# A review on Prior Theories that enables learning in Machine Learning

Ashish Bishnoi

*CCSIT, TMU, Moradabad*

ashish.computers@tmu.ac.in

*Abstract*— **Learning is a buzzword for human civilizations from ancient times. It affects the way human's lives, grows, educate themselves and made improvements in their life styles. The most fascinating task that human dream from ancient times for having machines that had thinking and learning capability like him. With invention of computers in middle of 19th century gives wings to this thought. New domains of Artificial Intelligence and Machine Learning emerged under the umbrella of computer science with focus on designing systems that have the ability to learn. Researchers from domains of mathematics, psychology, linguistics, philosophy and computer science contributes together to design algorithms using formalism for not only representation of knowledge but also for manipulation of the same. This paper provides a review on different theories that enables learning in system's based on Machine Learning. Generally it is observed that Machine Learning means learning from large number of examples focusing on introducing some hypothesis or model. Besides this other learning algorithms exists and introduced in different order. Theories or learning algorithms reviewed are commonly used as aid for learning from different sources in AI systems using somewhat similar or altogether different concepts.**

***Keywords***— **machine learning, logic, abduction, induction, deduction, inference techniques, predicates, Bayesian belief networks.**

***Abbreviation***— **AI (Artificial Intelligence), EBL (Explanation based learning), ML (Machine Learning), ILP (Inductive Logic Programming), CBR (Case Based Reasoning), RL (Reinforcement Learning).**

## *1. INTRODUCTION*

Emergence of computers in mid 1940's gives wings to the thoughts of logicians, mathematicians and science fiction writers. They began to feel that machines with brains are soon the part of human society. Some authors felt that century old dreams of human beings about having machines that could think and act like a human being soon be fulfilled. The task was not as easy as it was considered. Till date human beings are not able to develop such machines but continuous efforts towards the goal leads to the emergence of various technologies. One of the upcoming technologies is Machine Learning that focuses on unsupervised learning methodology.

A.M. Turing proposed the question in 1950 "Can machines think"? [1] This led to researchers started focusing on the meaning of the terms "machine" and "think". The summer of 1956 in Dartmouth College, Hanover, New Hampshire, US witnessed the coining of term Artificial Intelligence by JohnMcCarthy. [2] The commonly accepted definition of Artificial Intelligence is the study of computational activities that leads to perception, reasoning and act. AI provides the concepts and models, reasoning approaches, search techniques, usage of natural languages etc, to develop machines having ability to identify, learning and solving problems that are reserved for Human Beings [3]. Logic Theorist was considered as first AI program developed in 1955 by Allen Newell.

An autonomous system must have the ability to perceive, think, act and learn from its environment. The term Machine Learning was coined in 1959 by Arthur Samuel and program that plays checkers was considered as the first program to demonstrate machine learning capabilities. Machine learning is considered as the subset of AI which in turn is considered as subset of computer science. Machine Learning founds its usage in wide area of sectors such as medical diagnosis, driverless car, household robots, personal assistants, cyber security, image and speech recognition, data and text mining etc.

The paper focuses on various learning methodologies followed in machine learning to develop machines which are more human friendly, able to adapt to its environment, generate results with accuracy and promptness. The paper provides insights to various theoretical concepts that makes the learning to machines in smooth and fast.

## 2. LEARNING CONCEPT

The term "*learn"* means the process acquired to obtain knowledge. Learning is considered as most trivial feature of a human being which continuously present throughout the life. A human being starts learning from the time of his birth till and the learning process is stopped only after its death due to its continuous association with its environment or other living beings. The learning ability makes us superior from other living beings. Various authors from the field of psychology, social sciences, linguistics had consensus about learning which must includes (1) acquisition of knowledge, facts, culture, behaviour etc., (2) process of organizing and verifying ideas or solutions related to the problem, (3) extension of experience and future learning. The task which is simple for human beings is most difficult for machines, or challenge for researchers in the area to implement it. With advent of new technologies, man is able to develop computers that can store more data than a human brain still they are not able to develop machine or computer that could exhibit the learning capabilities of 3 year old child.

Learning in human beings is categorized as associative and non associative learning which in turn includes various subtypes depending on the environment. The process through which human being able to learn an association between two events is associative learning while in non associative learning permanent change in strength due to continuous exposure of stimuli is observed. The learning process in human beings also uses learning based on episode, observation, play, augmentation, multimedia, evidence, incidental, rote, dialogic etc.

Machine Learning is one of the successful paradigms that provide the way for developing autonomous machines that can learn from their experiences efficiently. To fulfil the goal various algorithms have been invented that helps the machine in theoretical learning of concepts and tasks. [4] To illustrate the way machine should handle the problem various designing approaches are proposed that helps in learning. Learning that helps the machine to model about an object by using examples is called concept learning. It uses Boolean functions for deriving values from input and output training examples. The next section focuses on various learning theories or concepts available for machine learning.

## 3. PRIOR LEARNING THEORIES

Most of the learning theories devised for Machine Learning is based on common inference techniques used in logics and reasoning. These are deduction, induction and abduction which specifies about how conclusion is obtained from axioms and observations. Deduction is generally used to derive conclusion from axioms and observations given where as induction derives axiom from observation and some background knowledge. Explanations are derived by using abductions that use some rules backwardly from axioms and some observation. Some learning theories use one or all of the reasoning techniques as the base for deriving conclusions. Common learning theories are-

### 3.1 Explanation Based Learning (EBL)

Human beings have the tendency to learn by seeing others or their activities. In general the approach of learning from existing problems and their solutions to solve similar kind of problems in future is called EBL. This approach creates the schemata for solving common problems by the way of analysis and observation of solutions to specific problems. The gamut in which EBL lies exists from deduction to abduction. To reduce the hypothesis space and increase its consistency, EBL uses knowledge relevant for specific domain. With the help of domain knowledge, EBL is able to

identify features related to target concept for better prediction [5]. The main objective of EBL is to found operational and efficient definition for target concept that provides consistency with examples based on domain theory. EBL approach typically has three main phases of- Explanation, Analysis and Addition of new rule to domain theory. Commonly it is assumed that accurate learning procedure requires a large number of training examples but EBL provides the ability of accurate learning with few training examples. For the latter case, it requires the presence of a strong domain theory.

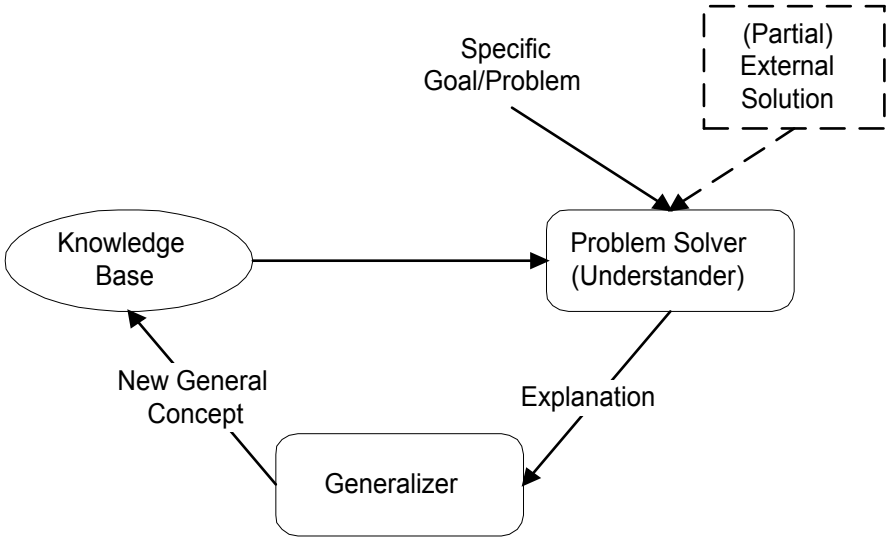


*Fig.1.Prototype of EBL Architecture*

3.2 Inductive Logic Programming (ILP)

A set of training examples consisting both positive examples and negative examples along with a domain theory are used to generate or find hypothesis in ILP. Hypothesis generated in ILP are general in nature and expressed in the form of logic programming clauses. Use of deduction and induction logic reasoning techniques in ILP geared it with better expression of hypothesis, domain knowledge and examples by using the same formalism. The notions of generalization and specialization present in ILP algorithm are used to handle subsumption. ILP systems can easily handle relationships between objects, events etc. specially multiple relationships, structured data and even some systems have the ability to generate new predicates and add them to existing domain theory. Bioinformatics and Natural Language Processing are two areas where ILP is extensively used. FOIL [6], PROGOL, GOLEM [7] and MARVIN are few ILP based systems.

3.3 Reinforcement Learning (RL)

It's an altogether a different learning paradigm and one of the basic learning paradigms in machine learning. Application area for Reinforcement Learning is real time environments which changes dynamically also the system rewards for every correct or wrong selection. Software agents extensively use RL for better adaptation and optimization. In this approach, software agents learns from their environment by interacting with it in order to achieve goals by taking sequence of actions and receiving reward

for them. The main objective of agent to make/take suitable decisions or actions in order to receive maximum rewards [8]. In RL, system is bound to learn from its experience as data set provided does not contain answers to inputs.

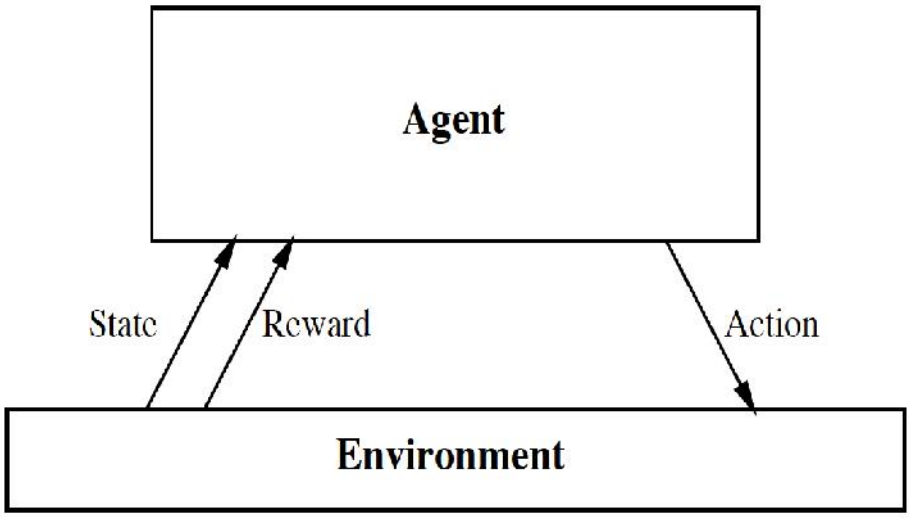


*Fig.2. Agents interaction with environment*

For handling different problems in agent environment, RL offers model based and model free approaches. Adaptive Dynamic Programming (ADP) is an model based approach whereas Direct Estimate, Monte Carlo Simulation (MC) and Temporal-Difference learning (TD) are examples of model free approach. It is widely used in the areas of gaming and software agents.

3.4 Case Based Reasoning (CBR)

Another approach for solving new problems using past experiences and solutions to related problems is case based reasoning. Main motivations behind this approach are general consideration that similar problems have similar solutions. The system stores the previous experiences and problems that are solved earlier in the form of cases in a case base. To find a solution for a problem it selects the case that is closest to the problem. It adapts the selected case for finding the solution for current problem. If solution is successfully found then it stores the adapted case into case base for increasing its competency. Learning behaviour in this approach is supported by similarity or memory based learning. CBR system cycle consists of four phases that includes retrieving of similar case, reuse the case or solution with or without adaptation, revise the case on the basis of solution generate and finally retaining which is a learning phase by storing new case to case base. Generalization is a key in inductive algorithms, CBR delays the its use till testing time, hence called lazy generalization [9] while other performs eager generalization at the time of training. Main limitation of this approach is that its dependency on statically relevant data. CBR systems are useful in the domains of customer service, recommender systems, medical applications specially in diagnosis, financial services, technical trouble shooting etc.

3.5 Learning of Bayesian Belief Networks

A probabilistic graphical model which is structurally represented by a Directed Acyclic Graph (DAG) represents BBN learning network. It is used to describe effects in terms of causes with the help of set of variables and dependencies

(conditional) among them. Learning in BBN network classified as parameter learning and structure learning. Former learning is done from conditional probabilities in a fixed variable structure while the latter deals with hidden variables. BBN network systematically uses Bayes theorem to identify posterior probability of subset hypothesis variables. These inferences are recursively applied to entire structure.

### 3.6 Model-based Clustering

Cluster refers to the group that have similar values. In machine learning, different methods based on clustering is available, out of them model based clustering is widely used as it is based on formal models. Constraints and geometric properties in covariance matrices are represented through model in family of clustering analysis. It eliminates the limitation of hierarchical clustering and k-means algorithm [10]. It uses domain knowledge represented in statistical form for harvesting instances of dataset. It has the ability to guide different clustering process up to some extent as it is based distribution or statistical based clustering. The model generates the hypothesis for instances of the aimed clusters. AUTOCLASS, COBWEB, CLASSIT [11] etc. are some systems based on clustering techniques.

## *4. CONCLUSION*

Paper starts with the discussion on Machine Learning, its importance and evolution. Application areas where ML systems could be useful in developing automated self thinking systems. In next section concept of learning is discussed and its relation with logical reasoning techniques. How these techniques act as base for ML learning methodologies. Then it focuses on different learning approaches or algorithms that are commonly used for learning. How efficiently these techniques enable the machine to learn from its experience or data sets for training? Also this section provides the list of systems that are a based on these learning approaches. In future, it is expected that ML systems not only learn but also able to select optimal algorithm for specific problem. The system is able to explain the selection.

## *5. REFERENCES*

[1] *A.M. Turing (1950). "Computing Machinery and Intelligencer".Mind..49: 433-460*.
[2] Russel and Norvig, *"Dartmouth Conference, Hamsphire"*, 2003, pp 17.
[3] A Bishnoi, H Kumar (2012), "*A Novel Designing methodologies for Artificial Intelligent Agents",* International conference proceeding, TMU, Moradabad, pp. 45
[4] Bruner, J.S. Goodnow, J.J., Austin, G.A. (1957), *"A Study of thinking"*, NewYork; John Wiley & Sons.
[5] DeJong, G.F., and R. Mooney, *"Explanation based learning: An Alternative View",* Machine Learning 1:2 (1986) 145-176
[6] Quinlan, J.R., *"Learning logical definitions from relations",* Machine Learning 5, (1990), 239-266
[7] Muggleton, S. and Feng, C. *"Efficient induction of logic problems",* In proceedings of the first conference on Algorithmic Learning Theory, Arikawa, S., Goto, S., Ohsuga, S. and Yokomori, T., Eds. Japanese Society for Artificial Intelligence, Tokyo, 1990, pp.368-391.
[8] Akar and C. Sahin., *"Reinforcement learning as a means of dynamic aggregate QoS provisioning",* Lecture Notes in Computer Science, Springer, Heidelberg, 2698:100-114,2003. Hall PTR, Upper Saddle River ,NJ,USA,2002
[9] Kolodner, J. (ED.), proceedings: " *Case Based ReasoningWorkshop"* DARP, Morgan-Kaufmann Publishers, Inc., San Mateo, CA.
[10] Everitt, B. (1980). "*Cluster Analysis",* London: Heinemann Educational Books
[11] Langley, P., & Carbonell, J.G. (1984), *"Approaches to Machine Learning",* Journal of the American Society for Information Science, 35, 306-316.