A Detailed Excursion on Machine Learning Approach, Algorithms and Applications

Padmavathi H G[1] Majid Shaikh[2] Dr. Sasikumar M[3]
[1,2,3]Brindavan College of Engineering, Bengaluru, India

Abstract— In the course of recent decades, Machine Learning (ML) has advanced from the undertaking of barely any PC lovers misusing the chance of PCs figuring out how to mess around, and a piece of Mathematics (Statistics) that only sometimes viewed as computational methodologies, to a free examination discipline that has not just given the vital base to measurable computational standards of learning strategies, yet additionally has created different calculations that are consistently utilized for text understanding, design acknowledgment, and a numerous other business purposes and has prompted a different exploration enthusiasm for information mining to distinguish covered regularities or inconsistencies in social information that developing by second. This paper centres on clarifying the idea and advancement of Machine Learning, a portion of the well-known Machine. Machine Learning is a system of automated data processing algorithms that help to make decision making more natural and enhance performance based on the results. The "learning" implies that the algorithm can glean new information and insights without being explicitly programmed. This paper brings out the different machine learning algorithms, which are used in different scenarios.

Keywords: Machine Learning

I. INTRODUCTION

performance on t, as measured by p, improves with experience e. Suppose we feed a learning algorithm a lot of historical weather data".Inside the field of information investigation, Machine Learning utilized to devise complex models and calculations that loan themselves to expectation; in business use, this is known as prescient examination.These systematic models permit specialists, information researchers, architects, and experts to "produce solid, repeatable choices and results" and reveal "shrouded bits of knowledge" through gaining from verifiable connections.

II. COMPARATIVE STUDY OF MACHINE LEARNING

A. Relating Machine Learning with Statistics and Computer Science:

The combined effect of statistics and computer science is the machine learning concept. The basic method that statistics includes is the data inference, modelling hypothesis and measuring reliability of the conclusion. In contrast Computer science is mainly on computer programming.

Nevertheless, Machine Learning is interpreted differently and as well as partially dependent on both.

Statistics spotlights on data inference and probability, contrast, Machine learning put forth the problem of reprogramming the computer themselves, in the case of new data, based on the previous learning strategies experienced. Machine Learning also focuses on feasibility and effectiveness of architecture and algorithms for the processing of data, combining several learning tasks into one and performance measures.

B. Relating Machine Learning and Human Learning

Neuroscience, psychology and other related fields incorporate the research associated with Machine learning and human brain. The researchers bring the idea that how a machine could understand the past experience as like as a human mind understand and learn with time and past experience.

Nonetheless the research focused on solving machine learning methods of human brain that did not yield much encouraging result so far than results perturbed with statistical-computational approach. However with these difficulties, association between human learning and machine learning is increasing, as the machine learning is being used to explain several learning techniques seen in human. To mention, temporal difference method was put forth to explain neural signals.

C. Relating Machine Learning, Data Mining and Artificial Intelligence

In real the three fields are so interlaced and imbricated that to unsheathe a boundary or hierarchy among them is arduous to differentiate. Put in differently, as these fields are symbiotically related, the amalgamation of these may be used in strategic way to make more efficient and effective outputs. Data mining elucidates any kind of data; data mining is the basis for both Artificial Intelligence and Machine Learning. The aim of Data mining is to discover the unseen patterns and relationships from large data sets.

Artificial Intelligence is nothing but machines having the capability to solve the problem without the interference of human. Artificial Intelligence System show latest few behaviours associated with human, to mention a few, planning, reasoning, problem solving, knowledge representation, perception, motion and manipulation, Machine learning is the neat step to the advanced level by giving the essentials to the machine to train and modify appropriately when exposed to the new data. Machine learning is the basis of Artificial Intelligence. However, Machines or systems are not fully intelligent; they lack the capability of learning from their past experience exposures.

D. Classification of Machine Learning Algorithms

A huge number of ML Algorithms have been structured and presented over past years. Not every one of them is generally known. Some of them didn't fulfill or take care of the issue, so another was presented in its place. Here the Algorithms are extensively gathered into two class and those two gatherings are further bifurcated. This paper try to name most popular ML algorithms and compares three most widely used ML algorithms.

E. Algorithms Grouped by Learning approach

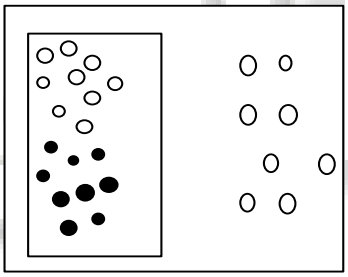
There are different ways an algorithm can model a problem based on its interaction with the experience or environment or whatever to call the input data. It is popular in machine learning and artificial intelligence to first consider the learning styles that an algorithm can adopt. There are only a few main learning styles or learning models that an algorithm can have and we'll go through them here with a few examples of algorithms and problem types that they suit. This taxonomy or way of organizing machine learning algorithms is useful because it forces you to think about the roles of the input data and the model preparation process and select one that is the most appropriate for your problem in order to get the best result.

Let's take a look at three different learning styles in machine learning algorithms:

1) Supervised Learning:

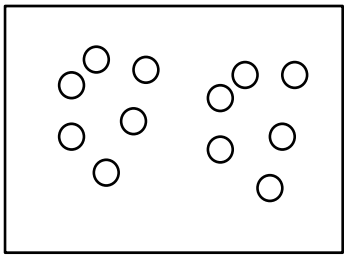
Information is called preparing information and has a referred to mark or result, for example, spam/not-spam or a stock cost at once.

A model is set up through a preparation procedure wherein it is required to make expectations and is amended when those forecasts aren't right. The preparation procedure proceeds until the model accomplishes an ideal degree of exactness on the preparation information. Model issues are arrangement and relapse. Model calculations include: Logistic Regression and the Back Propagation Neural Network.



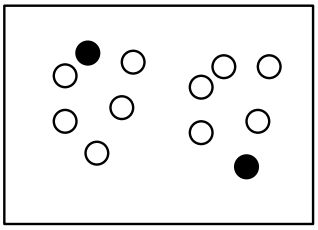
2) Unsupervised Learning:

Input data is not labeled and does not have a known result. A model is prepared by deducing structures present in the input data. This may be to extract general rules. It may be through a mathematical process to Info information isn't marked and doesn't have a known outcome. A model is set up by finding structures present in the information. This might be to extricate general principles. It might be through a scientific procedure to methodically diminish excess, or it might be to organize data by comparability. Model issues are bunching, dimensionality decrease and affiliation rule learning. Model calculations include: the Apriori calculation and K-Means.



3) Semi-Supervised Learning:

Information is a blend of named and unlabeled models. There is an ideal expectation issue yet the model must become familiar with the structures to compose the information just as make forecasts. Model issues are arrangement and relapse. Model calculations are expansions to other adaptable techniques that make suppositions about how to show the unlabeled information.



F. Algorithms Grouped By Similarity

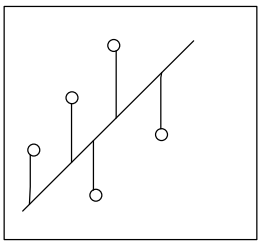
Algorithms are frequently assembled by similitude as far as their capacity (how they work). For model, tree-based techniques, and neural system enlivened strategies. This is the most helpful approach to assemble algorithms and it is the methodology we will use here. This is a valuable gathering strategy, however it isn't great. There are still calculations that could simply fit into various classifications like Learning Vector Quantization that is both a neural system roused technique and an occurrence based strategy. There are additionally algorithms that have a similar name that depict the issue and the class of algorithm, for example, Regression and Clustering. It could deal with these cases by posting algorithms twice or by choosing the gathering that subjectively is the "best" fit. It is to note that last methodology of not copying algorithms to keep things straightforward. In this area, a large number of the famous AI algorithms assembled the manner in which to believe is the most instinctive. The rundown isn't comprehensive in either the classifications or the algorithms. Here is the description of the algorithms

1) Regression Algorithms:

Relapse is worried about demonstrating the connection between factors that is iteratively refined utilizing a proportion of blunder in the forecasts made by the model. Relapse strategies are a workhorse of insights and have been co-picked into factual AI. This might be confounding in light of the fact that we can utilize relapse to allude to the class of issue and the class of calculation. Truly, relapse is a procedure.

The most famous regression algorithms are:

- Conventional Least Squares Regression (OLSR)
- Direct Regression
- Strategic Regression
- Stepwise Regression
- Multivariate Adaptive Regression Splines (MARS)
- Privately Estimated Scatterplot Smoothing (LOESS)

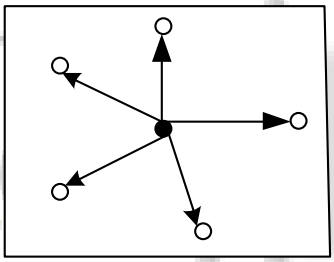


2) Instance-based Algorithms:

Occurrence based learning model is a choice issue with occasions or instances of preparing information that are considered significant or required to the model. Such techniques ordinarily develop a database of model information and contrast new information with the database utilizing a similitude measure so as to locate the best match and make an expectation. Thus, case based strategies are additionally called champ bring home all the glory techniques and memory-based learning. Concentrate is put on the representation of the put away occurrences and comparability estimates utilized between cases.

The most well-known case based algorithms are:

- k-Nearest Neighbor (kNN)
- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Locally Weighted Learning (LWL)
- Support Vector Machines (SVM)

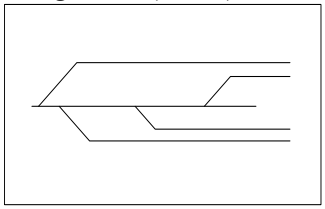


3) Regularization Algorithms:

An augmentation made to another technique (regularly relapse strategies) that punishes models dependent on their multifaceted nature, preferring less complex models that are likewise better at summing up. Regularization algorithms separately here in light of the fact that they are well known, ground-breaking and by and large basic changes made to different techniques.

The most mainstream regularization algorithms are:

- Ridge Regression
- Least Absolute Shrinkage and Selection
- Operator (LASSO)
- Elastic Net
- Least-Angle Regression (LARS)

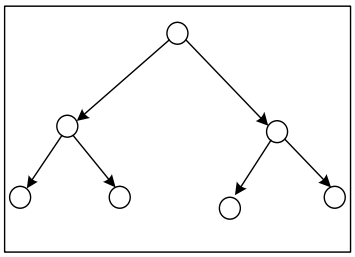


4) Decision Tree Algorithms:

Decision tree techniques build a model of decisions made dependent on genuine estimations of traits in the information. Decisions fork in tree structures until an expectation choice is made for a given record. Decision trees are prepared on information for grouping and replaces issues. Decision trees are regularly quick and exact and a big favorite in AI.

The most well-known choice tree algorithms are:

- Classification and Regression Tree (CART)
- Iterative Dichotomiser 3 (ID3)
- C4.5 and C5.0 (various renditions of an incredible methodology)
- Chi-squared Automatic Interaction Detectio (CHAID)
- Decision Stump
- M5
- Conditional Decision Trees

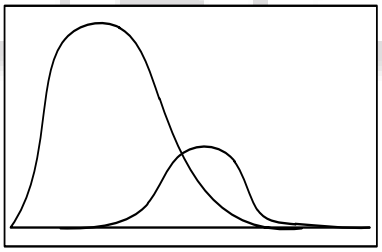


5) Bayesian Algorithms:

Bayesian strategies are those that explicitly apply Bayes' Theorem for issues, for example, classification and regression.

The most mainstream Bayesian Algorithms are:

- Naive Bayes
- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Averaged One-Dependence Estimators (AODE)
- Bayesian Belief Network (BBN)
- Bayesian Network (BN)

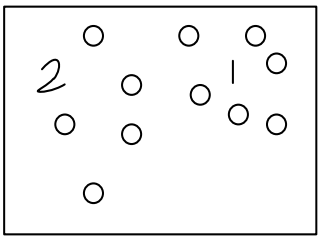


6) Clustering Algorithms:

Clustering, like regression, describes the class of problem and the class of methods. Clustering techniques are ordinarily sorted out by the displaying approaches, for example, centroid-based and hierarchal. All techniques are worried about utilizing the inborn structures in the information to best sort out the information into gatherings of most extreme shared trait.

The most famous Clustering Algorithms are:

- K-Means
- K-Medians
- Expectation Maximization (EM)
- Hierarchical Clustering

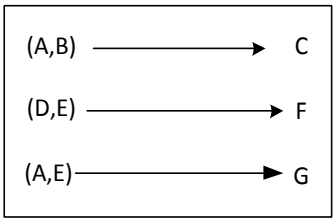


7) Association Rule Learning Algorithms:

Association rule learning methods extract rules that best explain observed relationships between variables in data. These rules can discover important and commercially useful associations in large multidimensional datasets that can be exploited by an organization.

The most popular association rule learning algorithms are:

- Apriori algorithm
- Eclat algorithm

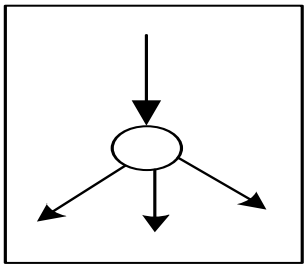


G. Artificial Neural Network Algorithms

Artificial Neural Networks are models that are enlivened by the structure as well as capacity of natural neural systems. They are a class of example coordinating that are usually utilized for relapse and grouping issues yet are actually a tremendous subfield included many Algorithms and variation for all way of issue types. Note that isolated out Deep Learning from neural systems on account of the huge development and prevalence in the field. Here it is concerned with the increasingly old style strategies.

The most well-known Artificial Neural Network Algorithms are:

- Perceptron
- Multilayer Perceptrons (MLP)
- Back-Propagation
- Stochastic Gradient Descent
- Hopfield Network
- Radial Basis Function Network (RBFN)

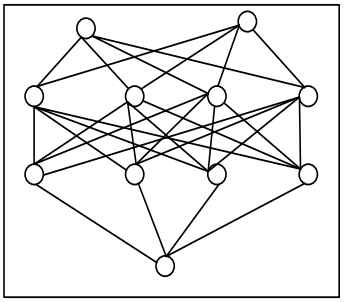


1) Deep Learning Algorithms:

Deep Learning methods are an advanced update to Artificial Neural Networks that abuse plentiful modest algorithm. They are worried about structure a lot bigger and increasingly complex neural systems and, as remarked on above, numerous strategies are worried about enormous datasets of labelled analog data for example, images, text. Audio and video.

Types of Deep learning Algorithms are:

- Convolutional Neural Network (CNN)
- Recurrent Neural Networks (RNNs)
- Long Short-Term Memory Networks (LSTMs)
- Stacked Auto-Encoders
- Deep Boltzmann Machine (DBM)
- Deep Belief Networks (DBN)

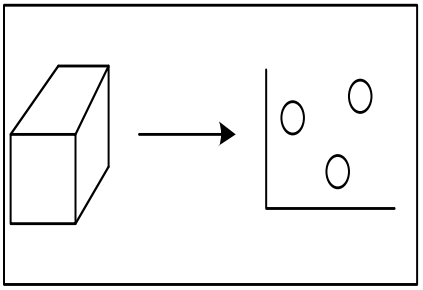


2) Dimensionality Reduction Algorithms:

Like Clustering techniques, dimensionality reduction look for and utilize the inborn structure in the information, however for this situation in a solo way or request to sum up or portray information utilizing less data. This can be useful to visualize dimensional data or to simplify data which can then be used in a supervised learning method. Many of these methods can be adapted for use in classification and regression.

The different Dimensionality Reduction Algorithms are:

- 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)



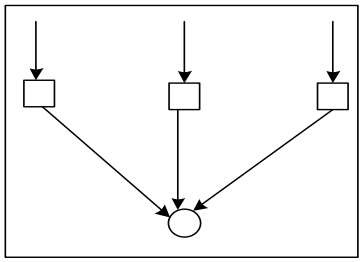
3) Ensemble Algorithms:

Ensemble techniques are models composed of multiple weaker models that are autonomously prepared and whose expectations are joined somehow or another to make the general forecast. Much exertion is placed into what sorts of feeble students to join and the manners by which to consolidate them. This is an amazing class of strategies and as such is extremely well known.

The different Ensemble Algorithms are:

- Boosting
- Bootstrapped Aggregation (Bagging)
- AdaBoost

- Weighted Average (Blending)
- Stacked Generalization (Stacking)
- Gradient Boosting Machines (GBM)
- Gradient Boosted Regression Trees (GBRT)
- Random Forest



H. Other Machine Learning Algorithms

Many algorithms were not covered.

This paper did not cover algorithms from specialty tasks in the process of machine learning, such as:

- Feature selection algorithms
- Algorithm accuracy evaluation
- Performance measures
- Optimization algorithms

It also did not cover algorithms from specialty subfields of machine learning, such as:

- Computational intelligence (evolutionary algorithms, etc.)
- Computer Vision (CV)
- Natural Language Processing (NLP)
- Recommender Systems
- Reinforcement Learning
- Graphical Models

And many more…

I. Comparison of machine Learning Algorithms

The following show comparison of six machine learning algorithms, the 10-fold cross validation procedure is used to evaluate each algorithm, importantly configured with the same random seed to ensure that the same splits to the training data are performed and that each algorithm is evaluated in precisely the same way.

Algorithm	Mean Accuracy	standard deviation accuracy
Linear Regression	0.769515	0.048411
Linear Discriminant Analysis	0.773462	0.051592
K-Nearest Neighbors	0.726555	0.061821
Classification and Regression Trees	0.695232	0.062517
Naive Bayes	0.755178	0.042766
Support Vector Machines	0.651025	0.072141

Table 1: Comparative study of Algorithm Accuracy and SD Accuracy

J. Applications of Machine Learning

1) Machine Learning in eLearning:

The eLearning industry is growing stronger with the help of technologies like machine learning. Machine learning is a sub-division of artificial intelligence. The technology involves algorithms that can draw conclusions and make predictions based on large data sets. The strength of the system lies in its ability to identify patterns and trends in the data and, based on those patterns, to make predictions that can benefit humans.

In terms of online learning, these systems can be invaluable in a number of ways.

- Machine learning can personalize content
- Automate time-consuming administrative tasks
- Provide personalized feedback
- Improve ROI for online learning

K. Machine Learning in Health Care:

The value of machine learning in healthcare is its ability to process huge datasets beyond the scope of human capability, and then reliably convert analysis of that data into clinical insights that aid physicians in planning and providing care, ultimately leading to better outcomes, lower costs of care, and increased.

There are limitless opportunities for machine learning in healthcare, for instance

- Reduce readmissions
- Prevent hospital acquired infections (HAIs)
- Reduce hospital Length-of-Stay (LOS)
- Predict chronic disease
- Reduce 1-year mortality
- Predict propensity-to-pay.
- Predict no-shows

L. Machine learning in protection of environment:

Ho and Handan-Nader hope that machine learning can complement the human monitoring efforts of environmental agencies and interest groups. Using machine learning for rote tasks can free people to do more complex ones, such as determining the possible environmental hazards of a facilityy.

The Ways Machine Learning Protects the Environment are:

- Recognizing animals in wildlife videos
- Audio-based identification of birds
- Fighting wildlife crime
- Anticipating environmental threats

III. CONCLUSION

The foremost target of ML scientists is to plan increasingly productive (as far as both time and space) and useful universally useful learning techniques that can perform better over an across the board area. With regards to ML, the productivity with which a technique uses information assets that is likewise a significant execution worldview alongside time also, space multifaceted nature. Higher precision of forecast and humanly interpretable expectation rules are additionally of high significance.

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

product to explain discernment undertakings 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, however planning a calculation to play out this errand is troublesome. Customisation of a product as indicated by nature it is conveyed to. Consider, discourse acknowledgment programming projects that must be redone as indicated by the requirements of the client. Like web based business destinations that tweaks the items shown by clients or email peruser that empowers spam location according to client inclinations. Direct programming comes up short on the capacity to adjust when presented to various conditions. ML provides a programming the adaptability and flexibility when fundamental. Regardless of some application (e.g., to compose lattice augmentation programs) where ML may neglect to be gainful, with increment of information assets and expanding request in customized customisable programming, ML will flourish in not so distant future. Other than programming advancement, MLwill 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 improvement of devices that are self-observing, self-diagnosing and self-fixing, and the uses of the information stream accessible inside the program as opposed to simply preparing it. Similarly, it will help change Statistical principles, by providing more computational position. Clearly, the two Statistics and Computer Science will likewise adorn ML as they create and contribute more advanced theories to change the method 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 information 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 Intelligence 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

