Machine Learning

A

Petr Berka
University of Economics, Prague, Czech Republic & University of Finance and Administration, Prague, Czech Republic

INTRODUCTION

Machine learning was a founding branch of artificial intelligence in the late 1950s. However, it has become very attractive, especially recently, thanks to the Expert Systems (ES), Knowledge Discovery from Databases (KDD), Data Mining (DM), and their developments and applications. This article covers the following topics:

- Basic definitions of machine learning.
- Formal definition of empirical concept learning from data.
- Understanding learning as search or as approximation.
- Description of basic forms of selected learning algorithms.

BACKGROUND

Human learning is one of the most important characteristics of human intelligence. In a similar way, machine learning (ML) is one of the most significant fields of artificial intelligence that attracts the interest of researchers for decades. Some researchers even argue that learning is so essential for intelligence, there can be no artificial intelligence without learning abilities.

There are numerous definitions of machine learning, most of them relating learning to the ability to improve a performance of a system over time:

The field of machine learning is concerned with the question of how to construct computer programs that automatically improve with experience. (Mitchell, 1997, p. XV)

Things learn when they change their behavior in a way that makes them perform better in the future. (Witten, 2005, p. 8)

Learning deals with "agents that can improve their behavior trough diligent study of their own experiences." (Russell, 2009, p. 649)

We can distinguish two basic types of learning: *concept learning* and *skill acquisition*. Concept learning consists of inferring and assimilating new material composed of concepts, general laws, procedures, etc. The knowledge acquired should be expressed in a formalized description that allows us to solve a problem, explain a situation, predict behavior, etc. Refinement of skills through practice consists of gradually correcting deviations between observed and desired behavior through repeated practice. So an example of concept learning can be a learning process that results in describing the concept "dog" (as a hairy animal with four legs that barks) while an example of skill refinement can be learning to drive a car or play a piano.

Based on the feedback available during learning, we can distinguish different forms of learning. *Reinforcement learning*, where the teacher observes the learning system (typically a robot) and gives his feedback to the system in the form of rewards (when learning correctly) or punishments (when learning incorrectly), is the form of learning used for skill refinement. The other forms of learning referred here are used for concept learning:

- **Supervised Learning:** Where pre-classified examples are available for learning and the task is to build a (classification or prediction) model that will work on unseen examples,

DOI: 10.4018/978-1-4666-5888-2.ch020

- **Unsupervised Learning:** Where there is no feedback at all (this is typical for clustering and segmentation tasks)
- **Semi-Supervised Learning:** Falls between supervised and unsupervised learning; it uses a small number of pre-classified examples and a large number of examples without known class,
- **Active Learning:** Is a form of supervised learning, where the algorithm can query the teacher for class membership of unclassified examples,
- **Apprenticeship Learning:** Where the system uses indirect hints derived from the teacher´s behavior, is also a form of supervised learning.

We can also distinguish between analytic and empirical learning. In analytic learning the system formulates a generalization by observing only a single example (or absent any examples at all), and by exploiting an extensive background knowledge about the given domain. In empirical learning, the generalization is based on a huge amount of data (examples) and only very limited (if any) background knowledge.

EMPIRICAL CONCEPT LEARNING

The main focus of the article is on empirical concept learning. So we will describe the machine learning methods that can be used for knowledge discovery in databases and data mining.

Learning Definition

The task of empirical concept learning can be defined as follows. Let the analyzed data have the form of a data table $\mathbf{D}_{TR}$ where each row corresponds to a single object (market basket, patient, bank client, etc.) and each column corresponds to an attribute (categorical or numerical) describing a property of the objects. Assume further that there is a special target attribute y we are interested in "computing" for unseen examples. So a description of an object $\boldsymbol{o}_i$ consists of values of "input" attributes $\boldsymbol{x}_i$ and the value of target attribute y_i

$\boldsymbol{o}_i = [\boldsymbol{x}_i, y_i]$.

When performing a classification or prediction task, we search for knowledge (represented by a decision function f) f: $\boldsymbol{x} \rightarrow y$, which, for input values $\boldsymbol{x}$ of an object, infers the value of target attribute $\hat{y} = f(\boldsymbol{x})$. During classification or prediction of an example we can make an error $Q_f(\mathbf{o}_i, \hat{y}_i)$ if the computed value $\hat{y}$ differs from the true value y. For the whole training data $\mathbf{D}_{TR}$ we can compute the total error $Err(f,\mathbf{D}_{TR})$, e.g., as

$$Err(f, \mathbf{D}_{\mathbf{TR}}) = \frac{1}{n}\sum_{i=1}^{n} Q_f(\mathbf{o}_\mathbf{i}, \hat{y}_i)$$

The goal of learning is to find such a knowledge function f^* that will minimize this error.

Empirical concept learning is based on two assumptions:

- Examples belonging to the same class have similar characteristics (similarity-based learning),
- General knowledge is inferred from a finite set of examples.

If we represent each example as a point in the attribute (feature) space, then the first assumption can be interpreted as a fact that examples belonging to the same class create clusters in the attribute space. Learning thus means identifying these clusters and representing them in such a way that this representation can be used for decision-making. The second assumption means that the learned representation of the classes must be sufficiently general to allow classification or prediction of examples that have not been used during learning. Both these assumptions have a significant impact on practical applications of empirical concept learning:

- We need to use such input attributes that fulfill the similarity assumption,
- We need to have a sufficiently representative set of training examples.

The learning process can be understood as:

- *Search* in the space of concept descriptions – in this case we learn both structure and parameters of the model,

Related Content