International Journal of Engineering Research and Advanced Technology (IJERAT)

DOI: 10.31695/IJERAT.2020.3607

E-ISSN : 2454-6135

Volume.6, Issue 8
August -2020

Machine Learning Algorithms, models and Applications: A Review

[1]Suleiman Saadon Fawzy and [2] Nada Hassan Jasem.

[1]Department of Computer center, I.T, Al-Mustansiriyah University, Baghdad, Iraq

[2]Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq

ABSTRACT

Machine Learning (ML) can be defined as unfolding from AI, also it is specified as a field related to the computer sciences. Also, ML is specified as one of the multi-disciplinary fields, a combination related to the statistics as well as the algorithms of computer sciences that has been majorly utilized in classification and predictive analyses. The other section of this study concentrating on the impact related to major ML algorithms and approaches. Furthermore, the presented study will specify different ML tools required to run the projects of ML. The major concern related to the work has been studying the major methods as well as the case studies related to utilizing ML with regard to forecasting in various areas like supply chain demand, tourism demand forecasting, stock price forecasting, solar irradiation forecasting as well as consideration related to ML and NN approaches.

Keywords: Supervised Learning, Un-Supervised Learning, Machine Learning, Support Vector Machine.

1. INTRODUCTION

In the last decades, the stream of AI has been general and significant field with regard to computer sciences since it is preparing the machines for performing the tasks which the individuals might be doing, also has the task of training the computers for solving real-world problems with highest rate of success. Since noticing the scientific growth as well as the advancements in the systems of AI come with the ability of learning and improving via past experience with no assistance code in the case when being exposed to the new data. Finally, it will be leading to a technology related to ML which applies learning algorithms for learning from provided data [1]. Also, the ML applies the methods of data mining for extracting information from data-sets of huge size. Data mining and ML methods exploring the data from end to end for finding hidden patterns in the data-set [2]. Data mining and ML algorithms were used in a lot of fields like electric load forecasting, finance forecasting, computer networking, telecommunication industry, tourism and travel industry, etc. [2].

2. LITERATURE REVIEW

A study conducted by **Rob Law (1998) [10]** applied NNs for forecasting the occupancy rates with regard to the rooms in the hotels of Hong Kong as well as finding that the NNs are outperforming the naïve extrapolation model, also outperforming multiple regression. Such studies examined the possibility to incorporate the NNs for predicting the occupancy rates of the rooms in the hotels of Hong Kong.

A study conducted by **Hua et al. (2006) [11]** specified the approach of SVM for predicting the occurrence related to the non-zero demands or the load time demands related to spare parts that is applied in the petrochemical enterprises in Chine for the inventory management. Also, they utilized integrated approach to establish correlations related to the explanatory variables as well as the auto-correlation regarding time series of the demands with the demands regarding the spare parts. With regard to achieving comparisons related to the SVM's performance with LRSVM model, Markov bootstrapping process, exponential smoothing model, Creston's model, and IFM approach, it is outperforming the others.

A study conducted by **Vahidov et al. (2008) [12]** compare the approached related to the prediction of the demands in last supply chain, naïve forecasting as well as the linear regression, also the trend moving average with the developed ML approaches like SVM, NN, the recurrent NNs finding that the recurrent NNs and SVMs showing the better performance.

A study conducted by **Wang (2007) [13]** describe ML approach with the genetic algorithm (GA)-SVR in addition to the real value GA. Also, experimental studied examining the fact that SVR outshining ARIMA models as well as the BPNN related to base normalized mean square errors in addition to the mean absolute percentage errors.

A study conducted by **Chen et al. (2011) [14]** presented an approach for forecasting the demands of tourism which has been SVR-built with the use of chaotic genetic algorithm (CGA), such as SVRCGA, that is overcoming the premature local optimum

problem. The presented study reveals that the supposed SVRCGA model is outperforming the other approaches indicated in the current study.

A study conducted by **Turksen et al. (2012) [15],** presented a model for the next-day stock price prediction that has been on the basis of 4-layer fuzzy multi agent system (FMAS) structures. Such model utilized the coordination regarding the intelligent agents with regard to such task. The researchers studying that the FMAS has been adequate tool for the problems of the stock price prediction since it is outperforming the previous approaches.

Shahrabi et al. (2013) [16] suggested an approach to estimate the tourism demands that is the new combined intelligent model, for instance the Modular Genetic-Fuzzy Forecasting System with the use of genetic fuzzy expert systems as well as finding that the precision related to predicting power regarding MGFFS has been better in comparison to the methods such as models of the Classical Time Series, thus it is of high importance in evaluating the problems of the tourism demand predictions.

A study conducted by **Chen Hung et al. (2014) [17]** suggested forecasting model with regard to the arrival of tourists in Hong Kong and Taiwan referred to as LLSSVR or the logarithm least-squares support vector regression approaches. Along with the fuzzy c-means (FCM) as well as the Genetic algorithms (GA) have been effectively utilized and indicating that the approach explains better performance in comparison to the other approaches with regard to prediction.

A study conducted by **Guang-Bin Huang et al. (2015) [18]** specified the major features related to ELMs like random neurons, random features, and kernels, comparing efficiency which is related to ELMs and showing its tendency for outshining the classifications, SVMs as well as the regression applications.

A study conducted by **Wang et al. (2016) [19]** suggested new forecasting approach CMCSGM based Markov-chain gray model that utilized the Cuckoo search optimization for making better performance regarding Markov chain grey model. Also, the resulting work indicated that the provided model has been in comparison to the conventional MCGM models.

A study conducted by **Barzegar et al. (2017) [20]** showing that the model predicts the multistep ahead electrical conductivity, for instance, indicator related to the quality of water that has been required to estimate salinity, mineralization, as well as purifications of the water on the basis of the wavelet extreme learning machine hybrid or the WAELM models as well as the extreme learning machine that exploit boosting ensemble approach. The results indicated that upgrading multi WA ELM in addition to the multi WAANFIS ensembles the models outperforming individual WAELM as well as the WA ANFIS constructions.

A study conducted by **Fouilloy et al. (2018) [21]** provided a statistical approach using ML model, also analyzing and applying it to the solar irradiation prediction functioning hourly. Such approach utilized medium, low, and high meteorological variability such as Tilos, Odeillo, and Ajacio. They performed the comparison of the model to the auto regressive moving average in addition to the multi-layer preceptor.

A study conducted by **Makridakis et al. (2018) [22]** presented ML approaches for the statistical time series forecasting, also put to comparison the correctness regarding such approaches with the correctness related to traditional statistical approaches and indicating that the first has been more efficient and outperforming the measures of the accuracy. Also, they provided the reason related to accuracy of the learning models has been less in comparison to statistical models and indicated certain approaches.

A study conducted by **Zhang et al. (2018) [23]** suggested a design which is related to the multi kernel ELM (MKELM) approach for the segregation which is related to the motor imagery electro-encephalogram (EEG) and examining the performance related to kernel ELM, also the effect related to 2 distinctive functions of the kernel like polynomial as well as the Gaussian kernel in comparison to the MKELM approach providing high segregation accuracy that the other algorithms indicating the efficiency of the indicated MKELM-based .

3. THEORY

The main aim of the learner has been generalizing from the experience.[3][4] Furthermore, the generalization can be considered as the capability of the learning machine for accurately performing on unseen, new tasks/examples following having experienced learning dataset. Also, the examples of training are coming from unrecognized probability distributions (specified representative regarding occurrence's space), also the learner must be building a model related to such space which is enabling it for producing precise predictions in the novel conditions.

The computational analysis that is related to the algorithms of ML in addition to their performance has been considered as a field related to theoretical computer sciences referred to as the computational learning theory. Due to the fact that the training sets have been finite, also the future has been uncertain, the learning theory generally doesn't provide grantee regarding the algorithm's performance. However, the probabilistic bounds related to the performance have been common. Also, bias-variance decompositions have been an approach for quantifying the generalization error.

With regard to optimum performances in terms of generalization, hypothesis's complexity must be matching the complexity related to the function underlying data. In the case when the hypothesis has less complexity in comparison to the function, the model will be under fitting the data. In the case when model's complexity has been subjected to increase, the training error will be

DOI : 10.31695/IJERAT.2020.3607

decreased. Yet, in the case when hypothesis has been extremely complex, the model will be subjected to overfitting, thus the generalization is going to be lesser.[5]

Along with performance bounds, the learning theorists have been studying time complexity as well as the feasibility related to learning. With regard to the computational learning theories, computation has been specified for being possible in the case when it might be achieved in polynomial time. There have been 2 types related to time complexity results. Furthermore, the positive results showing that specific class related to the functions might be learned in the polynomial time. Also, the negative results showing that specific classes might not be learned in the polynomial time.

4. APPROACHES

4.1 Learning Algorithm Types

Recently, a lot of algorithms related to ML has been provided. Just a few of them have the ability of solving the problem, thus they might be replaced via another one [6]. There have been 3 algorithms of ML (supervised learning, reinforcement learning, and un-supervised learning), that have been showing in the figure1.

4.1.1 Supervised learning: These algorithms include certain set related to the input variables (training data) that have been prelabelled as well as the target data [7]. Utilizing input variables, it might be generating mapping function for the purpose of mapping inputs to the needed outputs. The procedures of parameter adjustment will continue till the system show good accuracy extent related to teaching data.

Figure 1: Learning Algorithm Types

4.1.2. Un-supervised learning: This **algorithm** has just the training data instead of the outcome data. Also, the input data isn't formerly labeled. It has been utilized in the classifiers through specifying the current patterns or the clusters in input data-sets [8].

41.3 Reinforcement learning: This machine in this algorithm has been trained for mapping the action to certain decision, therefore feedback or reward signals will be created. The machine train itself for finding major rewarding actions through reward as well as punishment utilizing former experiences

5. MODELS

Regression Algorithms

This is considered as part related to the predictive analytics, also exploiting co-relations between the **independent** and the **dependent** (target) variables. The major models of regression have been: Locally Estimated Scatterplot Smoothing (LOESS), Ordinary Least Squares Regression (OLSR), Linear Regression, Stepwise Regression, Logistic Regression, Multivariate Adaptive Regression Splines (MARS), and so on[9], that have been showing in the figure2..

DOI : 10.31695/IJERAT.2020.3607

Figure 2: Linear Regression, Logistic Regression Algorithms

Instance-based Algorithms

Memory-based or instance-based learning model will be storing instances related to training data rather than accurate definition related to the target function. In the case when new example or problem has been faced, it will be studied along with stored instance for the purpose of predicting or determining target function value. Also, it might be replacing stored instances via new instances in the case when it is better fit in comparison to the former. Thus, they have been recognized as winner-take-it-all technique, examples of such approach are Locally Weighted Learning (LWL), Learning Vector Quantization (LVQ), K-Nearest Neighbors (KNN) that have been showing in the figure3 , Self-Organizing Map (SOM), and so on.

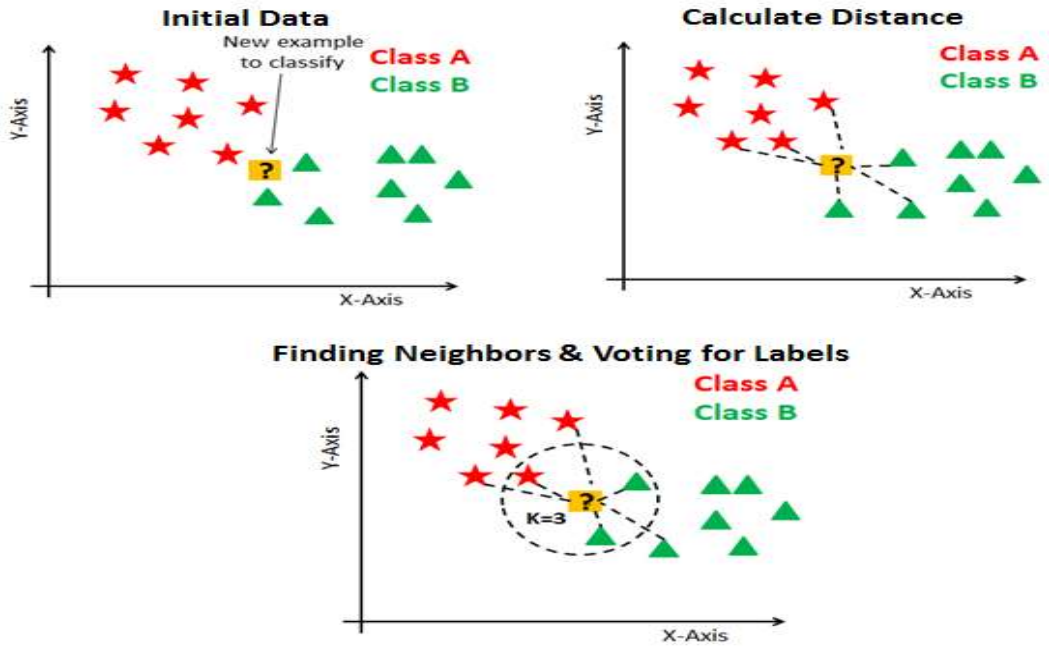

Figure 3: K-Nearest Neighbors (KNN) Algorithms

Regularization Algorithm

This algorithm is defined as the process related to counteracting overfitting or even abating outliers, also it is specified for being important and effective modification which has been augmented with the other current models of ML, generally the regressive Models. It is smoothing the regression line through castigating any bent regarding the curve which attempt on matching outliers. The examples of such algorithms are Least-Angle Regression (LARS), Least Absolute Shrinkage and Selection Operator (LASSO), Ridge Regression, Elastic Net, and so on.

Decision Tree Algorithms

These algorithms constructing tree like structures including possible solution to problems on the basis of certain constraints. Thus, it begins with root or single simple decision, that fork off in to some branches till prediction or decision has been carried out, they have been desired for their capability in formalizing the problem which will allow recognizing possible solutions with more

DOI : 10.31695/IJERAT.2020.3607

accuracy and speed in comparison to the others. The examples of such algorithms are Conditional Decision Trees show fig 4, C5.0, C4.5, Iterative Dichostomoiser 3 (ID3), Classification and Regression Tree (CART), and so on.

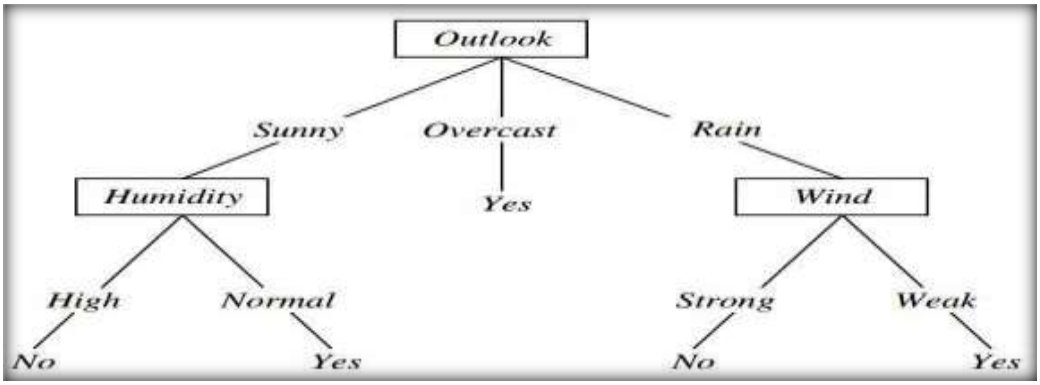

Figure (4): example of Conditional Decision Trees.

5. BAYESIAN ALGORITHMS

Some algorithms of ML are using the Bayes' Theorem for solving regression and classification problems. The examples of these algorithms have been Bayes Network (BN), Averaged One-Dependence Estimators (AODE), Naive Bayes, Multinomial Naive Bayes, Gaussian Naive Bayes, Bayesian Belief Network (BBN), and so on.

6. SUPPORT VECTOR MACHINE (SVM)

This is considered as significant approach of ML which might be group of its own. Also, it applies the decision plane or the separating hyperplane for demarcating the decision boundaries between set of the data points that have been classified with the various labels. Furthermore, it is supervised classification algorithm. Put differently, the algorithm will be developing optimum hyperplane using the training data or the input data and such decision plane will be categorizing new examples. According to used kernel, the SVM might be performing non-linear and linear classification show in fig 5.

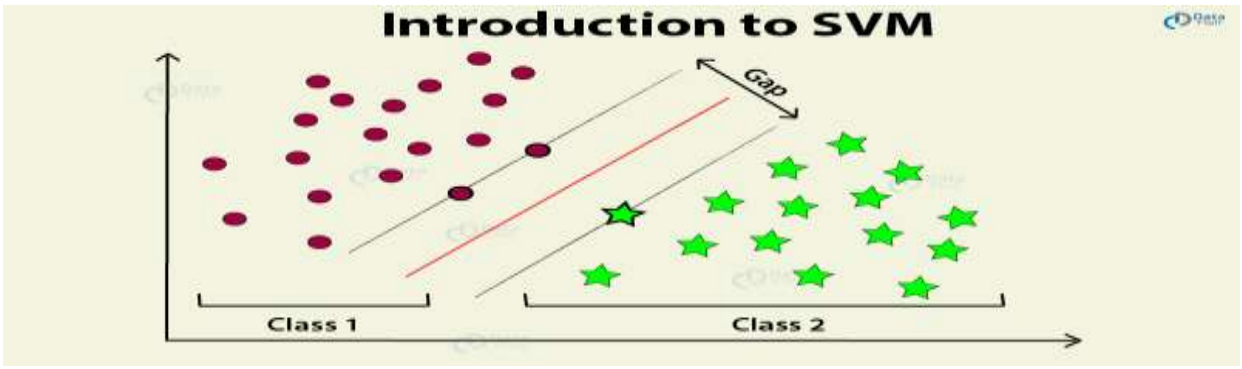

Figure 5: the SVM classification.

7. CLUSTERING ALGORITHMS

These algorithms are concerned with applying the ingrained pattern in the data-sets for classifying and labelling the data. The examples have been Expectation Maximization (EM), Gaussian Mixtures, Agglomerative clustering, Spectral Clustering, K-Medians, K-Means show in fig 6, Affinity Propagation, Ward hierarchical clustering, DBSCAN, Birch, Mean Shift, and so on.

DOI : 10.31695/IJERAT.2020.3607

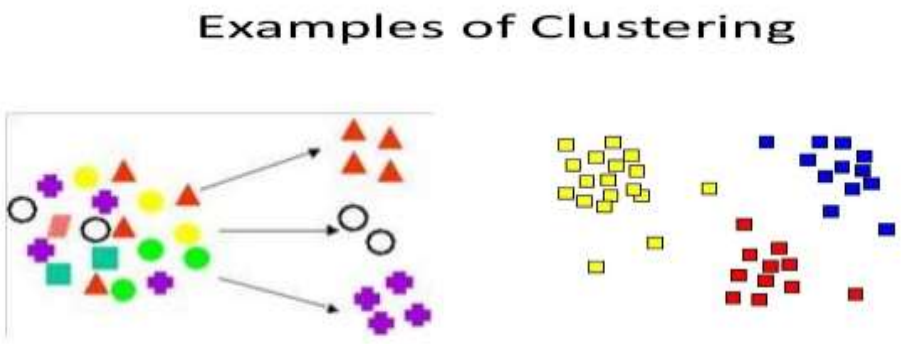

Figure 6: K-Means Clustering.

8. ASSOCIATION RULE LEARNING ALGORITHMS

These algorithms are helping in discovering the relations between seemingly unrelated data. They have been generally utilized via the ecommerce websites for predicting the behaviors of customers and the future requirements for promoting specific appealing products. The examples of such algorithms have been Eclat Algorithms, Apriori algorithm, and so on.

9. ARTIFICIAL NEURAL NETWORK (ANN) ALGORITHMS

This is considered as a model on the basis of operation and built related to actual NNs of the animals or humans. Also, the ANNs have been specified as non-linear models since they are trying for discovering the complex relations between output and input data. Yet, it is drawing samples from the data instead of specifying the whole set, thus decreasing time and costs. The examples of such algorithms are Back-Propagation, Perceptron, Radial Basis Function Network, Hopfield Network, and so on, show in fig 7.

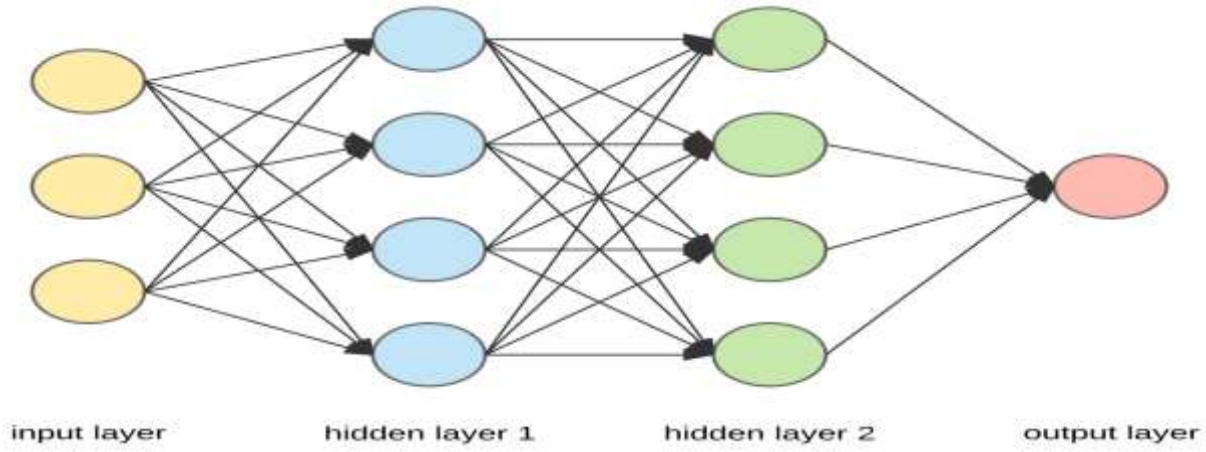

Figure 7: Artificial Neural Network.

10. DEEP LEARNING ALGORITHMS

These algorithms are considered as modernized versions related to the ANNs which capitalize on profuse supply related to the data. They have been utilizing larger NNs for solving the semi-supervised problems in which the main portion regarding abound data has not be classified or not labeled. The examples of these algorithms have been Stacked Auto-Encoders, Conventional Neural Networks (CNNs), Deep Belief Network (DBN), Deep Boltzmann Machine (DBM)show in fig 8 , and so on.

DOI : 10.31695/IJERAT.2020.3607

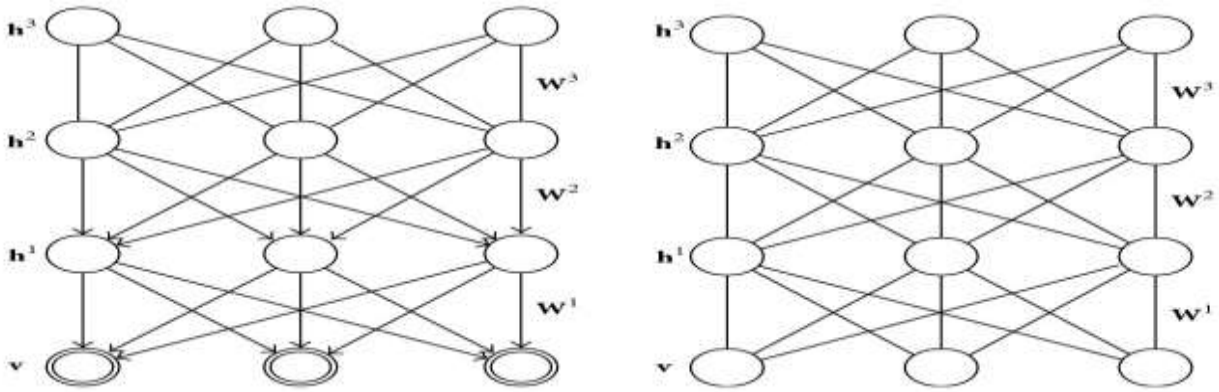

Figure 8: Deep Boltzmann Machine.

11. DIMENSIONALITY REDUCTION ALGORITHMS

These algorithms have been majorly used for reducing large datasets, thus they are major discriminative components for containing relative information as well as describing it with few features. This will be providing adequate visualization related to the data with various features or related to elevated dimensionality and help in achieving supervised classification with more efficiency. The examples of such algorithms have been Mixture Discriminant Analysis, Flexible Discriminant Analysis, Projection Pursuit, Sammon Mapping, Principal Component Regression, Principal Component Analysis, Partial Least Squares Regression, Multidimensional Scaling (MDS), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and so on.

12. ENSEMBLE ALGORITHMS

The major aim of these algorithms has been integrating projections related to some weaker estimators which have been singly trained for the purpose of boosting or enhancing the generalizability or the robustness over single estimator. The types related to learners as well as the means for incorporating them has been adequately selected for maximizing the accuracy. The examples of these algorithms are Extremely Randomized Trees, Gradient Boosted Regression Trees, Stacked Generalization, AdaBoost, Boosting, Bootstrapped Aggregation (Bagging), Gradient Boosting Machines (GBM), Random Forest, and so on show fig 9 .

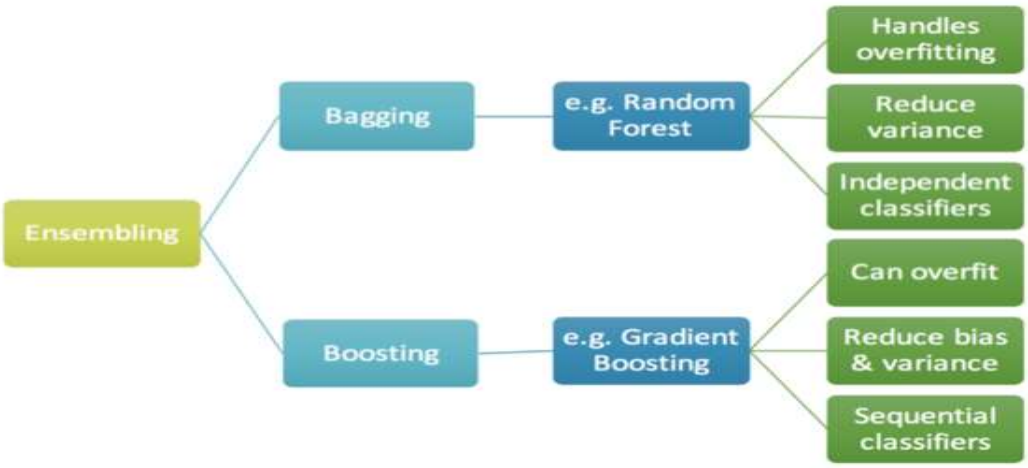

Figure 9: The types Ensemble Algorithms.

13. APPLICATIONS of ML

There are a lot of applications related to ML [24]:

Significant sign related to the ML's development is its vital real-life applications, a few of them will be described in the next sections. It must be indicated that until the year 1985, there has been no significant commercial applications related to the algorithms of ML.

DOI : 10.31695/IJERAT.2020.3607

A. Speech Recognition

All the existing systems of speech recognition provided in the market apply the approaches of ML for training the system to have better accuracy. The majority of these systems implementing learning in 2 different stages: the pre-shipping speaker independent training as well as the post-shipping speaker-dependent training.

B. Computer Vision.

Most of the current vision systems, for example, the facial recognition software, the system has the ability for automatic classification microscopic images regarding the cells, using the approaches of ML to have better accuracy. For instance, US Post Office applies a system of computer visions with the handwriting analyzer, therefore trained for sorting the letters with handwritten addresses automatically in addition to level of accuracy might be (85%_.

C. Bio-Surveillance

A lot of the government initiatives for tracking possible disease outbreaks applies the methods of ML. Considering RODS project in the west of Pennsylvania. Also, such project will be collecting reports of admission to emergency rooms in hospitals, ML software systems will be trained with the use of profiles related to the admitted patients for the purpose of detecting the aberrant symptoms, their patterns as well as the areal distribution. Researches are made for incorporating certain data in the system, such as the purchase history of over-the counter medicine for providing additional training data. The complexity related to such type of complex as well as dynamic datasets might be effectively handled with the use of just automated learning approaches.

D. Robot or Automation Control

The techniques of ML have been mainly utilized in automated and robot systems. For instance, utilizing ML for obtaining control tactics with regard to stable flights as well as aerobatics related to helicopter. Also, the self-driving cars that are created via Google utilize ML for being trained from the obtained terrain data.

E. Empirical Science Experiments

Major group data-intensive science discipline applies the approaches of ML in many studies. For instance, ML has been used in genetics, for identifying the rare celestial objects in astronomy, also in the Neuroscience as well as the psychological analysis. Furthermore, other significant application of ML includes topic identification, fraud detection, spam filtering as well as predictive analytics (market survey, weather forecast, stock market prediction, and so on).

14. TOOLS UTILIZED IN ML

The tools are making ML rapid and swift, also they are providing interface to ML machine learning programming language. The tools of ML are providing effective practices for implementation and process [25]. The tools of ML contain platforms that are providing ability to run project or module. The examples of platforms related to ML are as follows:

- **Python SciPy sub-parts like scikit-learn, Panda** R **Platform.**
- **WEKA ML Workbench.**

The tools of ML include a lot of libraries that are providing all the ability for completing projects as well as the libraries providing different algorithms. Examples of the libraries are as follows:

- **Accord Framework in .NET**
- **Scikit-learn in Python**
- **JSAT in Java.**

15. CONCLUSION

The algorithms and techniques of ML were indicated in the presented study, also this study specified the algorithms as well as the many types related to ML algorithms and methods. A lot of applications related to ML and various tools required for processing have been indicated. With regard to the section of literature review, the study examined a lot of ML algorithms that have been carried out before in a lot of areas along with conventional techniques, also they examined how they surpassed the previous models.

REFERENCES

1. Mariette Awad, Rahul Khanna. “Efficient Learning machines: Concepts and Applications”. As of press Publishers, (2015).

DOI : 10.31695/IJERAT.2020.3607

2. Teng Xiuyi1, Gong Yuxia1. "Research on Application of Machine Learning in Data Mining". IOP Conf. Series: Materials Science and Engineering, 2018.

3. Bishop, C. M., Pattern Recognition and Machine Learning, Springer, (2006).

4. Mohri, Mehryar; Rostamizadeh, Afshin; Talwalkar, AmeetFoundations of Machine Learning. USA, Massachusetts: MIT Press. ISBN 9780262018258, (2012).

5. Alpaydin, Ethem Introduction to Machine Learning. London: The MIT Press. ISBN 978-0-262-01243-0. Retrieved 4 February 2017, (2010).

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

7. S.B. Kotsiantis."Supervised Machine Learning: A Review of Classification Techniques", Informatica. pp 249-268, 2007. 6. 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. Kajaree Das, Rabi Narayan Behera. "A Survey on Machine learning: Concept, Algorithms and Applications", International Journal of Innovative Research in Computer and Communication Engineering. vol. 5, (2017).

9. A. Smola and S. Vishwanathan," Introduction To Machine Learning". The United Kingdom at the University Press, Cambridge, October 1,(2010).

10. Zhongsheng Hua , Bin 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).

11. Real Carbonneau, Kevin Laframboise, Rustam Vahidov, "Application of machine learning techniques for supply chain demand forecasting ", European Journal of Operational Research 184, pp 1140 1154,(2008).

12. Kuan-Yu Chen, Cheng-Hua Wang, "Support vector regression with genetic algorithms in forecasting tourism demand" , Tourism Management 28, pp 215–226,(2007).

13. Wei-Chiang Hong, Yucheng Dong, Li-Yueh Chen, Shih-Yung Wei, " SVR with hybrid chaotic genetic algorithms for tourism demand forecasting", Applied Soft Computing 11, pp 1881– 1890, (2011).

14. M. H. Fazel Zarandi, Esmaeil Hadavandi, B. Turksen, "A Hybrid Fuzzy Intelligent Agent-Based System for Stock Price Prediction", International Journal Of Intelligent Systems, Vol. 00, pp 1–23,(2012).

15. Jamal Shahrabi, Esmaeil Hadavandi, Shahrokh Asadi, "Developing a hybrid intelligent model for forecasting problems: A case study of tourism demand time series", Knowledge-Based Systems 43, pp 112–122,(2013).

16. Ping-Feng Pai, Kuo-Chen Hung, Kuo-Ping Lin, "Tourism demand forecasting using the novel hybrid system", Expert Systems with Applications 41, pp 3691–3702, (2014).

17. Guang-Bin Huang, "An Insight into Extreme Learning Machines: Random Neurons, Random Features, and Kernels", Springer, (2014).

18. Xu Sun, Wangshu Sun, Jianzhou Wang, Yixin Zhang, Yining Gao," Using a Greye Markov model optimized by Cuckoo search algorithm to forecast the annual foreign tourist arrivals to China", Tourism Management 52,(2016).

19. Rahim Barzegar, Asghar Asghari Moghaddam, Jan Adamowski, Bogdan Ozga-Zielinski," Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model", Springer, ISBN 00477-017-1394,(2017).

20. Alexis Fouilloy, Cyril Voyant, Gilles Notton, Fabrice Motte, Christophe Paoli, Marie-Laure Nivet,, Emmanuel Guillot, JeanLaurent Duchaud," Solar irradiation prediction with machine learning: Forecasting models selection method depending on whether variability", Energy 165,(2018).

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

22. Yu Zhang, Yu Wang, Guoxu Zhou, Jing Jin, Bei Wang, Xingyu Wang, Andrzej Cichocki," Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces", Expert Systems With Applications 96, pp 302–310,(2018).

23. Pedro Domingos, "A Few Useful Things to Know about Machin Learning"(2012).

24. Peters," The Need for Machine Learning is Everywhere" March 10,(2015).

25. Jason Brownlee," A Tour of Machine Learning algorithms" November 25, (2013).

DOI : 10.31695/IJERAT.2020.3607