International Journal of Advanced Technology in Engineering and Science
Vol. No. 10, Issue No. 04, April 2022
www.ijates.com
ijates
ISSN 2348 - 7550

Machine Learning: Review of systematic Approaches

Vaibhav Narkhede[1], Dr. Amit Bhusari[2],

Dr. Pradeep Jawandhiya[3]

[1]*Dept. of Comp. Sci. & Engg, Pankaj Laddhad Institute of Tech. Buldana , India*

[2] *Dept. of MCA, Trinity Academy of Engineering, Pune, India*

[3] *Dept. of Comp. Sci. & Engg, Pankaj Laddhad Institute of Tech. Buldana , India*

Abstract

In this 21st century, the world is evolving with the ease of automation in the every field of work whose backbone is the Artificial Intelligence. Machine Learning is the art of AI that provides ability to computer to think and provides the decision/prediction on the basis of analysis done on data provided or by gaining the previous experience. ML has huge impact on the data driven techniques and its applications becomes very important in day to day life. Machine Learning is further categorized into three types mainly as Supervised, Unsupervised and Reinforcement Learning. In this paper, we are going to review different machine learning approaches along with their pros and cons.

Keywords: *Machine Learning, Supervised Learning ,Data Classification ,Predictive analysis*

I. INTRODUCTION:

ML is a kind of Artificial Intelligence which provides the computers the ability to take decision by analyzing the input data provided to it. A computer is assigned with some type of task and the computer is able to generate more accurate results by repeating task again and again. This provides the fruitful experience to machine which helps to predict the accurate results[1].Use of data to gain the pattern in the data helps the machine to enhance the future values by predicting the outcomes is the key for every type of machine learning algorithms. ML has wide range of applications in every field today. viz. Healthcare, Agriculture, Education, Automobile Industry, Surveillance Systems etc[2]. ML is further classified into Supervised , Unsupervised and Reinforcement Learning mainly.

Supervised Learning is the ML technique where computer is supervised by number of inputs and outputs. Within this the computer is going to utilize the knowledge within the data and tries to predict the new label.ML algorithm compares the actual values with the predicted values to evaluate the model based on accuracy and score.

Unsupervised learning is the ML technique where computer is governed by number of inputs based on self learning. Computer has to find the hidden pattern in the data based on non labelled data feeded to the model. The system is going to use the observations from the data to generate the results.

Reinforcement Learning is the ML technique where computer takes the decisions on the previous experience to maximize the rewards. Here the agent is going to interact with the environment and takes the action to gain the reward on the basis of trial and error.

ML is also divided into semi supervised learning approach where computer is supervised with small amount of the featured data and unfeatured data largely. These systems are highly adapts themselves to attain the high accuracy[3]. More precisely the machine learning is based upon the processing of the large amount of data and investigate it to find the future results. Supervised Techniques are the most popular techniques used for training

the machine but other two are also important when the data is non classified.

II. METHODOLOGY USED BY ML:

The ML consist of the processing of the huge amount of the data to gain the knowledge by the machine which can be used to predict the future values. So the training process involves the sequence of steps as mentioned in the figure1. The implementation of the ML algorithm either in python or R involves this number of steps evertytime. The phases of ML technique actually begin with the problem identification and data gathering to solve the problem[2], [4].

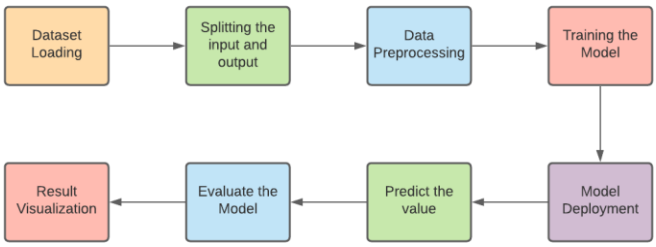

Figure 1 : Phases of Machine Learning technique

1) Dataset Loading: This phase deals with the loading of the dataset and preparing the input data for the preprocessing.
2) Splitting Input and Output: within this phase the data is divided into the input data and output data.
3) Data Preprocessing : This phase consist of the cleaning of the data where the one has to take care of missing values, categorical features in the dataset which are the important for maximizing the accuracy of ML model. The imbalanced data is also handled here.
4) Training the Model: The input and output is divided into training and testing data. Here Training data is used to train the machine to gain the knowledge and testing data is used to test the model by comparing the predicted values with the actual values.
5) Model Deployment: A ML model is used to fit to the input data along with some parameters as required by some algorithms.
6) Predict the value: This phase involves the prediction on the test values as well as more importantly to predict the new values.
7) Evaluate the Model: The model performance is evaluated based on the accuracy using the confusion matrix , the score and accuracy .
8) Result Visualization: Lastly the results are visualized with the help of drawing the graphs which shows the relation between the input parameters and the output parameters.

III. ML APPROACHES : Following are the approaches used for machine learning which can be summarized as follows.

1) Supervised learning:

The supervised learning algorithm creates a mathematical model of a dataset that contains both inputs and desired outputs. This data, called training data, consists of a series of training samples. Each training example has one or more inputs and a desired output, also known as a monitor signal. In a mathematical model, each training example is represented by an array or vector, sometimes referred to as a feature vector, and the training data is represented by a matrix. By repeatedly optimizing the objective function, the supervised learning algorithm learns a function that can be used to predict the output associated with a new input. Optimal performance allows the algorithm to correctly determine the output of inputs that were not part of the training data. Algorithms that improve the accuracy of output or prediction over time are said to have learned to perform this task.

2) Unsupervised Learning:

Unlike supervised learning, unsupervised learning does not label, classify, or classify input datasets. Mathematical models try to identify the similarity of the dataset and infer the structure present in the input data. Unsupervised learning problems can be further grouped into clustering and association problems. The clustering problem tries to find the groupings in the dataset, while the association problem tries to generalize the rules that describe most of the dataset. The most widely used learning algorithms in unsupervised learning are kmeans, neural networks, linear discriminant analysis, principal component analysis, and pre-algorithms.

3) Semi-Supervised Learning

In semi-supervised learning, the input dataset is a mixture of labeled and unlabeled data. Data sets typically contain a small amount of labeled data and a large amount of unlabeled data. Mathematical models use labeled data to learn the structure of unlabeled data and attempt to make predictions. Semi- supervised learning problems can also be further grouped into classification and regression problems.

4) Reinforcement learning

Reinforcement learning is an area of machine learning that explores how software agents need to behave in an environment to maximize the concept of cumulative rewards. Due to its generality, this area has been studied in many other areas, including: B. Game theory, control theory, operations research, information theory, simulation-based optimization, multi-agent systems, group intelligence, statistics, and genetic algorithms. In machine learning, the environment is usually represented as a Markov decision process (MDP). Many reinforcement learning algorithms use dynamic programming. Reinforcement learning algorithms do not require knowledge of the exact mathematical model of MDP and are used when the exact model is not feasible. Reinforcement learning algorithms are used in self - driving cars or when learning games with human opponents.

IV. CONCLUSION : We tried to summarize that machine learning techniques can be successfully applied to various applications like, broad pre-clinical systematic review; that they can detect and validate also update the human errors in the training and validation datasets. It significantly improves the system performance. These machine learning techniques are being integrated into existing systematic review applications to enable more wide-spread use. In the future, machine learning and error analysis techniques that are optimized for different types of review topics and research questions can be applied seamlessly within the existing methodological framework.

REFERENCES:

[1] Farhat, Ramzi. 2020. "An Overview of Machine Learning Technologies and Their Use in E-Learning," 8– 11.

[2] Binkhonain, Manal, and Liping Zhao. 2019. "Expert Systems with Applications : X A Review of Machine Learning Algorithms for Identification and Classification of Non-Functional Requirements" 1. https://doi.org/10.1016/j.eswax.2019.100001.

[3] Hoos, Holger H. 2020. "A Survey on Semi-Supervised Learning." Machine Learning 109 (2): 373–440. https://doi.org/10.1007/s10994-019-05855-6.

[4] Saranyaa, T, S Sridevib, C Deisyc, Tran Duc Chungd, M K A Ahamed Khane, S Sridevib, C Deisyc, Tran Duc Chungd, and M K A Ahamed Khane. 2020. "ScienceDirect ScienceDirect Performance Analysis of Machine Learning Algorithms in Intrusion Detection System : A Review Performance Analysis of Machine Learning Algorithms in Intrusion Detection System : A Review" 00 (2019). https://doi.org/10.1016/j.procs.2020.04.133.

[5] Webb, Mary E, Andrew Fluck, Johannes Magenheim, Joyce Malyn, Juliet Waters, Michelle Deschênes, and Jason Zagami. 2020. "Machine Learning for Human Learners : Opportunities , Issues , Tensions and Threats." Educational Technology Research and Development, no. 0123456789. https://doi.org/10.1007/s11423-020-09858-2.

[6] Dalal, Kushal Rashmikant. 2018. "Review on Application of Machine Learning Algorithm for Data Science," 2018–21.

[7] Saravanan, R, and Pothula Sujatha. 2018. "Algorithms : A Perspective of Supervised Learning Approaches in Data Classification." 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), no. Iciccs: 945–49.

[8] Gianey, Hemant Kumar. 2017. "Comprehensive Review On Supervised Machine Learning Algorithms," 37–43. https://doi.org/10.1109/MLDS.2017.11.

[9] Pahwa, Kunal. 2019. "Stock Market Analysis Using Supervised Machine Learning." 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 197–200.

[10] Ray, Susmita. 2019. "A Quick Review of Machine Learning Algorithms." 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 35–39.