Reason for taking up Machine Learning and Summary of 'Machine Learning In Nature Encyclopedia of Cognitive Science'

Sachin Srivastava

Rutgers University

180, College Avenue, New Brunswick

New Jersey - 08901

sachin.srivastava@rutgers.edu

www.rutgers.edu

Abstract

This is a submission for mini-project 1 of our class assignment. The first part of the paper deals with answering why I want to take machine learning course. The second part is a summary of the paper Dietterich, T. G. (2003). Machine Learning In Nature Encyclopedia of Cognitive Science, London: Macmillan, 2003.

1. Reason for taking up Machine Learning

I have worked as an Analyst for 5 years in Media and Telecommunication. In my work I constantly worked problems like - what is the expected sales for a region on a particular month? Are there any chance of a decline in subscription in a particular region on a particular month? We had a massive amount of data available with us but we were using intuition and experience to answer questions like these. I feel that with the amount of data available, and using Machine Learning Techniques we can better predict such events. I believe this course will help me in making models in the future to deal with such problems.

1.1. Background

In Airtel, a telecom operator, I initiated some of the work on predicting customer behavior using predictive modeling techniques. I created a simple logistic regression model to predict customer churn. The model helped us to be better prepared for such exigencies. I often faced similar problems in my work but I was using simple techniques based on deterministic models and off the shelf modeling software. I feel an access to advanced machine learning algorithms, will help me in applying to such problems. I am curious what can be achieved using these techniques on huge amount of data available. I believe that a deep understanding of machine learning applied on the massive amount of constantly changing data has the potential for revolutionizing the world. I am excited to learn and apply machine learning to different problems. I would like to learn how this technology is being used in different ways and in different fields. I feel that the real application of these techniques is ever expanding and I can play a role in the advancement of this field.

2. Summary of 'Machine Learning In Nature Encyclopedia of Cognitive Science', Dietterich, T. G.

Machine learning is a method for computer programs to learn. Machine Learning can be used to build an accurate and effective computer system for tasks where, human expertise is not available or cant be explained, some phenomena changes rapidly, customization is required, etc. Learning can be classified as Empirical or Analytical.
Empirical learning can be supervised or unsupervised learning. In supervised learning the system should be able to generalize using a smaller sample. Decision tree is a popular algorithm in this category. Supervised learning algorithms can be used for simple classifying tasks as well as complex tasks such as sequential, time series and spatial problems. Unsupervised learning is used when objects don't have an attached class label. Some such tasks are Understanding and Visualization, Object Completion, Information Retrieval, etc. A task may also require sequential decision making, such as steering robot, car, etc. These

can be modeled as MDP(Markov decision problems), POMDP(Partially-Observable Markov Decision Problem). Learning a control policy by interacting with an unknown environment in such tasks is called Reinforcement learning. Emperical learning faces a bias-variance tradeoff where high bias can lead to underfitting while high variance can result in overfitting. Analytical learning is a different category of learning algorithm as it does not involve interaction with an external source of data, and cannot learn knowledge with new empirical content. Analytical learning focuses on improving the speed and reliability of the inferences and decisions that are performed by the computer. Some examples of analytical techniques are caching, explanation based learning, search control heuristic.

2.1. Questions Machine Learning algorithm designer should consider

The first questions that the Machine Learning designer must ask is whether to use Analytical or Emperical learning for the task. Empirical learning is learning that relies on some form of external experience, while analytical learning requires no external inputs. The second question is whether it is a Supervised Learning Task or Unsupervised Learning Task. Supervised learning is employed when a given set of objects have a class associated with it and unsupervised learning is when a class label is absent.