A Review on Machine Learning Tasks , Algorithms and its Application in Image Processing

Rashi Varshney[1], Monika Verma[2]

[1]*Student, Department of Computer Science and Engineering, BIT Durg*

[2]*Assistant Professor, Department of Computer Science and Engineering, BIT Durg*

[1]*rashivarshney97@gmail.com*, [2]*monika04verma@gmail.com*

Abstract

The advancement of technology is increasing with time, society is looking to create and develop easier ways to live and lengthen their lives. The Machine learning plays a major role in it, as it is one of the latest technologies which has attracted the researchers and motivated the people to do research on it. Machine learning is a field of computer science which gives computers an ability to learn without being explicitly programmed. In this paper, tasks and algorithms of machine learning with its application in image processing have been discussed. After reviewing all the empirical studies done in the past, a new idea has been proposed.

Keywords: *Machine learning, Machine learning tasks and algorithms, ensemble learning, image processing with machine learning*

1. Introduction

In today's time, we come across ample of data from various sources like social networking sites, internet (search engines), surveys, and many others. All these data are the big source of knowledge for the people. And to extract maximum knowledge out of these data, the technology like machine learning comes into play. Machine learning is the ability of machines to automatically learn from the data and examples (provided by the humans, as input) and to improve itself from experiences, without any intervention of humans or explicitly programming, done by them. It is an application of Artificial intelligence (AI), whose foremost goal is to allow the computers to learn from the changes made to any of its structure, program or data in order to enhance its future performances.

Machine learning have large number of applications in various fields, which are useful as well as essential in our day to day lives, some of them are:-Spam filtering, Credit card fraud detection, Recommendation systems like search engines, and social sites, medical advancement- like MRI image analysis, handwriting recognition, scene classification and detection, with the help of image processing., and many more.

Let us now understand what really machine learning is by discussing its tasks and methods.

2. Machine learning tasks

Machine learning methods and tasks are broadly classified into five broad categories on the basis of type of learning

2.1 Supervised learning

The supervised learning algorithm receives a set of inputs along with the corresponding correct outputs, and the algorithm is learned by comparing its actual output with correct outputs to find errors. It then modifies the model accordingly [1] .

The major goal of supervised learning is to build a concise model of the distribution of class labels in terms of predictor features. The resulting classifier is then used to assign class labels to the testing instances where the values of the predictor features are known, but the value of the class label is unknown [2] . Examples of Supervised Learning: Regression, Decision Tree, Random Forest, KNN, Logistic Regression etc.

2.2 Unsupervised learning

This type of learning algorithm provides with sample inputs, without any desired output present. These data sets are called unlabelled data. Since samples or training sets given to the learner are unlabelled, there is no error to reward signal to evaluate a potential solution and thus can study how systems can infer a function to describe a hidden structure from unlabeled data [3].

The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.

The unsupervised learning algorithms learn few features from the data. When new data is introduced, it uses the previously learned features to recognize the class of the data. Examples of Unsupervised Learning: Apriori algorithm, K-means [4] .

2.3 Semi supervised learning

Semi – supervised learning algorithms is a technique which combines the power of both supervised and unsupervised learning [5] .

This learning is attractive because it can potentially utilize both labeled and un- labeled data to achieve better performance than supervised learning. It can be fruit-full in those areas of machine learning and data mining where the unlabeled data is already present and getting the labeled data is a tedious process. There are many categories of semi-supervised learning are [6] :-

1) Generative Models, 2) Self-Training, 3) Transductive SVM

2.4 Reinforcement

Reinforcement learning is a sub domain of machine learning inspired by behaviorist psychology, dealing with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. This type of learning has three primary components: the agent (the learner or decision maker), the environment (everything the agent interacts with) and actions (what

the agent can do). The objective is for the agent to choose actions that maximize the expected reward over a given amount of time. Example of Reinforcement Learning: Markov Decision Process[4]

2.5 Ensemble Learning

When various individual learners are combined to form only one learner then that particular type of learning is called ensemble learning. The individual learner may be Naïve Bayes, decision tree, neural network [7] . Ensemble learning is a hot topic since 1990s. It has been observed that, a collection of learners is almost always better at doing a particular job rather than individual learners. Two popular Ensemble learning techniques are given below

1) Boosting: Boosting is a technique in ensemble learning which is used to decrease bias and variance. Boosting creates a collection of weak learners and converting them to one strong learner. A weak learner is a classifier which is barely correlated with true classification.

2)Bagging: Bagging or bootstrap aggregating is applied where the accuracy and stability of a machine learning algorithm needs to be increased. It is applicable in classification and regression. Bagging also decreases variance and helps in handling overfitting [6] .

We have elaborate each type of task of machine learning which is given above, and now to understand the specific algorithms used in it, let us first discuss the various no. of algorithms in Machine learning.with some examples.

3.Machine learning algorithms

Machine learning have profusion of algorithms, we'll discuss the some of the basic and common once:-

3.1 Linear Regression Machine Learning Algorithm:

It depicts the dependence of one variable on to the other variable[4] . It majorly shows how the change of independent variable (also called explanatory variables) affects the dependent variables (also called the factor of interest or predictor).It is the most basic and easy to understand algorithm, which is primarily used for linear relationships (problems without transforming the inputs).It comes under supervised learning.

Example-polynomial curve lifting.

3.2 Decision trees

Decision trees is one of the powerful algorithm used in data mining which is a graphical representation that illustrate all the possible outcomes of a decisions, with respect to the conditions. Here, in a decision tree, the internal node represents attributes, branches represent the outcome of a test and leaf node represent a particular class label. It is the algorithm that can process the data which contains the missing values and errors [2]. This algorithm is used majorly to classify the data and store, to classify it again. It comes under supervised learning.

3.3 Artificial Neural networks

Artificial Neural Networks are models that are inspired by the structure and/or function of biological neural networks. They are a class of pattern matching that are commonly used for regression and classification problems but are really an enormous subfield comprised of hundreds of algorithms and variations for all manner of problem types [8] .

It works on three layers. The input layer takes input (much like dendrites). The hidden layer processes the input (like soma and axon). Finally, the output layer sends the calculated output (like dendrite terminals). There are basically three types of artificial neural network: supervised, unsupervised and reinforcement [1] Firstly, it helps us understand the impact of increasing / decreasing the dataset vertically or horizontally on computational time. Secondly, it helps us understand the situations or cases where the model fits best. Thirdly, it also helps us explain why certain model works better in certain environment or situations.

3.4. Support Vector Machine

Support Vector Machine is a supervised machine learning algorithm for classification or regression problems where the dataset teaches SVM about the classes so that SVM can classify any new data. SVM algorithm tries to maximize the distance between the various classes that are involved and this is referred as margin maximization. If the line that maximizes the distance between the classes is identified, the probability to generalize well to unseen data is increased[9] .

3.5 Clustering(k-means)

K-means is a popularly used unsupervised machine learning algorithm for cluster analysis[4] . K-Means is a non- deterministic and iterative method. The algorithm operates on a given data set through pre-defined number of clusters, k. The output of K Means algorithm is k clusters with input data partitioned among the clusters. K Means clustering algorithm can be applied to group the WebPages that talk about similar concepts.

Let us now discuss and survey a major section of application of machine learning.

4. Machine learning in image processing

Images have always played an important role in human life since vision is probably human beings' most important sense. As a consequence, the field of image processing has numerous applications (medical, military, etc.). Nowadays and more than ever, images are everywhere and it is very easy for everyone to generate a huge amount of images, thanks to the advances in digital technologies. With such a profusion of images, traditional image processing techniques have to cope with more complex problems and have to face their adaptability according to human vision

Let us discuss some findings of the past:-

Face recognition, one of the latest technologies, which has motivated the researcher to research upon. It is the application of machine learning in image processing.

Antonio J. Colmenarez and Thomas S. Huang presented research progress towards face recognition with video sequences. The main approach taken in face recognition is the comparison between the information present in a database and the probe data obtained from the person to be recognized. They

proceeded in following way- First; we briefly describe a visual-learning technique that leads to an automatic, real-time, robust face and facial feature detection and tracking system. Then they introduced a fast algorithm for 2D-template matching, which provides an efficient way for massive image comparison; and finally, we describe a template-matching based, real-time face recognition system. A major difficulty in face recognition comes from the appearance variation of the face due to facial expressions, so it was a setback [10] .

Breast cancer is one of the most common kinds of cancer, as well as the leading cause of mortality among women. Mammography is currently the most effective imaging modality for the detection of breast cancer and the diagnosis of the anomalies which can identify cancerous cells. Digital mammograms have become the most effective techniques for the detection of breast cancer.

The goal of the research by the author was to increase the diagnostic accuracy of image processing and machine learning techniques for optimum classification between malignant and benign abnormalities in digital mammograms by reducing the number of misclassified cancers [11].

Image processing techniques are applied in this research to enhance the mammogram images for the computerized detection of breast cancer. Image processing algorithms and machine learning algorithms like SVM (used for remarkable pattern classification) used for mammogram image processing.

5. Conclusion

Machine learning is the topmost technology used these days. Various sorts of research on it's different domains is being taking place. Technology is developing so fast with it, that this particular domain make the researchers curious and motivates many of us to do research on it. Various tasks and algorithms in machine learning and their applications on technologies like image processing have a large number of findings, some of them are discussed above like face recognition, digital monogram, handwriting detection. It has a scope of high development in future.

References

[1] A. Singh, N. Thakur, and A. Sharma, "A Review of Supervised Machine Learning Algorithms," *IEEE*, pp. 1310–1315, 2016.

[2] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Supervised Machine Learning: A Review of Classification Techniques," *Proc. 2007 Conf. Emerg. Artif. Intell. Appl. Comput. Eng. Real Word AI Syst. with Appl. eHealth, HCI, Inf. Retr. Pervasive Technol.*, vol. 31, pp. 3–24, 2007.

[3] M. Somvanshi and P. Chavan, "A review of machine learning techniques using decision tree and support vector machine," *2016 Int. Conf. Comput. Commun. Control Autom.*, pp. 1–7, 2016.

[4] S. Kaur and S. Jindal, "A Survey on Machine Learning Algorithms," vol. 3, no. 11, pp. 6–14, 2016.

[5] C. Olivier, B. Schölkopf, and A. Zien, *Semi-Supervised Learning*, vol. 1, no. 2. 2006.

[6] A. Dey and A. S. Learning, "Machine Learning Algorithms : A Review," vol. 7, no. 3, pp. 1174–1179, 2016.

[7] H. Crc, *Ensemble Methods*. 2014.

[8] D. Sharma and N. Kumar, "A Review on Machine Learning Algorithms , Tasks and Applications," vol. 6, no. 10, pp. 1548–1552, 2017.

[9] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: A review of classification and combining techniques," *Artif. Intell. Rev.*, vol. 26, no. 3, pp. 159–190, 2006.

[10] B. Freeman, "Face detection and recognition," *Science (80-.).*, 2005.

[11] J. Nagi, "the Application of Image Processing and Machine Learning Techniques for Detection and Classification of Cancerous Tissues in Digital Mammograms," 2011.