



A Review of Machine Learning

¹K. Srimitha, ²P. Kaviya Shree.

¹Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India.

²Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India.

ABSTRACT

This paper provides an extensive review of studies related to expert estimation of software development using Machine-Learning in this paper contains neural networks case based reasoning, classification and regression trees, rule induction, genetic algorithm & genetic programming for expert estimation in the field of software development etc. The main advantage of using machine learning is that, once an algorithm learns what to do with data, it can do its work automatically.

Keywords: Neural Networks, Case-based reasoning(CBR), Classification and regression, Rule induction, Genetic algorithm .

INTRODUCTION:

The poor performance results produced by statistical estimation models have flooded the estimation area for over the last decade. Their inability to handle categorical data, cope with missing data points, spread of data points and most importantly lack of reasoning capabilities has triggered an increase in the number of studies using non-traditional methods like machine learning techniques. Machine learning overlaps heavily with statistics. In fact, many machine-learning algorithms have been found to have direct counterparts with statistics. For example, boosting is now widely thought to be a form of stage wise regression using a specific type of loss function. Machine learning has a wide spectrum of applications including natural language processing, search engines, medical diagnosis, bioinformatics and cheminformatics, detecting credit card fraud, stock market analysis, classifying DNA sequences, speech and handwriting recognition, object recognition in computer vision, game playing and robot locomotion. Consequently, machine learning helps computers to find hidden insights without being programmed to do so. Moreover, machine learning works as a good predictive. The purpose of machine learning is to learn from the data. Many studies have been done on how to make machines learn by themselves . All the techniques of machine learning are explained in up-coming consequences.

1.1 Neural Networks:

A neural network is a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain. It is a type of machine learning process, called deep learning, that uses interconnected nodes or neurons in a layered structure that resembles the human brain. It creates an adaptive system that computers use to learn from their mistakes and improve continuously. Thus, artificial neural networks attempt to solve complicated problems, like summarizing documents or recognizing faces, with greater accuracy.

1.2 Case-based reasoning:

CBR is an experience-based approach to solving new problems by adapting previously successful solutions to similar problems. Addressing memory, learning, planning and problem solving, CBR provides a foundation for a new technology of intelligent computer systems that can solve problems and adapt to new situations. In CBR, the “intelligent” reuse of knowledge from already-solved problems, or cases, relies on the premise that the more similar two problems are, the more similar their solutions will be CBR developed its roots in the work of artificial intelligence theorist and cognitive psychologist, Roger Schank, and his students at Yale in the late 20th century. The researchers studied the problem-solving ability of humans and found that most people assemble solutions based on earlier experiences with similar situations.

1.3 Classification and Regression:

It is an algorithm that finds functions that help divide the dataset into classes based on various parameters. When using a Classification algorithm, a computer program gets taught on the training dataset and categorizes the data into various categories depending on what it learned. Classification algorithms find the mapping function to map the “x” input to “y” discrete output. The algorithms estimate discrete values (in other words, binary values such as 0 and 1, yes and no, true or false, based on a particular set of independent variables. To put it another, more straightforward way, classification algorithms predict an event occurrence probability by fitting data to a logit function. Regression finds correlations between dependent and independent variables. Therefore, regression algorithms help predict continuous variables such as house prices, market trends, weather patterns, oil and gas prices (a

critical task these days!), etc. The Regression algorithm's task is finding the mapping function so we can map the input variable of "x" to the continuous output variable of "y."

1.4 Rule Induction:

It is an area of machine learning in which formal rules are extracted from a set of observations. The rules extracted may represent a full scientific model of the data, or merely represent local patterns in the data. Data mining in general and rule induction in detail are trying to create algorithms without human programming but with analyzing existing data structures. In the easiest case, a rule is expressed with "if-then statements" and was created with the ID3 algorithm for decision tree learning. Rule learning algorithms are taking training data as input and creating rules by partitioning the table with cluster analysis. A possible alternative over the ID3 algorithm is genetic programming which evolves a program until it fits to the data.

1.5 Genetic Algorithm:

A genetic algorithm is an adaptive heuristic search algorithm inspired by "Darwin's theory of evolution in Nature." It is used to solve optimization problems in machine learning. It is one of the important algorithms as it helps solve complex problems that would take a long time to solve. The genetic algorithm works on the evolutionary generational cycle to generate high-quality solutions. These algorithms use different operations that either enhance or replace the population to give an improved fit solution. It basically involves five phases to solve the complex optimization problems, which are given as below:

Initialization

Fitness Assignment

Selection

Reproduction

Termination.

2. CONCLUSION:

Machine learning is a powerful tool for making predictions from data. However, it is important to remember that machine learning is only as good as the data that is used to train the algorithms. Furthermore, machine learning tools allow for the integration of traditional model-based engineering techniques and of existing domain knowledge in order to leverage the complementarity and synergy of the two solutions.

References

<https://www.investopedia.com/terms/n/neuralnetwork.asp>

<https://aws.amazon.com/what-is/neural-network>

<https://www.geeksforgeeks.org/ml-classification-vs-regression/>

<https://www.section.io/engineering-education/the-basics-of-genetic-algorithms-in-ml/>

<https://www.sciencedirect.com/topics/chemical-engineering/case-based-reasoning>