10:e005614. Medline:28956772 doi:10.1161/CIRCIMAGING.117.005614

[14] Stanford University. Algorithm outperforms radiologists at diagnosing pneumonia [Internet]. Stanford News. 2017. Available: https://news.stanford.edu/2017/11/15/algorithm-outperforms-radiologists-diagnosing-pneumonia/. Accessed: 20 March 2018.

[15] Johnson AE, Pollard TJ, Mark RG. 2017, November. Reproducibility in critical care: a mortality prediction case study. Machine Learning for Healthcare Conference 2017. JMLR W&C Track Volume 68. Available: http://proceedings.mlr. press/v68/johnson17a/johnson17a.pdf. Accessed: 20 March 2018.

[16] Celi LA, Moseley E, Moses C, Ryan P, Somai M, Stone D, et al. From pharmacovigilance to clinical care optimization. Big Data. 2014;2:134-41. Medline:26576325 doi:10.1089/big.2014.0008

[17] Brynjolfsson E, Mcafee AN. The business of artificial intelligence. Harv Bus Rev. 2017. Available: https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence. Accessed: September 2018.

[18] Helpman E, Trajtenberg M. Diffusion of general purpose technologies. National

Bureau of Economic Research. 1996. No. w5773.

[19] Trajtenberg M. AI as the next GPT: a Political-Economy Perspective. National Bureau of Economic Research. 2018. No. w24245.

[20] Kalli, Venkata & Gadde, Sai. (2020). Technology Engineering for Medical Devices - A Lean Manufacturing Plant Viewpoint. 9. 1-6. 10.17148/IJARCCE.2020.9401.

[21] Acemoglu D, Restrepo P. Artificial intelligence, automation and work. National Bureau of Economic Research 2018. No. w24196.

[22] Siddartha M. The algorithm will see you now. New Yorker. 2017;93:46-53.

[23] Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv:1711.05225v3 [cs.CV].

[24] Golden JA. Deep learning algorithms for detection of lymph node metastases from breast cancer: helping artificial intelligence be seen. JAMA. 2017;318:2184-6. Medline:29234791 doi:10.1001/jama.2017.14580

[25] Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep. 2018;8:3395. Medline:29467373 doi:10.1038/s41598-018-21758-3

[26] Ein-Dor L, Kela I, Getz G, Givol D, Domany E. Outcome signature genes in breastcancer: is there a unique set? Bioinformatics 2005;21:171–8.

[27] Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robustgene list for predicting outcome in cancer. Proc Natl Acad Sci 2006;103:5923–8.

[28] Ayer T, Alagoz O, Chhatwal J, Shavlik JW, Kahn CE, Burnside ES. Breast cancer risk es-timation with artificial neural networks revisited. Cancer 2010;116:3310–21.

[29] Platt JC, Cristianini N, Shawe-Taylor J. Large margin DAGs for multiclass classifica-tion; 1999 547–53.

[30] Adams S. Is Coursera the beginning of the end for traditional higher education?Higher Education; 2012.[22]Cicchetti D. Neural networks and diagnosis in the clinical laboratory: state of the art.Clin Chem 1992;38:9–10.

[31] Cochran AJ. Prediction of outcome for patients with cutaneous melanoma. PigmentCell Res 1997;10:162–7.[24]Exarchos KP, Goletsis Y, Fotiadis DI. Multiparametric decision support system for theprediction of oral cancer reoccurrence. IEEE Trans Inf Technol Biomed 2012;16:1127–34.

[32] Kononenko I. Machine learning for medical diagnosis: history, state of the art andperspective. Artif Intell Med 2001;23:89–109.[26]Park K, Ali A, Kim D, An Y, Kim M, Shin H. Robust predictive model for evaluatingbreast cancer survivability. Engl Appl Artif Intell 2013;26:2194–205.

[33] Sun Y, Goodison S, Li J, Liu L, Farmerie W. Improved breast cancer prognosis throughthe combination of clinical and genetic markers. Bioinformatics 2007;23:30–7.

[34] Bottaci L, Drew PJ, Hartley JE, Hadfield MB, Farouk R, Lee PWR, et al. Artificial neuralnetworks applied to outcome prediction for colorectal cancer patients in separate in-stitutions. Lancet 1997;350:469–72.

[35] Maclin PS, Dempsey J, Brooks J, Rand J. Using neural networks to diagnose cancer. JMed Syst 1991;15:11–9.

[36] Simes RJ. Treatment selection for cancer patients: application of statistical decisiontheory to the treatment of advanced ovarian cancer. J Chronic Dis 1985;38:171–86.

[37] Akay MF. Support vector machines combined with feature selection for breastcancer diagnosis. Expert Syst Appl 2009;36:3240–7.

[38] Chang S-W, Abdul-Kareem S, Merican AF, Zain RB. Oral cancer prognosis based onclinicopathologic and genomic markers using a hybrid of feature selection

andmachine learning methods. BMC Bioinforma 2013;14:170.
[39] Chuang L-Y, Wu K-C, Chang H-W, Yang C-H. Support vector machine-basedprediction for oral cancer using four snps in DNA repair genes; 2011 16–8.
[40] Eshlaghy AT, Poorebrahimi A, Ebrahimi M, Razavi AR, Ahmad LG. Using threemachine learning techniques for predicting breast cancer recurrence. J Health MedInform 2013;4:124.
[41] Exarchos KP, Goletsis Y, Fotiadis DI. A multiscale and multiparametric approach formodeling the progression of oral cancer. BMC Med Inform Decis Mak 2012;12:136.