UTILIZING A MACHINE LEARNING METHOD TO PREDICT POPULAR RESEARCH TOPICS

Poonam Raikwar,Vishal Paranjape
Global Nature Care Sangathans Group of Institutions

Abstract

Machine learning (ML) has changed over the past few decades from an endeavor of a few computer enthusiasts exploring the idea of computers learning to play games and a branch of mathematics (Statistics) that rarely considered computational approaches to an independent research field that has not only developed various algorithms that are frequently used in learning procedures but also provided the necessary foundation for statistical-computational principles of learning procedures. This essay aims to compare the three most prominent machine learning algorithms based on certain fundamental ideas, as well as to explain the concept and evolution of machine learning. The performance of each method in terms of training time, prediction time, and prediction accuracy has been documented and compared using the Sentiment140 dataset.

Keywords: Machine Learning, Algorithm, Data, Training, accuracy.

Introduction

Machine learning is a paradigm that can be used to describe learning from the past (in this case, historical data) to enhance performance in the future. This field only focuses on autonomous learning techniques. The term "learning" describes the automatic adjustment or enhancement of an algorithm based on prior "experiences" without any outside aid from a person.

The programmer always has a purpose in mind while creating a machine (a software system).

For instance, think about Robert Galbraith's Cormoran Strike Series and J. K. Rowling's Harry Potter Series. The London Sunday Times hired two specialists to investigate the claim that Rowling had actually written those works under the pen name Galbraith. Using forensic machine learning, the researchers were able to validate the claim. They created a machine learning algorithm and "trained" it using literary samples from Rowling and other authors in order to find and discover the underlying patterns and later "test" the Galbraith books. The algorithm came to the conclusion that Rowling and Galbraith's writing was the most similar in a number of ways.

Therefore, utilizing machine learning, a researcher looks for a method where the machine, or the algorithm, will come up with its own solution based on the example or training data set provided to it originally. This is opposed to building an algorithm to address the problem directly.

A. MACHINE LEARNING: INTERSECTION OF STATISTICS AND COMPUTER SCIENCE

The amazing result of the collaboration between computer science and statistics was machine learning. Computer science focuses on creating tools that address specific issues and seeks to determine whether issues are even solvable. Data inference, modeling of hypotheses, and assessing the veracity of the results constitute the basic methods that statistics fundamentally uses.

Even while the fundamental concept of machine learning is somewhat distinct from both, it nevertheless depends on both. ML addresses the issue of getting computers to re-program themselves whenever exposed to new data based on some initial learning strategies provided, as contrast to Computer Science which focuses on manually programming computers. Contrary to Statistics, which emphasizes data inference and probability, machine learning also addresses extra issues with the performance metrics, the viability of the structures and algorithms used to handle such data, and the efficiency of combining many learning tasks into a single task.

B. MACHINE LEARNING AND HUMAN LEARNING

The study of the human and animal brain in the domains of neuroscience, psychology, and allied disciplines is a third area of research that is strongly related to machine learning. The researchers hypothesized that a machine's capacity for experience-based learning would probably not differ noticeably from that of an animal or a human mind throughout the course of experience. However, compared to studies focused on utilizing a statistical and computational approach, research focused on employing human brain learning techniques to solve machine learning challenges has not yet produced many results that are particularly encouraging. This may be because psychology in both humans and animals is still not fully understood. Despite these challenges, there is an increase in collaboration between human learning and machine learning since machine learning is used to explain many human and animal learning strategies. For instance, a temporal difference machine learning approach was suggested to explain brain signals in animal learning. It is reasonable to anticipate that this relationship will expand significantly during the next few years.

C. DATA MINING, ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

In reality, these three disciplines are so entwined and overlapped that it is practically impossible to establish a line or hierarchy between them. In other words, these three domains are symbiotically linked, and combining these methods can be a strategy to create more sensitive and effective results.

Data mining essentially entails understanding any form of data, but it also serves as the basis for machine learning and artificial intelligence. In reality, it does more than just collect data from several sources; it also examines it, finds patterns, and correlates data that would have been challenging to decipher manually. Therefore, data mining is a tool for developing meaningful hypotheses rather than just proving a theory.This mined data, along with the accompanying patterns and hypotheses, can be used as the foundation for artificial intelligence and machine learning.

The ability of a machine to solve a problem without human assistance falls within the broad definition of artificial intelligence. The essential data and the AI's interpretation of that data develop a solution on their own rather than being directly programmed into the system. Underneath the interpretation is nothing more than a data mining program.

By supplying the data necessary for a machine to train and change appropriately when exposed to fresh data, machine learning advances the technique to a more sophisticated level. "Training" is what this is. In order to enhance its capacity to comprehend fresh data and generate more useful findings, it focuses on extracting information from sizable quantities of data. It then employs various statistical metrics to detect and identify underlying trends. Evidently, for increased production, some parameters need to be "tuned" at the beginning.

The foundation of artificial intelligence is machine learning. It seems unlikely that any computer with intelligence-related abilities, like language or vision, could be built to arrive there right away. It would have been nearly difficult to complete the assignment. In addition, a system cannot be deemed fully intelligent if it is incapable of learning from its experiences in the past.

Present Research Questions& Related Work

The Several applications mentioned earlier suggests considerable advancementso 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. Howeverthis paper describes the main research questions that are being pursued at present and provide references tosome 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 →Y for a particulargroup of pre-labeled training data {□xi, yi□}. The main challenge in this approach is pre-labeleddata 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, 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. TRANSFERRINGTHE LEARNING EXPERIENCE [12][13][14][15][16]

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

C. LINKING DIFFERENT ML ALGORITHMS

VariousML algorithms have been introduced and experimented on in a number of domains. One trail of research aims to discover thepossible 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 conceptualunderstanding of ML algorithms, theirconvergence features, and their respectiveeffectiveness 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 systemwhich try to learn the success of the drug while monitoring the exposed patients for possible unknown side effects and try to in turn minimising them.

E. PRIVACY PRESERVING DATA MINING [17][18][19][20]

This approach involvessuccessfully applying data mining and obtaining results without exploiting the underlying informationis attracting variety of research communities and beyond.

Consider,a medical diagnosis routine trainedwith 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 LEARNERS [21][22][23][24]

Most of the machine learning tasks entails training the learnerusing certain data sets, then setting aside the learner and utilise the output. Whereas, learning in humans and other animals learn continuously, adapting different skills in succession with experience, and use these learnings and abilities in a thoroughly synergistic way.Despite of sizeable commercial applications of ML algorithms, learning in machines(computers)to datehas remainedstrikinglylacking 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 webpages 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 oldinaccurateones, adding to a collection of confidence and provenance for each belief and there by improving each day than the last.

Categorization 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

1. 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.

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

3. Semi-supervised learning --- Training dataset contains both labeled and unlabelled data. The classifieris train to learn the patterns to classify and label the data as well as to predict.

4. Reinforcement learning --- The algorithm is trained to map action to situation so that the reward or feedback signal is maximised. The classifier is not programmed directlyto choose the action, but instead trained tofindthemost rewarding actions by trial and error.

5. Transduction --- Though it shares similar traits with supervise learning, but it does not develop a explicit classifier.Itattempts to predict the output based on training data, training label, and testdata.

6. Learning to learn --- The classifier is trainedto learn fromthe bias it induced during previous stages.

7. It is necessary and efficient to organise 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 Scatterplot 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 Neighbour (KNN), Learning Vector Quantisation (LVQ), Self-Organising Map (SOM), Locally Weighted Learning (LWL) etc.

3. Regularisation Algorithm

Regularisation is simply the process of counteracting overfitting or abate the outliers. Regularisation 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 castigatingany 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 atree like structure involving of possible solutions to a problem based on certain constraints. It is so namedfor 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 favoured for its ability to formalise 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 Dichotomiser 3 (ID3), C4.5 and C5.0, Chi-squared AutomaticInteraction 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. Ituses a separating hyperplane or a decision plane todemarcate decision boundaries among a set of data pointsclassified with different labels. It is a strictly supervised classification algorithm. In other words, the algorithm develops an optimal hyperplane utilising input data or training data and this decision plane in turnscategories 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 Maximisation (EM) etc.

8. Association Rule Learning Algorithms

Association rules help discovercorrelation between apparentlyunassociated data. They are widely used by e-commerce websites to predict customer behaviours 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 modelsas 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 modernised versions of ANNs that capitalise on the profuse supply of data today.

They are utiliseslarger 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 itwith fewer features. This gives a proper visualisation 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 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 Generalisation (blending), Gradient Boosting Machines (GBM), Gradient

Boosted Regression Trees (GBRT), Random Forest, Extremely Randomized Trees etc.

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 utilized 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 need to estimate certain parameters. Beside, NB classifiers have considerably outperformed even highly advanced classification techniques.

B. SUPPORT VECTOR MACHINE

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

C. DECISION TREE

A classification tree, popularly known as decision tree is one of the most successful supervised learning algorithms. 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 classifiers were exposed to same data. Same algorithm of feature selection, dimensionality reduction and k-fold validation were employed in each case. 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 algorithm may outperform the other.

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 practise, 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 softwares, 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 analyser thus trained to sort letters with handwritten addresses automatically with an accuracy level as high as 85%.

C. BIO-SURVEILLANCE

Severalgovernment initiatives to track probable outbreaks of diseasesuses 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 patientsin order to detect aberrant symptoms, their patterns and areal distribution. Research is ongoing to incorporatesome additional data in the system, like over-the-counter medicines' purchase history to provide more trainingdata. 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.).

Future scope

Machine learning is research area that has attracted a lot of brilliant minds and it has the potential to divulge further.

But the three most important future sub-problems are chosen to be discussed here.

A. EXPLAINING HUMAN LEARNING

A mentioned earlier, machine learning theories have been perceivedfitting to comprehendfeatures of learning in humans and animals. Reinforcement learning algorithms estimate the dopaminergic neurones induced activities in animals during reward-based learning with surprising accuracy. ML algorithms for uncovering sporadicdelineations of naturally appearing images predict visual features detected in animals' initial visual cortex. Nevertheless, the important drivers in human or animal learning like stimulation, horror, urgency, hunger, instinctive actions and learning by trial and error over numerous time scales, are not yet taken into account in ML algorithms. This a potential opportunity to discover a more generalised concept of learning that entailsboth animals and machine.

B. PROGRAMMING LANGUAGES CONTAINING MACHINE LEARNING PRIMITIVES

Inmajority of applications, ML algorithms are incorporated with manually coded programsas part of an application software. The need of a new programming language that is self-sufficient to support manually written subroutines as well as thosedefined as "to be learned." It could enablethe coder to definea set of inputs-outputs of every "to be learned" program andopt for an algorithm from the group of basic learning methodsalready imparted in the language. Programming languages like Python (Sckit-learn), R etc. already making use of this concept in smaller scope. But a fascinating new question is raised as to developa modeltodefinerelevant learning experience for each subroutines tagged as "to be learned", timing, and securityin case of any unforeseen modification to the program'sfunction.

C. PERCEPTION

A generalized concept of computer perception that can link ML algorithms which areused in numerous form of computer perception today including but not limited to highly advanced vision, speech recognition etc., is another potential research area. One thought-provoking problemis the integration of differentsenses (e.g., sight, hear, touch etc) to prepare a system which employ self-supervised learning to estimate one sensory knowledgeusing the others. Researches in developmental psychology have noted more effective learning in humans when various input modalities are supplied, and studies on co-training methods insinuate similar results.

Conclusion

The basic goal of ML research is to develop general purpose learning techniques that are more effective across a wide range of domains and are more efficient (in terms of both time and space). Along with time and space complexity, the effectiveness with which a method uses data resources is a key performance aspect in machine learning. Higher prediction accuracy and predictor rules that are easy to understand are also very important.

ML algorithms have an advantage over manual or direct programming since they are entirely data-driven and can analyze a lot of data in short bursts of time. Additionally, they are frequently more accurate and immune to human bias. Think about the following instances:

software creation for sensor-based perception tasks like speech recognition, computer vision, etc. Anyone can quickly identify an image of a letter by the alphabet it belongs to, but creating an algorithm to do this is challenging.

software customization for the environment in which it is used. Take speech recognition software as an example, which must be tailored to the demands of the user. Like e-commerce sites that adjust the products presented to customers' preferences or email clients that offer spam detection based on user preferences. Direct programming is incapable of adapting to various environments.

Software may be flexible and adaptable when needed thanks to machine learning (ML). Despite some applications (such writing matrix multiplication programs) where ML might not be useful, ML will flourish in the near future due to an increase in data resources and rising need for individualized, customizable software. In addition to software development, ML will definitely change how people view computer science in general. The development of self-monitoring, self-diagnosing, and self-repairing devices is prioritized by ML, which shifts the focus from how to program a computer to how to enable it to program itself. It also emphasizes the use of the data flow available within the program rather than just processing it. It will also assist in reforming statistical rules by supplying more computational support. Of course, as they advance and provide more sophisticated ideas to alter the process of learning, Statistics and Computer Science will likewise enrich ML.

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