Significant Impact of Improved Machine Learning Algorithm in The Processes of Large Data Sets

[1]Virendra Tiwari, [2]Balendra Garg, [3]Uday Prakash Sharma

[1]Assistant Professor, [2]Assistant Professor, [3]Lab Instructor

[2]Department of Computer Science & Application, AKS University, Satna, Madhya Pradesh, India

ABSTRACT

The machine learning algorithms are capable of managing multi-dimensional data under the dynamic environment. Despite its so many vital features, there are some challenges to overcome. The machine learning algorithms still requires some additional mechanisms or procedures for predicting a large number of new classes with managing privacy. The deficiencies show the reliable use of a machine learning algorithm relies on human experts because raw data may complicate the learning process which may generate inaccurate results. So the interpretation of outcomes with expertise in machine learning mechanisms is a significant challenge in the machine learning algorithm. The machine learning technique suffers from the issue of high dimensionality, adaptability, distributed computing, scalability, the streaming data, and the duplicity. The main issue of the machine learning algorithm is found its vulnerability to manage errors. Furthermore, machine learning techniques are also found to lack variability. This paper studies how can be reduced the computational complexity of machine learning algorithms by finding how to make predictions using an improved algorithm.

Keywords: Machine learning algorithms, Support vector machines, Deep Learning.

I. INTRODUCTION

Machine Learning (ML) is a growing technology that makes computers to learn automatically from past data which unfold from the Artificial Intelligence, a field of computer science. Machine Learning (ML) is a multidisciplinary field that is a combination of statistics and computer science algorithms. This mechanism is widely used in predictive analyses and classification. The second section of this study focuses to influence the different machine learning methods and algorithms.

Here we focus the main approaches and case studies of using machine learning for forecasting in various areas such as tourism demand forecasting, stock price forecasting, solar irradiation forecasting, supply chain demand, and working of neural network in machine learning methods.

Over the past decades, scientific growth and advancement in technologies now have enabled AI systems to learn and improve through past experiences without any explicit assistance or help if they exposed to new data. Eventually, it leads to the technology of Machine learning (ML) which uses various learning algorithms to learn from the available data [1]. Machine learning technology uses data mining techniques to extract the information from the huge amount of data helps to build a better model which predicts the output more accurately. Data Mining and Machine Learning techniques explore data from end to end to find the possible hidden patterns inside the dataset [3]. Machine Learning and data mining algorithms have been deployed in various fields such as Computer networking, image recognition, speech recognition, travel, and tourism industry, financial forecasting, telecommunication industry, and electric load forecasting, and so on [3]. Logistic Regression is one of the important and widely used Machine Learning algorithms for binary classification. This simple Algorithm that you can use as a performance baseline, it is simple and easy to implement and it does well enough in many tasks [33].

II. METHODS AND MATERIAL

Methods used in Machine Learning

Over the past few years, various ML algorithms were introduced. Only a few of them were able to solve the problem with accuracy so they replaced by another one [4]. At a broad level, ML algorithms can be classified into three types: supervised learning, unsupervised learning, and reinforcement learning, which are displayed in the following figure:
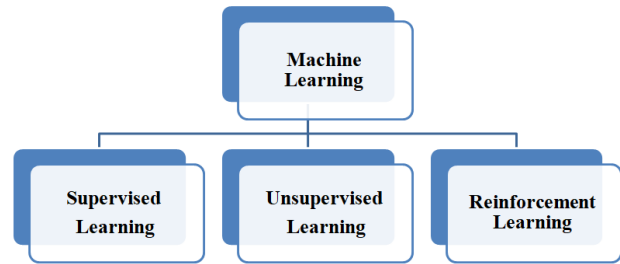

Figure 1. Types of Machine Learning

3.1 Supervised Learning

In this method, we provide sample labeled data as an input to the machine learning system to train it, and accordingly, it predicts the output. It consists of a given set of input variables that are trained, pre-labeled, and target data. This method uses the input variables it generates a mapping function to map inputs to required outputs [6]. The procedure of parameter adjustment continues until the system achieved a suitable accuracy extent regarding the teaching data. This method is the same as when a student learns something under the supervision of the teacher. An example of supervised learning method is **spam filtering**.

3.2 Unsupervised Learning

In this algorithm a machine learns without any supervision means we only have training data that has not been labeled rather an outcome data. It is used in classifiers by recognizing existing and available patterns or cluster in the input datasets without any supervision [5].

3.3 Reinforcement Learning

It is a feedback-based learning method, in which a learning machine gets a reward for each right action and gets a punishment for each wrong action. Applying this algorithm machine is actually trained to map action to a specific decision so the reward or feedback signals can be generated. The machine trained itself by this process to find the most

rewarding actions by reward and punishment using past experience.

III. Algorithms of Machine Learning

There are enormous numbers of algorithms used by machine learning technology. They are designed to enhance models of machine learning and implemented in it [5]. All these algorithms can be grouped by their working and learning methodology, as follows:

4.1 Regression Algorithms

Regression algorithms are used to evaluate if there is any relationship between the input variable and the output variable. It is used for the prediction of continuous variables, like Market Trends, Weather forecasting, etc. [26]. Continuously varying value is predicted by this technique and variable can be a price or a temperature. Some popular regression algorithms are as follows:

- Linear Regression algorithm
- Ordinary Least Squares Regression
- Multivariate Adaptive Regression Splines
- Logistic Regression
- Locally Estimated Scatter plot Smoothing
- Stepwise Regression

Logistic regression is less prone to over-fitting but it can overfit in high dimensional datasets and can't solve non-linear problems[36].

4.2 Instance Based Learning Algorithms

It is a supervised classification learning algorithm which performs operation after comparing the current instances with the previously trained instances, which have been stored in memory. Various instance based learning algorithms are:

- Learning Vector Quantization
- Self-Organizing Map
- k-Nearest Neighbor
- Locally Weighted Learning

4.3 Algorithms Using Decision Tree

In Decision Trees method, for predicting a class label for any record we start from the root of the tree.

Decision trees Algorithms are used mainly in classification problem. Here we split attributes in two or more groups by sorting basis on their values. Each tree has branches and nodes [5]. Groups attributes are represented by each node and each value represented by branch [6]. An example of decision tree is given in figure.

The most well known algorithms using decision tree are:

- Iterative Dichotomized 3
- M5
- Chi squared Automatic Interaction Detection
- C5.0 and C4.5 (different versions of a powerful approach)
- Decision Stump
- Classification and Regression Tree
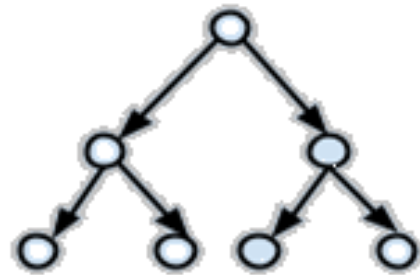- Conditional Decision Trees

Figure 2 Decision Tree Algorithm

4.4 Bayesian Algorithms

Machine Learning is one of the prominent multidisciplinary fields of Computer Science like Statistics and Algorithm. Statistics quantifies and manages the uncertainty and then represented by Bayesian algorithms which assumes that features are statistically independent that is based on probability theory and Bayes' Theorem. The most famous Bayesian algorithms are:

- Bayesian Belief Network (BBN)
- Multinomial Naive Bayes Bayesian Network (BN)
- Averaged One-Dependence Estimators (AODE)
- Gaussian Naive Bayes
- Naive Bayes

4.5 Data Clustering Algorithms

This technique that involves the grouping of data points which splits items into different types of batches. It makes group of item set into clusters where each subset share some similarity. It is unsupervised learning method, categorized as hierarchical or network clustering and partitioned clustering. The most discussed algorithms for clustering are:

- K Means
- Expectation Maximisation (EM)
- K Medians
- Hierarchical-Clustering

4.6 Learning Algorithms Using Association Rule

Learning algorithms is a rule-based machine learning method using Association rule for discovering interesting relations between variables in large databases. They are used as extraction methods, typically does not consider the order of items. It can explore or discover interesting relations between variables in large databases. The most well known learning algorithms using association rule are:

- Eclat algorithm
- Apriori algorithm

4.7 Algorithms Using Artificial Neural Network

Artificial neural networks models are inspired by neural networks basically based on the biological neuron structure and uses supervised learning. It consists of artificial neurons which have got weighted interconnections among units. They are also known parallel distributed processing networks. The most known algorithms for artificial neural network are:

- Radial Basis Function Network (RBFN)
- Back-Propagation
- Perceptron
- Hopfield Network

4.8 Deep Learning Algorithms

Deep Learning methods upgraded and enhanced the artificial neural networks that use several or multiple layers to progressively extract higher level features from the available raw input. They are concerned with building much larger and more complex neural networks. The most popular algorithms for deep learning are:

- Deep Belief Networks
- Stacked Auto Encoders
- Deep Boltzmann Machine (DBM)
- Convolution Neural Network (CNN)

4.9 Algorithms Using Dimensionality Reduction

Dimensionality reduction is a method of reducing the dimension of your feature set so widely used in case of large number of dimensions, large volume of space concerned. This algorithm requires a statistical significance. Dimensionality reduction process or method used for minimizing the number of dimensions outlined the item and eradicates unrelated and unessential data which lessen and reduce the computational cost. Few of these methods are used in classifying and regression. Well known algorithms using reduction in dimensionality are as follows:

- Partial Least Squares Regression
- Multidimensional Scaling
- Principal Component Analysis
- Flexible Discriminant Analysis
- Mixture Discriminant Analysis
- Sammon Mapping
- Projection Pursuit
- Linear Discriminant Analysis
- Principal Component Regression
- Quadratic Discriminant Analysis

4.10 Ensemble Algorithms

These methods are based on unsupervised Learning. It groups the teaching data into various types of classes of data. Self-supporting models for learning gets built for those groups. To make accurate hypothesis all learning models are combined. The popular ensemble algorithms are:

- Gradient Boosting Machines
- Boosting
- Gradient Boosted Regression Trees
- Bagging
- Bootstrapped Aggregation
- Stacked Generalization (blending)
- AdaBoost
- Random Forest

5. Objectives of Study

The development of machine learning technique is very fast now. Its usage has spread to various fields, such as learning machines currently used in medical science, pharmacology, agriculture, archeology, games, business and so forth. Many researches has been performed to create a more intelligent machines that can replace or relieve human tasks such as analyzing, communicating, learning, or making decisions. In this research performed a systematic review of research from 1998 to 2019 in the literature about the use of the machine learning technique. The purpose of this study is to determine the techniques and problems in the use of machine learning that may be used as a reference for conducting research in the future. Few of the major objectives are:

- Exploring and applying machine learning algorithms to datasets that are too large to fit into memory is pretty common.
- Analyze the problem of building scalable machine learning solutions.
- Discover how machine learning algorithms work particularly that the data preparation for logistic regression is much like linear regression.
- Analyzing the amount of research on machine learning increasingly from year to year and Growth in the total of research on machine learning techniques,
- Identify the strengths and weaknesses of the machine learning techniques.
- How can reduce computational complexity of machine learning algorithms.
- Finding that how making predictions using logistic regression is so easy that you can do it in excel.

IV. Review of Literature

The method of this research is performed by adapting the systematic literature review procedure including planning, conducting and reporting the review [28].

Teng Xiuyi (2018) [27] summarizes and analyzes machine learning technology. It also discusses their advantages and disadvantages in data mining. This study describes that machine learning technology in data mining has been applied in various industries, including financial industry, retail industry, insurance industry, telecommunication industry and so on.

Rob Law (1998) [7] applies neural networks to forecasts occupancy rates for the rooms of Hong Kong hotels and finds that neural networks outperforms naive extrapolation model and also superior to multiple regression. This research studied the feasibility incorporating the neural network to predict the rate of occupancy of rooms in Hong Kong hotel industry.

Authors **Hua et al. (2006) [8]** described support vector machines approach to predict occurrences of non zero demand or load time demand of spare parts which used in petrochemical enterprise in china for inventory management. They used an integrated procedure for establishing a correlation of explanatory variables and autocorrelation of time series of demand with demand of spare parts. On

performing the comparison the performance of SVM method with this LRSVM model, Croston's model, exponential smoothing model, IFM method and Markov bootstrapping procedure, it performs best across others.

Authors **Vahidov et al. (2008) [9]** compares the methods of predicting demand in the last of a supply chain, the naive forecasting and linear regression and trend moving average with advanced machine learning methods such as neural networks and support vector machines, recurrent neural networks finds that recurrent neural networks and support vector machines show the best performance.

Wang (2007) [10] describes the machine learning method with genetic algorithm (GA)-SVR with real value GAs. The experimental findings investigates this, SVR outshines the ARIMA models and BPNN regarding the base the normalized mean square error and mean absolute percentage error.

Author **Chen et al. (2011) [11]** presents a method forecast the tourism demands that is SVR built using chaotic genetic algorithm (CGA), like SVRCGA, which overcome premature local optimum problem. This paper reveal that suggested SVRCGA model outclass other methodologies reviewed in the research paper.

Turksen et al. (2012) [12], presents next-day stock price prediction model which is based on a four layer fuzzy multi agent system (FMAS) structure. This artificial intelligence model used the coordination of intelligent agents for this task. Authors investigates that FMAS is a suitable tool for stock price prediction problems as it outperforms all previous methods.

Shahrabi et al. (2013) [13] proposed a method for estimating tourism demand which is a new combined intelligent model i.e. Modular Genetic-Fuzzy Forecasting System using a genetic fuzzy expert systems and finds that accuracy of predicting power of MGFFS is better than approaches like Classical

Time Series models, so it is suitable estimating tool in tourism demand prediction problems.

Chen Hung et al. (2014) [14] proposes forecasting model for tourists arrival of Taiwan and Hong Kong named as LLSSVR or logarithm least-squares support vector regression technologies. In combinations with fuzzy c-means (FCM) and Genetic algorithms (GA) were optimally used and indicate that method explains a better performance to other methods in terms of prediction.

Guang-Bin Huang et al. (2015) [15] explores the basic features of ELMs such as kernels , random features and random neurons, compares the performance of ELMs and shows it tend to outshine classification, support vector machine and regression applications

Wang et al. (2016) [16] proposed a novel forecasting method CMCSGM based Markov-chain grey model which used algorithm of Cuckoo search optimization to make better the performance of the Markov chain grey model. The resultant study indicates that the given model is systematic and fine than the traditional MCGM models.

Barzegar et al. (2017) [17] demonstrates model predict multi-step ahead electrical conductivity i.e. indicator of water quality which is needed for estimating the mineralization, purification and salinity of water based on wavelet extreme learning machine hybrid or WAELM models and extreme learning machine which exploiting the boosting ensemble method. The findings showed that upgrading multi WA ELM and multi WAANFIS ensemble models outshines the individual WAELM and WA ANFIS constructions.

Author **Fouilloy et al. (2018) [18]** suggested a statistical method employing machine learning model and to analyze and applied it to solar irradiation prediction working hourly. This methodology used the high, low and medium meteorological variability like Ajacio, Odeillo, Tilos. They compared model

with auto regressive moving average and multi-layer preceptor.

Makridakis et al. (2018) [19] presents Machine Learning methods to statistical time series forecasting and compared the correctness of those methods with the correctness of conventional statistical methods and found that the first one is better and out top using the both measures of accuracy. They provide the reason for the accuracy of learning models is less that of statistical models and suggested some other achievable ways.

Zhang et al. (2018) [20] suggests a design of multi kernel ELM or MKELM method for segregation of motor imagery electroencephalogram or EEG and investigate performance of kernel ELM and impacts of two different functions of kernel such as polynomial and Gaussian kernel Compares MKELM method gives greater segregation accuracy than other algorithms indicates betterment of the suggested MKELM based .

E. Christodoulou et al. (2019) [35] conducted a study; objective was to compare performance of <u>logistic regression</u> (LR) with machine learning (ML) for clinical prediction modeling in the literature. Found no evidence of superior performance of ML over LR. Improvements in methodology and reporting are needed for studies that compare modeling algorithms.

X. Zhang et al. (2019) [36] suggests that logistic regression is more accurate on datasets with fewer attributes and balanced data distribution. Logistic regression is less prone to over-fitting but it can overfit in high dimensional datasets.

Here numbers of literatures are reviewed that how Machine learning research advances are transforming our technology and tried to find out answers of few questions like: What are the basics of machine learning and the techniques of machine learning? What are the uses of machine learning techniques?

What are the strengths and weaknesses of the machine learning techniques? How machine learning technology in data mining has been applied majorly in many industries, including financial industry, retail industry, insurance industry, telecommunication industry? How designing a resource efficient machine learning methods and so on.

The machine learning algorithms are capable of managing multi-dimensional data under the dynamic environment. Despite its so many vital features, there are some challenges to overcome.

The established research areas have achieved new momentum in big data analysis. Tremendously growth in the quantity of digital data, optimization algorithms, and affordable computing resources has enabled the application of machine learning techniques for the breakthrough of artificial intelligence. For instance, large quantities of medical data are being analyzed for diagnosis and treatment by machine learning techniques like analyze the medical data and determine the patterns in biosignals. The research scope in the future on machine learning algorithms attempts to solve the following challenges:

- Designing algorithms that don't require large amounts of labeled data,
- Improving logistic regression algorithm so it may not lead to overfit.
- Developing the machine learning algorithms that can computationally scale to Big data in order to obtain the effective and efficient parameters that will be used in the analysis,
- Discovering a privacy preservation techniques for various applications and
- Designing a resource efficient machine learning methods.

V. Research Outline

Machine learning methods and algorithms have been reviewed in this paper. This paper also reviewed algorithms describing the various types of machine learning techniques, algorithms, and methodology.

Various applications of Machine learning and many tools needed for processing are also being reviewed. In the Literature review section, we studied various machine learning algorithms implemented in past years in different areas in combination with the traditional methods and studied how they outperformed the previous models. Supervised learning is applied in classification problems like face recognition, medical diagnosis, pattern recognition, character recognition, web advertising [23]. Unsupervised learning can be applied in clustering, association analysis, CRM, summarization, image compression, bioinformatics. Reinforcement learning is widely applied in game-playing and robot control [24].

In response to the presented challenges, various approaches have been developed.

Although designing entirely new algorithms would appear to be a possible solution [29], researchers have mostly preferred other methods. Many approaches have been suggested and surveys have been published on specific categories of solutions; examples include surveys on platforms for Big Data analytics [30], [31], and review of data mining with Big Data [32]. This paper reviews and organizes various proposed machine learning approaches and discusses how they address the identified challenges. Eventually this study analyzed and approached Logistic Regression in a way so if the number of observations is lesser than the number of features and Logistic Regression is being used, so it may not lead to overfit[36].

VI. Conclusion

In this study, various machine learning algorithms have been discussed, identified, and presented challenges in machine learning with Big Data, reviewed emerging machine learning approaches, and discussed how each approach is capable of addressing the identified challenges. The main advantage of using machine learning is that, once an algorithm learns efficiently what to do with data, it can do its work automatically. In the current scenario, each and every person is using machine learning knowingly or unknowingly like getting a recommended product in online shopping to updating photos on social networking sites. This paper gives an introduction to most of the popular machine learning algorithms.

The direction of future work can be focused to obtain larger data. Large data needs not only to obtain the number of samples to be used in the machine learning process. But it is also to obtain more statistical parameters as input to find a much better correlation. Currently, the need for larger data represented in big data analysis. When dealing with big data, not only pay attention to the technique used to perform the analysis but also need to pay attention to develop the feature selection technique in order to obtain the effective and efficient parameters that will be used in the analysis. Logistic regression is a simple and moderate scale machine learning technique that gets cumbersome if you have a high number of parameters in your feature. The less is parameters count there are fewer chances of overfitting but it can overfit in high dimensional datasets. This study proposes a smart algorithm where training with more data can help to reduce overfitting issue.

VII. REFERENCES

[1]. P. Louridas and C. Ebert, "Machine learning and its applications: A review", International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), vol: 33, Issue: 5, Sept.-Oct. 2016.

[2]. A. Dey, "Machine Learning Algorithms: A Review", International Journal of Computer Science and Information Technologies, Vol. 7 (3), 2016.

[3]. Teng Xiuyi and Gong Yuxia, "Research on Application of Machine Learning in Data

Mining". IOP Conf. Series: Materials Science and Engineering, 2018.

[4]. M. Praveena and V. Jaiganesh, "Literature Review on Supervised Machine Learning Algorithms and Boosting Process". International Journal of Computer Applications, ISSN No. 0975– 8887, vol. 169, 2017.

[5]. K. Das and R. N. Behera, "A Survey on Machine learning: Concept, Algorithms and Applications", International Journal of Innovative Research in Computer and communication Engineering. vol. 5, 2017.

[6]. S.B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques", Informatica. pp 249-268, 2007.

[7]. Rob Law, "Room occupancy rate forecasting: a neural network approach", International Journal of Contemporary Hospitality Management, vol. 10 Issue 6, pp 234 – 239, 1998.

[8]. Z. Hua and B. Zhang, "A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts", Applied Mathematics and Computation 181, pp 1035–1048, 2006.

[9]. R. Carbonneau, K. Laframboise and R. Vahidov, "Application of machine learning techniques for supply chain demand forecasting", European Journal of Operational Research 184, pp 1140 1154, 2008.

[10]. K.Y. Chen and C.H. Wang, "Support vector regression with genetic algorithms in forecasting tourism demand", Tourism Management 28, pp 215–226, 2007.

[11]. W. C. Hong, Y. Dong, and S.Y. Wei, "SVR with hybrid chaotic genetic algorithms for tourism demand forecasting", Applied Soft Computing 11, pp 1881– 1890, 2011.

[12]. E. Hadavandi and B. Turksen, "A Hybrid Fuzzy Intelligent Agent-Based System for Stock Price Prediction", International Journal Of Intelligent Systems, Vol. 00, pp 1–23, 2012.

[13]. J.Shahrabi, E. Hadavandi and S. Asadi, "Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series", Knowledge-Based Systems 43, pp 112–122, 2013.

[14]. K.C. Hung and K.P. Lin, "Tourism demand forecasting using novel hybrid system", Expert Systems with Applications 41, pp 3691–3702, 2014.

[15]. G.B. Huang, "An Insight into Extreme Learning Machines: Random Neurons, Random Features and Kernels", Springer, 2014.

[16]. X. Sun, W. Sun, and Y. Gao, "Using a Greye Markov model optimized by Cuckoo search algorithm to forecast the annual foreign tourist arrivals to China", Tourism Management 52, 2016.

[17]. R. Barzegar and B. O. Zielinski, "Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model", Springer, ISBN 00477-017-1394, 2017.

[18]. A. Fouilloy and C.Voyant, "Solar irradiation prediction with machine learning: Forecasting models selection method depending on weather variability", Energy 165, 2018.

[19]. Spyros Makridakis, Evangelos Spiliotis and Vassilios Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward", PLoS ONE 13, 2018.

[20]. Yu Zhang, Yu Wang, and A. Cichocki, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces", Expert Systems With Applications 96, pp 302– 310, 2018.

[21]. P. Domingos, "A Few useeful Things to Know about Machin Learning", 2012.

[22]. Y. Singh, P. K. Bhatia and O.P Sangwan, "A review of studies in machine learning

technique". International Journal of Computer Science and Security, vol.1, pp 70 – 84, 2007.

[23]. Petersp, "The Need for Machine Learning is everywhere", March 2015.

[24]. J. Brownlee, "A Tour of Machine Learning algorithms" November 25, 2013.

[25]. J. Brownlee, "Machine Learning Tools", December 28, 2015.

[26]. T. Ayodele, "Types of Machine Learning algorithms", 2010.

[27]. T. Xiuyi1 and G. Yuxia1, Research on Application of Machine Learning in Data Mining, IOP Conf. Series: Materials Science and Engineering 392 (2018).

[28]. B.A. Kitchenham, "Guidelines for performing systematic literature review in software engineering", Technical report EBSE-2007-001, UK, 2007.

[29]. I. H. Witten, E. Frank, M. A. Hall and C. J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, San Mateo, CA:Morgan Kaufmann, 2016.

[30]. D. Singh and C. K. Reddy, "A survey on platforms for big data analytics", J. Big Data, vol. 2, no. 1, pp. 1-20, 2015.

[31]. P. D. C. de Almeida and J. Bernardino, "Big data open source platforms", Proc. IEEE Int. Congr. Big Data, pp. 268-275, Jun. 2015.

[32]. X. Wu, X. Zhu, G.-Q. Wu and W. Ding, "Data mining with big data", IEEE Trans. Knowl. Data Eng., vol. 26, no. 1, pp. 97-107, Jan. 2014.

[33]. Z. Yang and D. Li, "Application of Logistic Regression with Filter in Data Classification", Chinese Control Conference (CCC), 8865281, 30 July 2019.

[34]. T. Rymarczyk, E. Kozłowski, G. Kłosowski, and Konrad Niderla1, "Logistic Regression for Machine Learning in Process Tomography", Sensors. doi: 10.3390/s19153400. Aug 2019.

[35]. E. Christodoulou, J. Ma and E. W. Steyerberg, "A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models Auth", Journal of Clinical Epidemiology, Vol. 110, June 2019, Pages 12-22

[36]. X. Zhang, Di Wang, and Y. Yang, "Prediction accuracy analysis with logistic regression and CART decision tree", Fourth International Workshop on Pattern Recognition, vol. 11198 Nanjing-China,31 July 2019

Cite this article as :