[ICESTM-2018]
ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

A SURVEY ON MACHINE ALGORITHMS LEARNING

Varun Chadha[*1], Mohit Anand[2] & Vaibhav Chittlangia[3]
*[1,2&3]3yr B.Tech, VIT University, Vellore, Tamilnadu

ABSTRACT

Over the previous couple of decades, Machine Learning (ML) has developed from the undertaking of few PC fans abusing the likelihood of PCs figuring out how to play amusements, and a piece of Mathematics (Statistics) that only from time to time considered computational methodologies, to an autonomous research teach that has not just given the essential base for measurable computational standards of learning systems, yet additionally has developedvarious calculations that are consistently utilized for content translation, design acknowledgment, and a numerous other business purposes and has prompted a different research enthusiasm for information mining to distinguish concealed regularities or inconsistencies in social information that developing by second. This paper centers around clarifying the idea and development of Machine Learning, a portion of the prominent Machine Learning calculations and attempt to think about three most prominent calculations in view of some fundamental ideas. Sentiment140 dataset was utilized and execution of every calculation regarding preparing time, forecast time and precision of expectation have been reported and thought about.

Keywords- *Machine Learning, Algorithm, Data, Training, precision*

I. INTRODUCTION

Machine learning is a worldview that may allude to gaining from past understanding (which for this situation is past information) to enhance future execution. The sole focal point of this field is programmed learning strategies. Learning alludes to alteration or change of calculation in view of past "encounters" naturally with no outside help from human.

While planning a machine (a product framework), the developer dependably has a particular reason at the top of the priority list. For example, think about J. K. Rowling's Harry Potter Series and Robert Galbraith's Cormoran Strike Series. To affirm the claim that it was without a doubt Rowling who had composed those books under the name Galbraith, two specialists were connected by The London Sunday Times and utilizing Forensic Machine Learning they could demonstrate that the claim was valid. They build up a machine learning calculation and "prepared" it with Rowling's and in addition different authors composing cases to look for and take in the basic examples and after that "test" the books by Galbraith. The calculation closed that Rowling's and Galbraith's written work coordinated the most in a few viewpoints. So as opposed to outlining a calculation to address the issue specifically, utilizing Machine Learning, a specialist look for an approach through which the machine, i.e., the calculation will think of its own answer in view of the case or preparing informational collection gave to it at first.

A. Machine learning : intersection of statistics and computer science

Machine Learning was the wonderful outcomewhen Computer Science and Statistics united. PC Science centers around building machines that take care of specific issues, and endeavors to recognize if issues are resolvable at all. The primary approach that Statistics in a general sense utilizes is information deduction, demonstrating estimates and estimating dependability of the conclusions.

The characterizing ideaof Machine Learning is somewhat extraordinary however in part subject to both in any case. While Software engineering focus on physically programming PCs, MLaddressesthe issue of getting PCs to re-program themselves at whatever point presented to new information in view of some underlying learning techniques gave. On the other hand, Statistics centers around information surmising and likelihood, Machine Learning incorporates extra worries about the feasibility and effectiveness of architectures and algorithms to process those data, compounding several learning tasks into a compact one and performance measures.

ISSN 2348 – 8034
Impact Factor- 5.070

B. Machine learning and human learning

A third research area closely related to Machine Learning is the study of human and animal brain in Neuroscience, Psychology, and related fields. The researchers proposed that how a machine could learn from experience most probably would not be significantly different than how an animal or a human mind learn with time and experience. However, the research concentrated on solving machine learning problems using learning methods of human brain did not yield much promising result so far than the researches concerned with statistical - computational approach. This might be due to the fact that human or animal psychology remains not fully understandable to date. Regardless of these difficulties, collaboration between human learning and machine learning is increasing for machine learning is being used to explain several learning techniques seeing in human or animals. For example, machine learning method of temporal difference was proposed to explain neural signals in animal learning. It is fairly expected that this collaboration is to grow considerably in coming years.

C. Data mining, artificial intelligence and machine learning

In practise, these three disciplines are so intertwined and overlapping that it's almost to draw a boundary or hierarchy among the three. To put it in other words, these three fields are symbiotically related and a combination of these approachesmay be used as a tactic to produce more efficient and sensitiveoutputs. Roughly, Data mining is basically about interpreting any kind of data, but it lays the foundation for both artificial intelligence and machine learning. In practice, it not only sample information from various sources but it analyses and recognises pattern and correlations that exists in those information that would have been difficult to interpret manually. Hence, data mining is not a mere method to prove a hypothesis but method for drawing relevant hypotheses.That mined data and the corresponding patterns and hypotheses may be utilised the basis for both machine learning and artificial intelligence. Artificial intelligence may be broadly defined asmachinesthose having the ability to solve a given problem on their own without any human intervention. The solutions are notprogrammed directly into the system but the necessary data and the AI interpreting that data produce a solution by itself. The interpretation that goes underneath is nothing but a data mining algorithm. Machine learning takes promote the approach to an advanced level by providing the data essential for a machine to train and modify suitably when exposed to new data. This is known as "training". It focuses onextracting information from considerably largesets of data, and then detects and identifies underlying patterns using various statistical measures to improve its ability to interpret new data and produce more effective results. Evidently, some parameters should be "tuned" at the incipient levelfor better productivity. Machine learning is thefoothold of artificial intelligence. It is improbable to design any machinehaving abilitiesassociated with intelligence, like language or vision, to get there at once. That task would have been almost impossible to solve. Moreover, a system can not be considered completely intelligent if it lacked the ability to learn and improve from its previous exposures.

II. PRESENT RESEARCH QUESTIONS & RELATED WORK

The Several applications mentioned earlier suggests considerable advancement so far in ML algorithms and their fundamental theory. The discipline is divulging in several direction, probing a range of learning problems. ML is a vast discipline and over past few decades numerous researchers have added their works in this field. The enumeration of these works are countably infinite and mentioning every work is out of the scope of this paper. However this paper

describes the main research questions that are being pursued at present and provide references to some of the recent notable works on that task.

A. Using unlabelled data in supervised learning[10][11][25][26][27]

Supervised learning algorithms approximate the relation between features and labels by defining anestimator $f : X \rightarrow Y$ for a particulargroup of pre-labeled training data $\{\square x_i, y_i \square\}$. The main challenge in this approach is pre-labeled data is not always readily available. So before applying Supervised Classification, data need to be preprocessed, filtered and labeled using unsupervised learning, feature extraction, dimensionality reduction etc. there by adding to the total cost. This hike in cost can be reduced effectively if the Supervised algorithm can make use of unlabelled data (e.g.,images) as well. Interestingly, in many special instances of learning problems with additional assumptions,

ISSN 2348 – 8034
Impact Factor- 5.070

unlabelled data can indeed be warranted to improve the expected accuracy of supervised learning. Like,consider classifying web pages or detecting spam emails. Currently active researchers are seriously taking into account new algorithms or new learning problems to exploit unlabelled data efficiently.

B. Transferring the learning experience[12][13][14][15][16]

In many real life problem, the supervised algorithm may involve learning a family of related functions (e.g., diagnosis functions for hospitals across the globe) rather than a single function. Even if the diagnosis functions for different cities (e.g., Kolkata and London) are presumed to be relatively different, some commonalities are anticipated as well. ML algorithms like hierarchical Bayesian methods give one approach that assumes the learning parameters of both the functions, say for Kolkata and London respectively, have some common prior probabilities, and allows the data from different city hospitals to over rule relevant priors as fitting. The subtlety further increases when the transfer among the functions are compounded.

C. Linking different ml algorithms

Various ML algorithms have been introduced and experimented on in a number of domains. One trail of research aims to discover the possible correlations among the existing ML algorithms, and appropriate case or scenarios to use a particular algorithm. Consider, theses two supervised classification algorithms, Naive Bayes and Logistic Regression. Both of them approach many data sets distinctly, but their equivalence can be demonstrated when implemented to specific types of training data (i.e., when the criteria of Naive Bayes classifier are fulfilled, and the number of examples in trying set tends to infinity). In general, the conceptual understanding of ML algorithms, their convergence features, and their respective effectiveness and limitations to date remain a radical research concern.

D. Best strategical approach for learners which collects their own data

A border research discipline focuses on learning systems that instead of mechanically using data collected by some other means, actively collects data for its own processing and learning. The research is devoted into finding the most effective strategy to completely hand over the control to the learning algorithm. For example consider a drug testing system which try to learn the success of the drug while monitoring the exposed patients for possible unknown side effects and try to in turn minimizing them.

E. Privacy preserving data mining[17][18][19][20]

This approach involves successfully applying data mining and obtaining results without exploiting the underlying information's attracting variety of research communities and beyond.

Consider, a medical diagnosis routine trained with data from hospitals all over the world. But due to privacy concerns, this kind of applications is not largely pursued. Even if this presents a cross road between data mining and data privacy, ongoing research says a system can have both. One proposed solution of the above problem is to develop a shared learning algorithm instead of a central database. Each of the hospitals will only be allowed to employ the algorithm under pre-defined restrictions to protect the privacy of the patients and then hand it over to the next. This is an booming research domain, combining statistical exploitation of data and recent cryptographic techniques to ensure data privacy.

F. Never-ending earners[21][22][23][24]

Most of the machine learning tasks entails training the learner using certain data sets, then setting aside the learner and utilize the output. Whereas, learning in humans and other animals learn continuously, adapting different skills in succession with experience, and use these learning's and abilities in a thoroughly synergistic way. Despite of sizeable commercial applications of ML algorithms, learning in machines(computers)to date has remained strikingly lacking compared to learning in human or animal. An alternative approach that more diligently capture the multiplicity, adeptness and accumulating character of learning in human, is named as never- ending learning. For instance, the Never Ending Language Learner (NELL)[8] is a learner whose function is learning to read web pages and has been reported to read the world wide web every hour since January 2010. NELL has obtained almost 80 million confidence- weighted opinions (Example, served With(tea, biscuits)) and has been able to learn million pairs of features and parameters that capacitate it to acquire these beliefs. Furthermore, it has become competent in reading (extracting) more beliefs, and overthrow old in accurate ones, adding to a collection of confidence and provenance for each belief and there by improving each day than the last.

III. CATEGORISATION OF ML ALGORITHMS

An overwhelming number of ML algorithm have been designed and introduced over past years. Not everyone of them are widely known. Some of them did not satisfy or solve the problem, so another was introduced in its place. Here the algorithms are broadly grouped into two category and those two groups are further sub-divided. This section try to name most popular ML algorithms and the next section compares three most widely used ML algorithms.

A. Group by learning style

- Supervised learning — Input data or training data has a pre-determined label e.g. True/False, Positive/Negative, Spam/Not Spam etc. A function or a classifier is built and trained to predict the label of test data. The classifier is properly tuned (parameter values are adjusted)to achieve a suitable level of accuracy.
- Unsupervised learning --- Input data or training data is not labelled. A classifier is designed by deducing existing patterns or cluster in the training datasets.
- Semi-supervised learning --- Training dataset contains both labeled and unlabelled data. The classifier's train to learn the patterns to classify and label the data as well as to predict.
- Reinforcement learning --- The algorithm is trained to map action to situation so that the reward or feedback signal is maximized. The classifier is not programmed directly to choose the action, but instead trained to find the most rewarding actions by trial and error.
- Transduction --- Though it shares similar traits with supervise learning, but it does not develop a explicit classifier. It attempts to predict the output based on training data, training label, and test data.
- Learning to learn --- The classifier is trained to learn from the bias it induced during previous stages.
- It is necessary and efficient to organize the ML algorithms with respect to learning methods when one need to consider the significance of the training data and choose the classification rule that provide the greater level of accuracy.

B. Algorithms grouped by similarity

1. Regression Algorithms

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

2. Instance-based Algorithms

Instance-based or memory-based learning model stores instances of training data instead of developing an precise definition of target function. Whenever a new problem or example is encountered, it is examined in accordance with the stored instances in order to determine or predict the target function value.It can simply replace a stored instance by a new one if that is a better fit than the former. Due to this, they are also known as winner-take-all method.

Examples:
K-Nearest Neighbor (KNN), Learning Vector Quantization (LVQ), Self-Organizing Map (SOM), Locally Weighted Learning (LWL) etc.

3. Regularization Algorithm

Regularization is simply the process of counteracting over fitting or abate the outliers. Regularization is just a simple yet powerful modification that is augmented with other existing ML models typically Regressive Models. It smoothes up the regression line by castigating any bent of the curve that tries to match the outliers. Examples: Ridge

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

4. Decision Tree Algorithms

A decision tree constructs a tree like structure involving of possible solutions to a problem based on certain constraints. It is so named for it begins with a single simple decision or root, which then forks off into a number of branches until a decision or prediction is made, forming a tree.

They are favored for its ability to formalize the problem in hand process that in turn helps identifying potential solutions faster and more accurately than others. Examples: Classification and Regression Tree (CART), Iterative Dichotomies 3 (ID3), C4.5 and C5.0, Chi-squared Automatic Interaction Detection (CHAID) , Decision Stump, M5, Conditional Decision Trees etc.

5. Bayesian Algorithms

A group of ML algorithms employ Bayes' Theorem to solve classification and regression problems. Examples: Naive Bayes, Gaussian Naive Bayes, Multinomial Naive Bayes, Averaged One-Dependence Estimators (AODE), Bayesian Belief Network (BBN), Bayesian Network (BN) etc.

6. Support Vector Machine (SVM)

SVM is so popular a ML technique that it can be a group of its own. It uses a separating hyper plane or a decision plane to demarcate decision boundaries among a set of data points classified with different labels. It is a strictly supervised classification algorithm. In other words, the algorithm develops an optimal hyperplane utilizing input data or training data and this decision plane in turns categories new examples. Based on the kernel in use, SVM can perform both linear and nonlinear classification.

7. Clustering Algorithms

Clustering is concerned with using ingrained pattern in datasets to classify and label the data accordingly. Examples: K-Means, K-Medians, Affinity Propagation, Spectral Clustering, Ward hierarchical clustering, Agglomerative clustering. DBSCAN, Gaussian Mixtures, Birch, Mean Shift, Expectation Maximization (EM) etc.

8. Association Rule Learning Algorithms

Association rules help discover correlation between apparently unassociated data. They are widely used by ecommerce websites to predict customer behaviors and future needs to promote certain appealing products to him. Examples: Apriori algorithm, Eclat algorithm etc.

9. Artificial Neural Network (ANN)

Algorithms A model based on the built and operations of actual neural networks of humans or animals. ANNs are regarded as non-linear models as it tries to discover complex associations between input and output data. But it draws sample from data rather than considering the entire set and thereby reducing cost and time. Examples: Perceptron, Back-Propagation, Hop-field Network, Radial Basis Function Network (RBFN) etc.

10. Deep Learning Algorithms

These are more modernized versions of ANNs that capitalize on the profuse supply of data today. They are utilizes larger neural networks to solve semi-supervised problems where major portion of an abound data is unlabelled or not classified. Examples: Deep Boltzmann Machine (DBM), Deep Belief Networks (DBN), Convolutional Neural Network (CNN), Stacked Auto-Encoders etc.

11. Dimensionality Reduction Algorithms

Dimensionality reduction is typically employed to reduce a larger data set to its most discriminative components to contain relevant information and describe it with fewer features. This gives a proper visualization for data with numerous features or of high dimensionality and helps in implementing supervised classification more efficiently. Examples: Principal Component Analysis (PCA), Principal Component Regression (PCR), Partial Least Squares

ISSN 2348 – 8034
Impact Factor- 5.070

Regression (PLSR), Sammon Mapping, Multidimensional Scaling (MDS), Projection Pursuit, Linear Discriminant Analysis (LDA), Mixture Discriminant Analysis (MDA), Quadratic Discriminant Analysis (QDA), Flexible Discriminant Analysis (FDA) etc.

12. Ensemble Algorithms

The main purpose of an ensemble method is to integrate the projections of several weaker estimators that are singly trained in order to boost up or enhance generalisability or robustness over a single estimator. The types of learners and the means to incorporate them is carefully chosen as to maximise the accuracy. Examples: Boosting, Bootstrapped Aggregation (Bagging), AdaBoost, Stacked Generalization (blending), Gradient Boosting Machines (GBM), Gradient Boosted Regression Trees (GBRT), Random Forest, Extremely Randomized Trees etc.

IV. MEASURING AND COMPARING PERFORMANCES OF POPULAR ML ALGORITHMS

Though various researchers have contributed to ML and numerous algorithms and techniques have been introduced as mentioned earlier, if it is closely studied most of the practical ML approach includes three main supervised algorithm or their variant. These three are namely, Naive Bayes, Support Vector Machine and Decision Tree. Majority of researchers have utilised the concept of these three, be it directly or with a boosting algorithm to enhance the efficiency further. These three algorithms are discussed briefly in the following section.

A. Naive bayes classifier

It is a supervised classification method developed using Bayes' Theorem of conditional probability with a 'Naive' assumption that every pair of feature is mutually independent. That is, in simpler words, presence of a feature is not effected by presence of another by any means. Irrespective of this over-simplified assumption, NB classifiers performed quite well in many practical situations, like in text classification and spam detection. Only a small amount of training data is needto estimate certain parameters. Beside, NB classifiershave considerably outperformed even highly advanced classification techniques.

B. Support vector machine

SVM, another supervised classification algorithm proposed by Vapnik in 1960s have recently attracted an major attention of researchers. The simple geometrical explanationof this approach involves determining an optimal separating plane or hyper plane that separates the two classes or clusters of data points justly and is equidistant from both of them. SVM was defined at first for linear distribution of data points. Later, the kernel function was introduced to tackle nonlinear Data as well.

C. Decision tree

A classification tree, popularly known as decision tree is one of the most successful supervised learning algorithm. It constructs a graph or tree that employs branching technique to demonstrate every probable result of a decision. In a decision tree representation, every internal node tests a feature, each branch corresponds to outcome of the parent node and every leaf finally assigns the class label. To classify an instance, a top-down approach is applied starting at the root of the tree. For a certain feature or node, the branch concurring to the value of the data point for that attribute is considered till a leaf is reached or a label is decided.

Now, the performances of these three were roughly compared using a set of tweets with labels positive, negative and neutral. The raw tweets were taken from Sentiment140 data set. Then those are pre-processed and labeled using a python program. Each of these classifier were exposed to same data. Same algorithm of feature selection, dimensionality reduction and k-fold validation were employed in each cases. The algorithms were compared based on the training time, prediction time and accuracy of the prediction. The experimental result is given below.

Table - 1: Comparison Between Gaussian NB, SVM and Decision Tree

Algorithm	Training Time (In sec.)	Prediction Time (In sec.)	Accuracy
Naïve Bayes (Gaussian)	2.708	0.328	0.692
SVM	6.485	2.054	0.6565
Decision Tree	454.609	0.063	0.69

But efficiency of an algorithm somewhat depends on the data set and the domain it is applied to. Under certain conditions, a ML algorithms may outperform the other.

V. APPLICATIONS

One clear sign of advancement in ML is its important real-life applications, some of which are briefly described here. It is to be noted that until 1985 there was no significant commercial applications of ML algorithms.

A. Speech recognition

All current speech recognition systems available in the market use machine learning approaches to train the system for better accuracy. In practice, most of such systems implement learning in two distinct phases: pre-shipping speaker independent training and post-shipping speaker-dependent training.

B. Computer vision

Majority of recent vision systems, e.g., facial recognition software's, systems capable of automatic classification microscopic images of cells, employ machine learning approaches for better accuracy. For example, the US Post Office uses a computer vision system with a handwriting analyzer thus trained to sort letters with handwritten addresses automatically with an accuracy level as high as 85%.

C. Bio-surveillance

Several government initiatives to track probable outbreaks of diseases uses ML algorithms. Consider the RODS project in western Pennsylvania. This project collects admissions reports to emergency rooms in the hospitals there, and the an ML software system is trained using the profiles of admitted patients in order to detect aberrant symptoms, their patterns and areal distribution. Research is ongoing to incorporate some additional data in the system, like over-the counter medicines' purchase history to provide more training data. Complexity of this kind of complex and dynamic data sets can be handled efficiently using automated learning methods only.

D. Robot or automation control

ML methods are largely used in robot and automated systems. For example, consider the use of ML to obtain control tactics for stable flight and aerobatics of helicopter. The self driving cars developed by Google usesML to train from collected terrain data.

E. Empirical science experiments

A large group data-intensive science disciplines use ML methods in several of it researches. For example, ML is being implemented in genetics, to identify unusual celestial objects in astronomy, and in Neuroscience and psychological analysis.

The other small scale yet important application of ML involves spam filtering, fraud detection, topic identification and predictive analytics (e.g., weather forecast, stock market prediction, market survey etc.). Machine learning is look into region that has pulled in a considerable measure of splendid personalities and it can possibly uncover further.

ISSN 2348 – 8034
Impact Factor- 5.070

Be that as it may, the three most imperative future sub-issues are been talked about here.

A. Clarifying human learning

A specified before, machine learning hypotheses have been perceived fitting to comprehend features of learning in people and creatures. Fortification learning calculations appraise the dopaminergic neurones actuated exercises in creatures amid compensate based learning with astonishing precision. ML calculations for revealing sporadic delineations of normally showing up pictures foresee visual highlights identified in creatures' underlying visual cortex. By and by, the vital drivers in human or creature learning like incitement, awfulness, desperation, hunger, instinctual activities and learning by trial also, blunder over various time scales, are not yet considered in ML calculations. This a potential chance to find a more summed up idea of discovering that entails both creatures and machine.

B. Programming languages containing machine learning primitives

In majority of uses, ML calculations are fused with physically coded programs as part of an application programming. The need of another programming dialect that is independent to help physically composed subroutines as well as those defined as "to be educated." It could enable the coder to define a set of information sources yields of each "to be learned" program and opt for a calculation from the gathering of essential learning methods already granted in the dialect.

Programming dialects like Python (Sckit-learn), R and so on as of now making utilization of this idea in littler degree. Be that as it may, a interesting new inquiry is raised as to develop a model to define relevant learning knowledge for every subroutine labeled as "to be picked up", timing, and security in instance of any unforeseen modification to the program's function.

C. Discernment

A summed up idea of PC perception that can interface ML calculations which are used in various type of PC discernment today including yet not restricted to exceedingly propelled vision, discourse acknowledgment and so on., is another potential research territory. One idea provoking gproblemis the incorporation of different senses (e.g., locate, hear, contact and so on) to set up a framework which utilize self-administered figuring out how to evaluate one tactile knowledge using the others. Inquires about in formative brain research have noted more viable learning in humans when various input modalities are provided, and thinks about on co-preparing strategies insinuate similar comes about.

VI. CONCLUSION

The foremost target of ML scientists is to plan more productive (regarding both time and space)and reasonable broadly useful learning techniques that can perform better finished an across the board area. With regards to ML, the proficiency with which a technique uses information assets that is additionally a critical execution worldview alongside time furthermore, space intricacy. Higher precision of expectation and humanly interpretable forecast rules are additionally of high significance.

Being totally information driven and being able to look at a lot of information in littler interims of time, ML calculations has an edge over manual or direct programming. Additionally they are regularly more precise and not inclined to human predisposition. Think about the accompanying situations:

Improvement of a product to comprehend discernment assignments utilizing sensors, similar to discourse acknowledgment, PC vision and so on. It is simple for anybody to mark a picture of a letter by the letter set it means, yet planning a calculation to play out this errand is troublesome. Customization of a product as indicated by the earth it is conveyed to. Consider, discourse acknowledgment programming projects that must be redone as indicated by the requirements of the client. Like online business locales that modifies the items shown by clients or email peruser that empowers spam location according to client inclinations. Coordinate programming does not have the capacity to adjust when presented to various condition.

ML provides a programming the adaptability and versatility when important. Notwithstanding some application (e.g., to compose grid duplication programs) where ML may neglect to be valuable, with increment of information assets and expanding request in customized customizable programming, ML will flourish in not so distant future. Other than programming advancement, ML will most likely however help reform the general outlook of Computer Science. By evolving the characterizing question from "how to program a PC" to "how to empower it to program itself," ML cloisters the advancement of devices that are self-checking, self-diagnosing and self-repairing, and the uses of the information stream accessible inside the program instead of simply handling it. In like manner, it will help change Statistical standards, by providing more computational position. Clearly, the two Statistics and Computer Science will likewise decorate ML as they create and contributemore advancedtheories to alter the method for 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.
[9] 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
[10]Pedregosa et al.,Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830, 2011.
[11]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
[12]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
[13]J. Baxter. A model of inductive bias learning. Journal of Artificial Intelligence Research, 12:149–198, 2000.
[14]S. Ben-David and R. Schuller. Exploiting task relatedness for multiple task learning. In Conference on Learning Theory, 2003.
[15]W. Dai, G. Xue, Q. Yang, and Y. Yu, Transferring Naive Bayes classifiers for text classification.AAAI Conference on Artificial Intelligence, 2007.
[16]H. Hlynsson. Transfer learning using the minimum description length principle with a decision tree application. Master's thesis, University of Amsterdam, 2007.
[17]Z. Marx, M. Rosenstein, L. Kaelbling, and T. Dietterich. Transfer learning with an ensemble of background tasks. In NIPS Workshop onTransfer Learning, 2005.
[18]R Conway and D Strip, Selective partial access to a database, In Proceedings of ACM Annual Conference, 85 - 89, 1976
[19]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
[20]R Oppliger, Internet security: Firewalls and beyond, Comm. ACM, Volume 40, Issue 5, 92 -102, 1997
[21]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

[22]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

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

[24]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

[25]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

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

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

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