Control and Optimization in Applied Mathematics (COAM)
Vol. 1, No. 2, Autumn-Winter 2016(63-75), ©2016 Payame Noor University, Iran

# Comparative Analysis of Machine Learning Algorithms with Optimization Purposes

R. Alesheykh*

Department of Information Technology, Payame Noor University,
P.O. BOX. 19395-3697, Tehran, Iran.

Received: February 5, 2016; Accepted: November 27, 2016.

**Abstract.** The field of optimization and machine learning are increasingly interplayed and optimization in different problems leads to the use of machine learning approaches. Machine learning algorithms work in reasonable computational time for specific classes of problems and have important role in extracting knowledge from large amount of data. In this paper, a methodology has been employed to optimize the precision of defect detection of concrete slabs depending on their qualitative evaluation. Based on this idea, some machine learning algorithms such as C4.5 decision tree, RIPPER rule learning method and Bayesian network have been studied to explore the defect of concrete and to supply a decision system to speed up the defect detection process. The results from the examinations show that the proposed RIPPER rule learning algorithm in combination with Fourier Transform feature extraction method could get a defect detection rate of 93% as compared to other machine learning algorithms.

**Keywords.** Decision tree, Bayesian network, Rule learning algorithm, Optimization, Soft computing.

**MSC.** 97N80; 97R40.

* Corresponding author
alesheykh@pnu.ac.ir
http://mathco.journals.pnu.ac.ir

## 1 Introduction

A qualitative evaluation of concrete can be easily obtained by tapping methods that receive sounds with human ears. When the hammer is struck on normal concrete, a ringing sound is created. However, on areas where delaminations or cracks occur, the striking of the hammer produces a drum-like sound. These methods completely depend on the detection of unclear sounds to identify internal defects. Evaluation by human ears is greatly affected by the experiences and subjectivity of inspectors. In order to solve these problems, methods using devices such as microphone, to receive sounds and to analyze the signals have been employed [1]. Development of a system that can recognize the condition of concrete and guarantee its health can lead to build high speed digital signal processing platforms to automate the recognition procedure. Condition data plays a fundamental role not only in the control of concrete safety, e.g. identification of potential failures such as porosity and delaminations inside concrete, but also for decision support, improving the power in planning optimized maintenance and renewal works.

In this study, machine learning algorithms such as decision tree, rule learning method and Bayesian network will be examined across feature extraction methods such as Fourier and Wavelet Transforms for classification of signals acquired from concrete slabs to detect the destruction of concrete. While the origins of these machine learning approaches are distinct and the underlying algorithms differ substantially, the fundamental processes are the same; they are all inductive methods. The mentioned algorithms were examined using WEKA software which provides a safe chance of testing several machine learning algorithms. The objective of this research is to assess the relative performance of some well-known machine learning techniques. The result of this study can aid to increase the reliability and consistency of the classification.

The interplay between optimization and machine learning is one of the most important developments in modern computational science. Optimization formulations and methods are proved to be vital in designing algorithms to extract essential knowledge from huge volumes of data. Machine learning, however, is not simply a consumer of optimization technology but a rapidly evolving field that is itself generating new optimization ideas [29]. Computer-based techniques, offer advantages of improved speed and accuracy of analysis, especially for large-volume inspection information [20].

Next section reviews some previous related researches. Section 3 explains how data was prepared. Section 4 briefly indicates the feature extraction methods. Section 5 describes employed machine learning algorithms. Section 6 presents results and discussion and finally the last section concludes the paper.

## 2 Related Works

This section briefly surveys some previous work related to the current representation in this paper. Several works have addressed the issue of using machine learning algorithms for solving different problems.

The study [24] used machine learning algorithms as a research methodology to develop a housing price prediction model. The authors developed a housing price prediction model based on machine learning algorithms such as C4.5, RIPPER, Naïve Bayesian, and AdaBoost and compared their classification accuracy performance. An improved housing price prediction model was then proposed to assist a house seller or a real estate agent to make better informed decisions based on house price valuation. The experiments demonstrate that the RIPPER algorithm, based on accuracy, consistently outperforms the other models in the performance of housing price prediction.

A malware detection system based on the data mining and machine learning technique has been proposed in [22]. Malware represents a serious threat to the security of computer systems. Traditional malware detection techniques like signature-based, heuristic-based, Specification-based detection are used to detect the known malware. These techniques detect the known malware accurately, but unable to detect the new, unknown malware. The proposed method in [22] consists of disassemble process, feature extraction process and feature selection process. Three classification algorithms were employed on dataset to generate and train the classifiers named as Ripper, C4.5, IBk.

The goal of study [9] was to find an effective machine learning method for classifying ElectroMyoGram (EMG) signals by applying de-noising, feature extraction and classifier. The study presented a framework for classification of EMG signals using multi-scale principal component analysis for de-noising, discrete wavelet transform for feature extraction and decision tree algorithms for classification. The presented framework automatically classified the EMG signals as myopathic, ALS or normal, using CART, C4.5 and random forest decision tree algorithms. Decision tree algorithms are extensively used in machine learning field to classify biomedical signals. De-noising and feature extraction methods were also utilized to get higher classification accuracy.

Since the application of microarray data for cancer classification is important, researchers have tried to analyze gene expression data using various computational intelligence methods. A novel method for gene selection has been proposed in [2]. This method utilizing particle swarm optimization, is combined with a decision tree as the classifier to select a small number of informative genes from the thousands of genes in the data that can contribute in identifying cancers.

A class of relevant speech signal processing algorithms as probabilistic inference problems has been described in [19]. Starting with an observation model that relates all involved random variables, the authors converted the respective joint probability density function into its Bayesian network representation in order to infer the desired signal estimates.

In a novel research, impact acoustic parameters obtained from received sound generated by impact on concrete surface were investigated to develop an evaluation system of defects in concrete [1]. As impact acoustic parameters, frequency distribution was employed. From analytical and experimental results, it was likely possible to estimate defect sizes using the relation between the resonance frequencies of impact sounds and defects diameter.

To ensure the safety and the serviceability of civil infrastructure it is essential to visually inspect and assess its physical and functional condition. Another work [17] presented the current state of practice of assessing the visual condition of vertical and horizontal civil infrastructure; in particular of reinforced concrete bridges, precast concrete tunnels, underground concrete pipes, and asphalt pavements. Since the rate of creation and deployment of computer vision methods for civil engineering applications has been exponentially increasing, the main part of the paper presented a comprehensive synthesis of the state of the art in computer vision based defect detection and condition assessment related to concrete and asphalt civil infrastructure.

Another work [3] was concerned to explore and develop a heart sound based diagnostic system. classification of the heart murmurs by their associated heart conditions led to the development of a modularized approach to the computer-aided auscultation based on the conventional cardiac auscultation. It was proposed that the murmurs can be characterized based on their acoustic qualities. The pattern classification framework was able to classify innocent murmurs and abnormal murmurs with an accuracy of better than the average cardiologists. Different signal processing techniques such as Fourier Transform and Continuous Wavelet Transform were used for feature extraction. In another similar work [33], a novel method was put forward for automatic identification of the normal and abnormal heart sounds. After the original heart sound signal was pre-processed, it was analyzed by the optimum multi-scale wavelet packet decomposition and then the wavelet-time entropy was applied to extract features from the decomposition components. The extracted features were then applied to a support vector machine for identification of the normal and five types of abnormal heart sounds.

A research project was also developed for classifying the environmental sounds. In paper [16], an environmental sound classification algorithm using spectrogram pattern matching along with neural network and k-nearest neighbour classifiers was proposed. The recognition of environmental sounds can benefit crime investigations, warning systems for elderly persons, and security systems.

## 3 Data Preparation

The most accessible way of testing concrete slabs is through a process called sounding. Sounding involves striking the concrete surface and interpreting the sound produced. Solid concrete will produce a ringing sound, while concrete that is spalled, delaminated, or contains voids will produce a flat or hollow sound. Data collection was performed by collecting signals from the concrete slabs. All measurements were made by a microphone and a conventional PC sound card which could sample in stereo. The process was accomplished on 160 flat thick piece of concrete of which 90 pieces were in normal situation and 70 pieces were in damaged form. It should be mentioned that the final situation of concrete slabs are determined by an engineer to specify their target class.

## 4 Feature Extraction Methods

The aim of feature extraction is to present sound signals compactly and efficiently. Features are extracted to obtain the most significant information from the original data with an aim of reducing computational load for further classification task. In order to analyze a signal whose frequency components vary in time, a time-frequency distribution of the signal is a good choice. The main time-frequency techniques that are commonly mentioned in this work are:

1. Short-term Fourier Transform (STFT)
2. Discrete Wavelet Transform (DWT)

These two techniques use different algorithms to produce a time-frequency representation of a signal. While STFT uses a standard Fourier transform over several windows and specifies complex amplitude versus time and frequency for any signal, Wavelet-based techniques apply a mother wavelet to a waveform and analyzes the signal by decomposing it into its approximate and detailed information [32], which is accomplished by the use of successive high-pass and low-pass filtering and sub-sampling operations. The STFT technique is the simplest method to be used to analyze a time-varying signal with frequency fluctuation over time. To compute the STFT of an entire signal, a sliding window is used to divide the signal into several blocks. The Fast Fourier transform (FFT) is then applied to each data block to obtain the frequency contents (See [23] for more details about FFT). The STFT aligns the center of the first sliding window with the first sample of the signal and extends the signal at the beginning with zeros or the signal itself. In Figure 1 the computing procedure of the STFT is shown.

The concept of the DWT is that filters with different cut-off frequencies are utilized to analyze the signal at different scales. Firstly, the signal is passed through a high-pass filter to analyze high frequencies, and then it is passed through a low-pass filter to analyze low frequencies. Generally, by using the DWT, a multi-resolution analysis can be performed at different frequency bands with different resolutions by decomposing the time domain signal [6, 28]. Two sets of functions called the wavelet function and the scaling function, which are associated with the high-pass (HP) and low-pass (LP) filters are used, respectively. At the first level, the original signal is decomposed by passing it through both of these filters

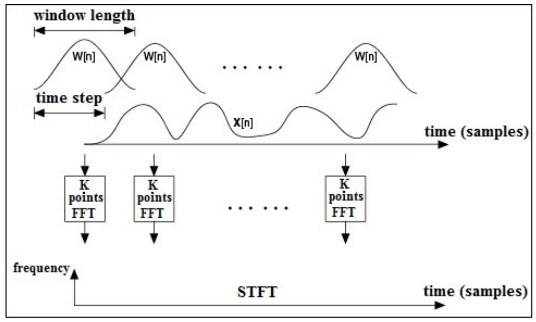


**Figure 1:** Short-term Fourier Transforms (STFT) Principle

and emerges as two signals, each one having the same number of samples as the original signal, and are termed as coefficients. In order to keep the total number of coefficients in the produced filtered signals equal to the original signal samples they are then down-sampled by a factor of 2, by keeping only one sample out of two successive samples [15]. The signal decomposition using DWT is shown in Figure 2.

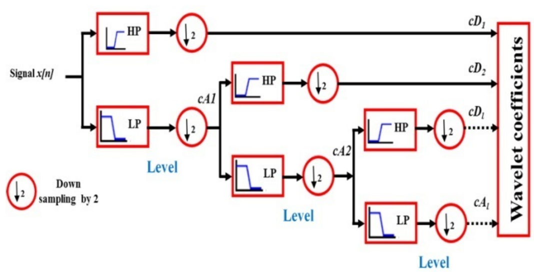


**Figure 2:** Signal decomposition using discrete wavelet transform

In this article, sounds collected by making tests on concrete slabs were used in pre-processing and feature extraction stages. These tasks were performed using signal processing toolbox in MATLAB. The output of feature extraction process together with Principle Component Analysis (PCA) was 80 features extracted by STFT technique and 80 in the case of using DWT. The feature vector is then presented to the machine learning algorithms for classification task concerning the condition assessment of the concrete slabs.

## 5 Machine Learning Algorithms

The field of machine learning is concerned with the question of how to construct computer programs that automatically improves with experience [21, 30]. Given that, each machine learning method has its strengths and limitations and real world problems do not always satisfy the assumptions of a particular method, one approach is to apply many appropriate methods and select the one that provides the best solution. This article explores the application of effective machine learning to overcome challenges associated with data analysis and demonstrates how machine learning algorithms and signal processing techniques have contributed and are contributing to the research [35]. In this work, some machine learning algorithms such as decision tree, rule learning method and Bayesian network are examined in WEKA software and the results concluded from the mentioned methods will be compared. What follows next is a brief discussion concerning the above mentioned methods.

### 5.1 Decision Tree Algorithm

Decision tree is a predictive model that maps target values from observations. It is a flow-chart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent classes or class distributions [11, 18]. In order to classify an unknown sample using decision trees, the attribute values of the sample are tested against the decision tree. To learn which attribute should be tested at the root of the tree, each instance attribute is evaluated using a statistical test to determine how well it alone classifies the training examples. A path is traced from the root to a leaf node that holds the class prediction for that sample. The popular algorithm which has been used for generating decision tree in the current work is C4.5 [21]. C4.5 algorithm uses a divide-and-conquer approach for growing decision trees. The splitting node strategy is based on the computation of the information gain ratio. The basic idea is that each node should hold a question concerning the attribute which is the most informative among the set of attributes not yet considered in the path from the root to that node. Information value, called entropy, also measures how informative is the association of an attribute with a node [10]. The sub-trees are spanned by splitting the training dataset according to this strategy. Once the initial decision tree is constructed, a pruning procedure is initiated to decrease the overall tree size and decrease the estimated error rate of the tree [26].

### 5.2 Rule Learner Algorithm

Rule learner (rule induction) method performs an iterative process consist of two steps. In the first step, a rule that covers a subset of the training examples is generated and then all examples covered by the rule are removed from the training set before subsequent rules are learned. This process is iteratively repeated until there are no examples left be covered. The final rule set is the collection of the rules discovered at every iteration of the process. Rule learner algorithms expect positive and negative examples for an unknown concept. If any of the learned rules fires for a given example, the example is classified as positive and if no rule fires, it is classified as negative [7, 12]. The rule learner algorithm employed in this work is Repeated Incremental Pruning to Produce Error Reduction (RIPPER). RIPPER [4] builds a rule set by repeatedly adding rules to an empty rule set until all positive examples are covered. Rules are formed by greedily adding conditions to the antecedent of a rule (starting with empty antecedent) until no negative examples are covered. The pruning stage then attempts to simplify the rule by removing a sequence of conditions at the end of the rule. This greedy process examines which deleted sequence maximizes the proportion of positive examples over total examples covered. Afterwards, a rule set is constructed, an optimization post pass massages the rule set so as to reduce its size and improve its fit to the training data. The optimization stage examines each rule in sequence and decides whether the rule needs to be replaced, revised or kept.

Rule induction and decision tree methods both split a data set into subgroups on the basis of the relationships between predictors and the output field. Rules can be symmetric whereas trees must select one attribute to split on first, and this can lead to trees that are much larger than an equivalent set of rules [34].

### 5.3 Bayesian Network Learning

Bayesian networks (BNs) [8, 25, 27] are a probabilistic framework for reasoning under uncertainty. BNs are directed acyclic graphs where the nodes are random variables which denote attributes, features or hypothesis and the arcs specify the conditional independencies between the random variables. Associated with each node (child node) is a probability distribution on that node given the state of its parent nodes. A Bayesian network specifies a joint distribution in a structured form. The joint distribution described by a graph is computed by the product of conditional probabilities for each node conditioned on the variables corresponding to the parents of that node in the following way:

$$\mathrm{P}\left(\mathrm{y}_1, ..., \mathrm{y}_\mathrm{n}\right) = \prod_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{P}\left(\mathrm{y}_\mathrm{i} \,|\mathrm{Parents}\left(\mathrm{Y}_\mathrm{i}\right)\right)$$

where $\mathrm{y}_i$ represents the value of the random variable $\mathrm{Y}_i$ and parent($\mathrm{Y}_i$) denotes the value of the parents of $\mathrm{Y}_i$.

In order to specify the probability distribution of a BN, one must give prior probabilities to all root nodes and conditional probabilities for all other nodes, given all possible combinations of their direct predecessors. Once the network is constructed, it constitutes an efficient device to perform probabilistic inference. Many algorithms have been proposed on learning Bayesian network structure. One method is score-and-search approach [13, 31], which poses the learning problem as a structure optimization problem. Namely, it uses a score metric to evaluate every candidate network structure, and then, finds a network structure with the best score.

In the current work, the Bayesian network represents the probabilistic relations between extracted features and the target class which is a normal or damaged concrete. Given the features, the network computes the probabilities of being kept in either normal or damaged class. BN learning algorithm also uses the general purpose search method of simulated annealing to find a well scoring network structure.

## 6 Results and Discussion

In this work, machine learning algorithms such as C4.5 decision tree, RIPPER rule learning method and Bayesian networks were tested for classifying the concrete slabs into normal and damaged classes. WEKA, an open source machine learning framework which is a collection of machine learning algorithms was employed to do the classification task. Feature extraction techniques such as STFT and DWT together with PCA technique were used to extract the most important features of the signals. Using PCA, the redundant features were removed effectively and this made the efficiency improved.

Before proceeding any further with the classification process, it is worth mentioning that the vocal signals collected by making experiments on 160 concrete slabs were partitioned into training and test sets.

Since the classification rate reported for the current work is based on the analysis of a very small set of data and to investigate how the discussing methods are performed on new or different data sets, cross-validation has been used. Cross-validation is a method for evaluating machine learning algorithm by dividing data into training and testing sets. In cross-validation a fixed number of partitions of the data called folds is determined. In this experiment 10-fold cross-validation has been chosen for partitioning the dataset. This means that the data is split into ten approximately equal partitions and each in turn is used for testing and the remainder is used for training. The procedure is repeated ten times so that, by the end, every instance has been used exactly once for the testing. Many experiments on numerous datasets have shown that 10-fold cross validation is about the right number of folds to get more robust results as classification rate.

The results obtained from each algorithm with the mentioned feature extraction technique have been reported in Tables 1 and 2. Table 1 reports the percentage of classification using 10-fold cross-validation while Table 2 shows classification rate using partitioning the data into 70% and 30% for predetermined training and testing set respectively. Results achieved in this work demonstrate that the combination of STFT feature extraction technique together with RIPPER rule learning algorithm will result in the best classification rate which is 93% (See Table 1). Results show that the entire machine learning algorithm employed in this work performed much better when using cross-validation (See Table 2). Cross-validation is intended to avoid the possible bias introduced by relying on

**Table 1:** Results by using 10-fold cross-validation

| Machine learning algorithms | Classification rate using STFT | Classification rate using DWT |
|---|---|---|
| C4.5 | 91% | 79% |
| RIPPER | 93% | 82% |
| Bayesian network | 90% | 76.5% |

**Table 2:** Results by using training and testing sets

| Machine learning algorithms | Classification rate using STFT | Classification rate using DWT |
|---|---|---|
| C4.5 | 90% | 78% |
| RIPPER | 90% | 80% |
| Bayesian network | 88% | 76% |

any one particular division into test and train components. By partitioning the original set into several parts and compute an average score over the different partitions i.e. average number of corrected classified samples over all the samples in every partition, more reliable result will be concluded. The results presented in this paper present the obvious superiority of STFT technique compared to DWT technique due to the characteristic of DWT. The fact that DWT is usually used for encoding and decoding signals might be responsible for the slightly low classification rate of DWT in the current case [5].

According to the results comprehended from the Tables 1 and 2, RIPPER rule learning algorithm has achieved better classification accuracy in contrast with C4.5 decision tree. The reason might be that rules are much more compact than trees and a default rule can cover cases not specified by other rules [34]. When a decision tree is built, many of its branches may reflect anomalies in training data. In addition, when adding new rules to an existing rule set, there is no need to disturb previous rules, but to add a tree structure may require modifying the whole tree. Results achieved in the current work also indicate that Bayesian network showed lower performance compared with two other techniques. This is because Bayesian network requires initial knowledge for assigning probabilities. Either an expert must provide prior probabilities for all root nodes and conditional probabilities for all other nodes or they can be obtained from an algorithm which automatically induces them. The quality of the results of the network strongly depends on the quality of the prior beliefs. The nearly same work [14] which has already been done also demonstrates the higher performance of RIPPER rule learning algorithm and decision tress against some kinds of networks on sound data.

## 7 Conclusion

In the present paper, with the features extracted via STFT and DWT techniques, an inspection approach was proposed to facilitate the robust assessment of defect detection of concrete slabs. Machine leaning algorithms such as C4.5 decision tree, RIPPER rule learning algorithm and Bayesian network were chosen to be compared for classifying the concrete slabs into normal and damaged classes. The fact that condition assessment innovations has outstanding benefits, demands for increasing focus and investment in many organizations around the world. After several experiments, the final classification rate demonstrated an accuracy of 93% using the combination of RIPPER rule learning algorithm with STFT feature extraction technique. The relative effectiveness and classification efficiency of the techniques used in the current case will become apparent when they are applied to a larger database of sounds. Although 93% classification accuracy seems so powerful, it would be more efficient if one tries more machine learning algorithms for a classification stage to achieve the highest possible performance.

**References**


[1] Asano M., Kamada T., Kunieda M., Rokugo K. (2003). " Impact acoustics methods for defect evaluation in concrete ", International Symposium Non-Destructive Testing in Civil Engineering.

[2] Chen H., Wang J., Wang M., Angelia M. (2014). " Applying particle swarm optimization-based decision tree classifier for cancer classification on gene expression data ", Applied Soft Computing, 24, 773-780.

[3] Chia-Hsuan S. (2012). " Acoustic based condition monitoring ", The Graduate Faculty of The University of Akro, Ph.D dissertation.

[4] Cohen W. W. (1995). " Fast effective rule induction ", Proceedings of International Conference on Machine Learning, Lake Tahoe, CA.

[5] Cowling M., Sitte R. (2003). " Comparison of techniques for environmental sound recognition ", Pattern Recognition Letters, 24(15), 2895-2907.

[6] Debdas S., Quereshi M., Reddy A., Chandrakar D., Pansari D. (2011). " A wavelet based multiresolution analysis for real time condition monitoring of AC machine using vibration analysis ", International Journal of Scientific and Engineering Research.

[7] Furnkranz J. (1999). " Separate-and-conquer rule learning ", Artificial Intelligence Review, 13, 3–54.

[8] Fuster-Parra P., Tauler P., Bennasar-Veny M., Ligęza A., López-González A., Aguiló A. (2016). " Bayesian network modelling: A case study of an epidemiologic system analysis of cardiovascular risk ", Computer Methods and Programs in Biomedicine, 126, 128-142.

[9] Gokgoz E., Subasi A. (2015). " Comparison of decision tree algorithms for EMG signal classification using DWT ", Biomedical Signal Processing and Control, 18, 138-144.

[10] Gray R. (1990). " Entropy and information theory ", New York, Springer-Verlag.

[11] Han J., Kamber M. (2001). “ Data mining: concepts and techniques ”, San Francisco, Morgan Kaufmann.

[12] Han J., Kamber M., Pei J. (2011). “ Data mining: concepts and techniques ”, Third Edition, Morgan Kaufmann.

[13] Heckerman D. (1999). “ A tutorial on learning with Bayesian networks ”, In Learning in Graphical Models, MIT Press, Cambridge, MA.

[14] Henry M. (2011). “ Learning techniques for identifying vocal regions in music using the wavelet transformation, Version 1.0 ”, Graduate School of Arts and Sciences of Georgetown University, Washington DC, M.S. dissertation.

[15] Jaber A., Bicker R. (2016). “ Industrial robot backlash fault diagnosis based on discrete wavelet transform and artificial neural network ”, American Journal of Mechanical Engineering, 4(1), 21-31.

[16] Khunarsal P., Lursinsap C., Raicharoen T. (2013). “ Very short time environmental sound classification based on spectrogram pattern matching ”, Information Sciences, 243, 57-74.

[17] Koch C., Georgieva K., Kasireddy V., Akinci B., Fieguth P. (2015). “ A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure ”, Advanced Engineering Informatics, 29(2), 196-210.

[18] Ma L., Destercke S., Wang Y. (2016). “ Online active learning of decision trees with evidential data ”, Pattern Recognition, 52, 33-45.

[19] Maas R., Huemmer C., Hofmann C., Kellermann W. (2014). “ On Bayesian networks in speech signal processing ”, 11th ITG Conference on Speech Communication, Germany.

[20] McNab A., Dunlop I. (1990). “ Artificial intelligence techniques for the automated analysis of ultrasonic NDT data ”, IEE Colloquium on Advances in Transducers, Equipment and Data Processing to Improve the Reliability of NDT, 1-8.

[21] Mitchell T. (1997). “ Machine learning ”, McGraw-Hill.

[22] More S., Gaikwad P. (2016). “ Trust-based voting method for efficient malware Detection ”, Procedia Computer Science, 79, 657-667.

[23] Oran Brigham E. (1988). “ The fast Fourier transform and its applications ”, Prentice-Hall.

[24] Park B., Bae J. K. (2015). “ Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia ”, Expert Systems with Applications, 42(6), 2928-2934.

[25] Pearl J. (1988). “ Probabilistic reasoning in intelligent systems: networks of plausible inference ”, Morgan Kaufmann.

[26] Quinlan J. (1993). “ C4.5: Programs for machine learning ”, San Francisco, Morgan Kauffman.

[27] Rancoita P., Zaffalon M., Zucca E., Bertoni F., de Campose C. (2016). “ Bayesian network data imputation with application to survival tree analysis ”, Computational Statistics and Data Analysis, 93, 373-387.

[28] Sawicki J., Sen A., Litak G. (2009). “ Multiresolution wavelet analysis of the dynamics of a cracked rotor ”, International Journal of Rotating Machinery.

[29] Suvrit S., Nowozin S., Wright S. J. (2012). “ Optimization for machine learning ”, MIT Press.

[30] Sugiyama M. (2016). “ Statistical machine learning ”, Introduction to Statistical Machine Learning, 3-8.

[31] Suzuki J. (1999). ” Learning Bayesian belief networks based on the minimum description length principle: basic properties ”, IEICE Transactions on Fundamentals, 82(10), 2237−2245.

[32] Vetterli M., Kovacevic J. (1995). “ Wavelets and subband coding ”, Prentice Hall.

[33] Wang Y., Li W., Zhou J., Li X., Pu Y. (2014). “ Identification of the normal and abnormal heart sounds using wavelet-time entropy features based on OMS-WPD ”, Future Generation Computer Systems, 37, 488-495.

[34] Witten I., Eibe F. (2005). “ Data mining: practical machine learning tools and techniques ”, Morgan Kaufmann.

[35] Yu J., Zhou H., Gao X. (2015). “ Machine learning and signal processing for human pose recovery and behaviour analysis ”, Signal Processing, 110, 1-4.

# تحلیل تطبیقی الگوریتم‌های یادگیری ماشین با اهداف بهینه‌سازی

آل‌شیخ ر.
مربی مهندسی فناوری اطلاعات
ایران، تهران، گروه مهندسی فناوری اطلاعات، دانشگاه پیام نور، صندوق پستی ۳۶۹۷-۱۹۳۹۵.
alesheykh@pnu.ac.ir

**چکیده**

مبحث بهینه‌سازی و یادگیری ماشین به‌صورت گسترده‌ای به‌هم مرتبط هستند و بهینه‌سازی در مسایل مختلف منجر به استفاده از روش‌های یادگیری ماشین می‌گردد. الگوریتم‌های یادگیری ماشین برای کلاس‌های ویژه‌ای از مسایل در یک زمان محاسباتی منطقی کار می‌کنند و نقش مهمی در استخراج دانش از حجم انبوهی از داده‌ها دارند. در این مقاله یک روش برای بهینه‌سازی دقت تشخیص نقص قطعه‌های بتنی بر اساس ارزیابی کیفی آن‌ها به‌کار گرفته شده است. بر این اساس، چند الگوریتم یادگیری ماشین از جمله درخت تصمیم‌گیری C۴,۵ ، روش یادگیری قاعده ریپر و شبکه بیزین، برای بررسی نقص در بتن مورد مطالعه قرار گرفته‌اند تا یک سیستم تصمیم‌گیری برای سرعت بخشیدن به فرآیند تشخیص نقص مهیا گردد. نتایج آزمایش‌ها نشان می‌دهد که میزان تشخیص نقص ۹۳ درصد با استفاده از الگوریتم یادگیری قاعده ارائه شده به همراه روش استخراج ویژگی تبدیل فوریه در مقایسه با سایر الگوریتم‌های یادگیری ماشین حاصل شده است.

**کلمات کلیدی**

درخت تصمیم‌گیری، شبکه بیزین، روش یادگیری قاعده، بهینه‌سازی، محاسبات نرم.