Enrich, XII(II): 30 - 38, Jan – June, 2021

A Brief Study on the Concepts, Algorithms and Applications of Machine Learning

J. Arockia Jeyanthi, Assistant Professor
PG Department of Computer Science
A.P.C. Mahalaxmi College for Women, Thoothukudi

Abstract

Machine Learning, a branch of artificial intelligence which is one of the most interesting subfield of Computer Science is the latest buzzword floating all around the World. It is a data analysis method which is employed to automate analytical model building. Using Machine Learning with minimal human intervention, the systems can learn from the data, identify the patterns and take decisions (i.e.) "Without being explicitly programmed ML provides computers the learning capability"**.**

Till now, Machine Learning acts as the base for statistical-computational principles of learning procedures, has also developed many algorithms which are very regularly used for pattern recognition, text interpretation and various other purposes and has paved the way for a separate research interest in data mining - in identifying hidden regularities or irregularities in data that is growing by second.

This paper focuses on specifying the concepts, evolution, importance, applications and some of the popular Machine Learning algorithms

Keywords: Machine Learning, Artificial Intelligence, Algorithms, Data

I. Introduction

An application of Artificial Intelligence (AI), Machine Learning (ML) provides systems with the ability to learn automatically and to improve from experience without the necessity for explicitly programmed. ML always focuses to develop computer programs which can access data and use it to learn for themselves.

So using Machine Learning a researcher, carries out his/her work based on the training data set or examples provided to the machine at the start. Instead of developing an algorithm to directly address the problem he/she seeks an approach through which the machine that is the algorithm will come up with its own solution

The learning process begins with data or observations, instruction or direct experience, in order to look for patterns in data and to make better decisions in the future based on the examples that we provide. The main aim is to make the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Example: Training the students for exam.

Students prepare themselves for the exams with complete understanding of the subject. They feed their machine (brain) with a good amount of high-quality data before the examination. They actually train their brain with input and output. Gradually, the performance

keeps on increasing by solving practice test papers and find the performance (accuracy /score) by comparing answers with answer key given and thus gaining more confidence with the adopted approach.

The same way models are built, train the machine with data (by giving both inputs and outputs to model) and at the right time test on data (with input only) and thus it achieves our model scores by comparing its answer with the actual output which has not been fed while training. Researchers are working tirelessly to improve algorithms and techniques so that these models perform even much better.

While artificial intelligence (AI) is the broad science of mirroring human abilities, machine learning is a specific subset of AI that trains a machine how to learn.

II. Evolution of Machine Learning

With the advent of modern computing technologies, machine learning today is not like machine learning of the past. Current ML has originated from pattern recognition and the theory that computers without being programmed to do specific tasks can learn by itself. The reiterative aspect of ML does play a very important role because as models are exposed to new data, they are able to independently adapt. The models study from former computations to produce repeatable, reliable decisions and results. It's a science which is not new but it is the one that has gained fresh momentum.

This evolution has started from Turing machines to currently very high intelligent robots. The reachability of ML starting from its origin is very difficult to judge and measure but the results are visible and clear though. In recent years, the term 'machine learning has become very popular among developers, researchers and business alike, even though research in the field has been going on for more than 5 decades plus.

III. Applications of ML

ML plays a very important role in Speech Recognition, Computer Vision, Bio-Surveillance, Robot or Automation control and Empirical Science Experiments as listed below.

A. Speech Recognition

Machine Learning approaches are used to train the currently available speech recognition systems for better accuracy.

B. Computer Vision

Almost all the vision systems of today's world (e.g. facial recognition softwares) employ machine learning approaches for better accuracy.

C. Bio-Surveillance

ML algorithms are used by several government initiatives to track probable outbreaks of diseases.

D. Robot or Automation Control

ML methods finds its large applications in automated systems and robot.

E. Empirical Science Experiments

ML methods are incorporated in several researches of a large group data-intensive science disciplines.

Below are the other widely known examples of ML applications you may be working or familiar with:

- The self-driving Google car - the heavily hyped one
- Amazon and Netflix online recommendation offers
- On Twitter, getting to know about the customers review
- Fraud detection
- Management of email box
- Weather Forecasting
- Stock market prediction

IV. Categorisation of Ml Algorithms

Over past years, an overwhelming number of ML algorithm have been designed and introduced. Not every one of them are widely known. So, another algorithm was introduced in its place when some of them did not satisfy or solve the problem. Here the algorithms are broadly grouped into two categories and those two groups are further sub-divided. In this section I have named few of the very popular ML algorithms.

A. Group by Learning Style

1. Supervised learning — Training data or input data has a pre-determined label e.g. Positive/Negative, True/False, Spam/Not Spam etc. A classifier or a function is built and trained to predict the label of test data. The classifier is tuned properly (adjustment of parameter values) to reach a suitable level of accuracy.

2. Unsupervised learning --- Training data or input data is not labelled. A classifier is thus designed by deducing cluster or existing patterns in the training datasets.

3. Semi-supervised learning --- Training dataset comprises of both labelled and unlabelled data. The classifier is trained so as to learn the patterns to classify and to label the data as well as to predict.

4. Reinforcement learning --- This algorithm is trained such that it will map action to situation so that the reward or feedback signal is maximized. The classifier here is not directly programmed to choose the action, but instead it is trained to find the most rewarding actions by trial and error.

5. Transduction --- Though it is similar to supervised learning but it does not develop an explicit classifier. It predicts the output based on the training data, the training label and the test data.

6. Learning to learn --- Here the classifier is trained to learn from the bias it induced during its previous stages.

When one needs to consider the significance of the training data and to choose the classification rule, it is necessary and efficient to organise the ML algorithms with respect to learning methods that provide the greater level of accuracy.

B. Algorithms Grouped By Similarity

1. Regression Algorithms

The co-relation between dependent (target) and independent variables is exploited by Regression analysis, a part of predictive analytics. Few of the regression models are: Logistic Regression, Linear Regression, Ordinary Least Squares Regression (OLSR), Stepwise Regression, Multivariate Adaptive Regression Splines (MARS), Locally Estimated Scatterplot Smoothing (LOESS) etc.

2. Instance-based Algorithms

Whenever a new problem or example is encountered, instance-based or memory-based learning model stores instances of training data. The problem is examined in accordance with the stored instances. if that is a better fit than the former it can simply replace a stored instance by a new one. Because of this, they are also called as winner-take-all method.

Examples of Instance-based algorithms are: K-Nearest Neighbour (KNN), Self-Organising Map (SOM), Learning Vector Quantisation (LVQ), Locally Weighted Learning (LWL) etc.

3. Regularisation Algorithm

The process of counter acting overfitting or decrease the outliers is Regularisation. It is just a simple and very powerful modification that is improved with other ML models namely Regressive Models. The regression models soothe up the regression line by castigating any bent of the curve which tries to match the outliers.

Examples: Least Absolute Shrinkage and Selection Operator (LASSO), Ridge Regression, Least-Angle Regression (LARS), Elastic Net, etc.

4. Decision Tree Algorithms

Based on certain constraints, a tree like structure that provides the possible solutions to a problem is constructed as a decision tree. It starts with a single simple decision or root, which then forks off into a number of branches until a prediction or decision is made, forming a tree. They help in formalising the problem which in turn helps to identify the potential solutions faster and more accurately than others.

Examples: Iterative Dichotomise 3 (ID3), C4.5 and C5.0, Classification and Regression Tree (CART Chi-squared Automatic Interaction Detection (CHAID) Conditional Decision Trees Decision Stump, M5, etc.

5. Bayesian Algorithms

Bayes' Theorem is employed by a group of ML algorithms to solve classification and regression problems.

Examples: Naive Bayes, Multinomial Naive Bayes, Gaussian Naive Bayes, Averaged One-Dependence Estimators (AODE), Bayesian Network (BN), Bayesian Belief Network (BBN), etc.

6. Support Vector Machine (SVM)

SVM which is a very popular ML technique can be a group of its own. It uses a separating decision plane or a hyperplane to define decision boundaries among a set of data points classified with various labels. SVM is strictly a supervised classification algorithm. In other words, this algorithm develops an optimal hyperplane utilising training data or input data and this decision plane in turn categorizes new examples. SVM can perform both linear and nonlinear classification based on the kernel in use.

7. Clustering Algorithms

Clustering works upon ingrained pattern in datasets to classify the data and label the data accordingly.

Examples: K-Means, Ward hierarchical clustering, Agglomerative clustering. , K-Medians, Affinity Propagation, Spectral Clustering, DBSCAN, Gaussian Mixtures, Birch, Mean Shift, Expectation Maximisation (EM) etc.

8. Association Rule Learning Algorithms

The correlation between apparently unassociated data is discovered with the help of Association rules. The Ecommerce websites widely uses these algorithms to predict customer behaviours and future needs and thus helps in promoting certain appealing products to him.

Examples: Apriori algorithm, Eclat algorithm etc.

9. Artificial Neural Network (ANN) Algorithms

These are models which are based on the built and operations of actual neural networks of humans or animals.ANNs are termed as non-linear models as it tries to discover complex associations between the input and the output data. Rather than considering the entire set it draws sample from data and thereby reducing cost and time.

Examples: Perceptron, Hop-field Network, Back-Propagation, Radial Basis Function Network (RBFN) etc.

10. Deep Learning Algorithms

Deep Learning algorithms are the most modernised versions of ANNs that capitalise on the abundant supply of data nowadays. Where all major portion of an abound data is unlabelled or not classified they utilises larger neural networks to solve semi-supervised problems.

Examples: Deep Boltzmann Machine (DBM), Stacked Auto-Encoders, Deep Belief Networks (DBN), Convolutional Neural Network (CNN), etc.

11. Dimensionality Reduction Algorithms

Dimensionality reduction finds it's usage to reduce a larger data set to its most discriminative components which contain relevant information and describe them with very less features. This provides a proper visualisation of high dimensionality data with many features and thus helps in implementing supervised classification in an efficient manner.

Examples: Principal Component Analysis (PCA), Partial Least Squares Regression (PLSR), Principal Component Regression (PCR), Sammon Mapping, Multidimensional Scaling (MDS), Mixture Discriminant Analysis (MDA), Projection Pursuit, Quadratic Discriminant Analysis (QDA), Linear Discriminant Analysis (LDA), Flexible Discriminant Analysis (FDA) etc.

12. Ensemble model

The main aim of an ensemble method is to integrate the projections of various weaker estimators which are singly trained in order to enhance generalisability or boost up or robustness over a single estimator. To maximise the accuracy, the types of learners and the means to incorporate them is carefully chosen.

Examples: Boosting, Bootstrapped Aggregation (Bagging), Stacked Generalisation (blending), AdaBoost, Gradient Boosted Regression Trees (GBRT), Gradient Boosting Machines (GBM), Random Forest, Extremely Randomised Trees etc.

Though numerous algorithms and techniques have been introduced as specified above contributed by many researchers, it is in fact that most of the practical ML approach uses three main supervised algorithm or their variants. The three algorithms viz Naive Bayes, Support Vector Machine and Decision Tree. Most of the researchers have used the concept of these three, directly or with a boosting algorithm to enhance the efficiency further.

V. Conclusion

Machine learning is the most popular research area which has attracted a lot of intelligent minds and has the potential to reveal further. Perform better over a widespread domain is the foremost target of ML researchers which requires to design more efficient and practical general-purpose learning methods. Along with time and space complexity an important performance paradigm in the context of ML is that the efficiency with which a

method utilises data resources. Higher accuracy of prediction values and humanly interpretable prediction rules are also of great importance.

Being completely data-driven and having the ability, ML algorithms has an edge over direct or manual programming to examine a large amount of data in smaller intervals of time. Also, they are very often not prone to human bias and more accurate.

ML will surely help reform the general outlook of Computer Science by changing the definition of the question from "how to program a computer" to "how to empower it to program itself," ML facilitates in the development of self-diagnosing , self- monitoring and self-repairing devices and utilises the data flow available within the program rather than just processing it. Likewise, it helps to reform Statistical rules, by providing more computational stance. Obviously, both Computer Science and Statistics will also enhance ML as they develop and contribute more advanced theories which modifies the way of learning.

References

1. T. M. Mitchell, Machine Learning, McGraw-Hill International, 1997.

2. T.M. Mitchel, The Discipline of Machine Learning, CMU-ML-06-108, 2006

3. N. Cristianini and J. Shawe-Taylor. An Introduction to Support Vector Machines. Cambridge University Press, 2000.

4. E. Osuna, R. Freund, and F. Girosi. Support vector machines: training and applications. AI Memo 1602, MIT, May 1997.

5. V. Vapnik. Statistical Learning Theory. John Wiley & Sons, 1998.

6. C.J.C. Burges. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 2(2):1-47, 1998.

7. Taiwo Oladipupo Ayodele, Types of Machine Learning Algorithms, New Advances in Machine Learning, Yagang Zhang (Ed.), InTech, 2010

8. T. Mitchell, W. Cohen, E. Hruschka, P. Talukdar, J. Betteridge, A. Carlson, B. Dalvi, M. Gardner, B. Kisiel, J. Krishnamurthy, N. Lao, K.

Mazaitis, T. Mohamed, N. Nakashole, E. Platanios,A. Ritter, M. Samadi, B. Settles, R. Wang, D. Wijaya, A. Gupta, X. Chen, A. Saparov,M. Greaves, J. Welling, Never-Ending Learning, Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, 2014

9. Pedregosa *et al.*,Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830, 2011.

10. Wang, J. and Jebara, T. and Chang, S.-F. Semi-supervised learning using greedy max-cut.Journal of Machine Learning Research , Volume 14(1), 771-800 2013

11. Chapelle, O. and Sindhwani, V. and Keerthi, S. S. Optimization Techniques for Semi-Supervised Support Vector Machines, Journal of Machine Learning Research , Volume 9, 203–233, 2013

12. J. Baxter. A model of inductive bias learning. Journal of Artificial Intelligence Research, 12:149–198, 2000.

13. S. Ben-David and R. Schuller. Exploiting task relatedness for multiple task learning. In Conference on Learning Theory, 2003.

14. W. Dai, G. Xue, Q. Yang, and Y. Yu, Transferring Naive Bayes classifiers for text classification.AAAI Conference on Artificial Intelligence, 2007.

15. H. Hlynsson. Transfer learning using the minimum description length principle with a decision tree application. Master's thesis, University of Amsterdam, 2007.

16. Z. Marx, M. Rosenstein, L. Kaelbling, and T. Dietterich. Transfer learning with an ensemble of background tasks. In NIPS Workshop on Transfer Learning, 2005.

17. R Conway and D Strip, Selective partial access to a database, In Proceedings of ACM Annual Conference, 85 - 89, 1976

18. P D Stachour and B M Thuraisingham Design of LDV A multilevel secure relational databasemanagement system, IEEE Trans. Knowledge and Data Eng., Volume 2, Issue 2, 190 - 209, 1990

19. R Oppliger, Internet security: Firewalls and beyond, Comm. ACM, Volume 40, Issue 5, 92 -102, 1997

20. Rakesh Agrawal, Ramakrishnan Srikant, Privacy Preserving Data Mining, SIGMOD '00 Proceedings of the 2000 ACM SIGMOD international conference on Management of data, Volume 29 Issue 2,Pages 439-450, 2000

21. A. Carlson, J. Betteridge, B.Kisiel, B.Settles,E. R.Hruschka Jr,and T. M. Mitchell, Toward an architecture for never-ending language learning, AAAI, volume 5, 3, 2010

22. X. Chen, A. Shrivastava, and A. Gupta, Neil: Extracting visual knowledge from web data, In Proceedings of ICCV, 2013.

23. P. Donmezand J. G. Carbonell, Proactive learning: cost-sensitive active learning with multiple imperfect oracles. In Proceedings of the 17th ACM conference on In- formation and knowledge management, 619–628. ACM, 2008

24. T. M.Mitchell, J. Allen, P. Chalasani, J. Cheng, O. Etzioni, M. N. Ringuetteand J. C. Schlimmer, Theo: A framework for self-improving systems, Arch. for Intelli- gence 323–356, 1991

25. Gregory, P. A. and Gail, A. C. Self-supervised ARTMAP Neural Networks, Volume 23, 265-282, 2010

26. Cour, T. and Sapp, B. and Taskar, B. Learning from partial labels, Journal of Machine Learning Research, Volume 12, 1501-1536 2012

27. Adankon, M. and Cheriet, M. Genetic algorithm-based training for semi-supervised SVM, Neural Computing and Applications , Volume 19(8), 1197-1206, 2010