

Chapter 1.2

Machine Learning

João Gama

University of Porto, Portugal

André C. P. L. F. de Carvalho

University of São Paulo, Brazil

INTRODUCTION

Machine learning techniques have been successfully applied to several real world problems in areas as diverse as image analysis, Semantic Web, bioinformatics, text processing, natural language processing, telecommunications, finance, medical diagnosis, and so forth.

A particular application where machine learning plays a key role is data mining, where machine learning techniques have been extensively used for the extraction of association, clustering, prediction, diagnosis, and regression models.

This text presents our personal view of the main aspects, major tasks, frequently used algorithms, current research, and future directions of machine

learning research. For such, it is organized as follows: Background information concerning machine learning is presented in the second section. The third section discusses different definitions for Machine Learning. Common tasks faced by Machine Learning Systems are described in the fourth section. Popular Machine Learning algorithms and the importance of the loss function are commented on in the fifth section. The sixth and seventh sections present the current trends and future research directions, respectively.

BACKGROUND

Machine learning can be seen as a subfield of artificial intelligence (Bratko, 1984) and is influenced by works on statistics (inference and pat-

DOI: 10.4018/978-1-60960-818-7.ch1.2

tern recognition [Duda & Hart, 1973; Fukunaga, 1990]), databases (analytical and multivariate databases [Berson & Smith, 1997]).

Machine learning is strongly linked to search, optimization, and statistics. Several models present optimization mechanisms, like support vector machines. Others are based on statistics inference, for instance, Bayesian classifiers.

Machine learning models have been extensively used in data mining. Data mining is concerned with the discovery of useful information in large databases. Very often, the observations need to be collected, selected, and preprocessed before machine learning techniques can be employed. It is important to mention that data mining relies not only on machine learning, but also on statistics, artificial intelligence, databases, and pattern recognition.

WHAT IS MACHINE LEARNING?

Informally speaking, the main goal of machine learning is to build a computational model from past experience of what has been observed. For such, machine learning studies the automated acquisition of domain knowledge looking for the improvement of systems performance as result of experience.

In the beginning of the 1980s, Michaslky, Carbonell, and Mitchell (1983) presented one of the first definitions of machine learning “Self-constructing or self-modifying representations of what is being experienced for possible future use” (p. 10).

The focus of this definition is on programs that modify themselves in response to feedback from their environment. This definition reflects the main research lines at that time: expert systems (Weiss & Kulikowski, 1991), automatic programming, and reinforcement learning (Sutton, 1998).

A more recent definition appears in (Hand, Mannila, & Smyth, 2001) “Analysis of observational data to find unsuspected relationships and

to summarize the data in novel ways that is both understandable and useful for the data owner” (p. 1).

An even more recent definition is due to (Alpaydin, 2004), where machine learning is defined as “Programming computers to optimize a performance criterion using example data or past experience” (p. 3).

Clearly, the task here is much closer to a data analysis task, enlarging the range of practical applications, mainly industrial and commercial, where machine learning is frequently employed. In any case we can define machine learning as the acquisition of a useful (understandable) representation of a data set from its extensional representation.

MACHINE LEARNING TASKS

In a basic learning task, observations take the form of pairs. $\{\vec{x}, y\}$ The elements of the vector \vec{x} are named *independent variables* or *attributes* and the dependent variable $y=f(\vec{x})$ is an unknown function. The learning task is to obtain a predictive model or an approximation function \hat{f} able to predict \hat{y} for future observations of the independent variables \vec{x} (Mitchell, 1997). In this framework, we can consider two different problems: *classification problems*, whenever y takes values in a finite set of unordered values (e.g., $y \in \{C_1, \dots, C_n\}$), and *regression problems*, when y takes values in a subset of R . In machine learning theory, these observations are assumed to be independent and generated at random, according to a stationary probability distribution. This task is referred as *predictive or supervised learning*, because the value of the target variable y in the observations or training set is known.¹ When there is no clear target variable in the training set, we have an unsupervised learning task.

Several areas of human activity can involve *supervised machine learning*: predicting the use of

8 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/machine-learning/56127

Related Content

Models for Strategic Thinking in Intelligence Analysis

Luisa Dall'Acqua (2021). *Computational Thinking for Problem Solving and Managerial Mindset Training* (pp. 196-207).

www.irma-international.org/chapter/models-for-strategic-thinking-in-intelligence-analysis/282287

The Formal Design Model of a Real-Time Operating System (RTOS+): Conceptual and Architectural Frameworks

Yingxu Wang, Cyprian F. Ngolah, Guangping Zeng, Philip C.Y. Sheu, C. Philip Choyand Yousheng Tian (2010). *International Journal of Software Science and Computational Intelligence* (pp. 105-122).

www.irma-international.org/article/formal-design-model-real-time/43900

A Comparison of Open Source Data Mining Tools for Breast Cancer Classification

Ahmed AbdElhafeez Ibrahim, Atallah Ibrahim Hashadand Negm Eldin Mohamed Shawky (2017). *Handbook of Research on Machine Learning Innovations and Trends* (pp. 636-651).

www.irma-international.org/chapter/a-comparison-of-open-source-data-mining-tools-for-breast-cancer-classification/180964

An Empirical Study on Initializing Centroid in K-Means Clustering for Feature Selection

Amit Saxena, John Wangand Wutiphol Sintunavarat (2021). *International Journal of Software Science and Computational Intelligence* (pp. 1-16).

www.irma-international.org/article/an-empirical-study-on-initializing-centroid-in-k-means-clustering-for-feature-selection/266225

Pedagogical Software Agents for Personalized E-Learning Using Soft Computing Techniques

Mukta Goyaland Rajalakshmi Krishnamurthi (2019). *Nature-Inspired Algorithms for Big Data Frameworks* (pp. 319-338).

www.irma-international.org/chapter/pedagogical-software-agents-for-personalized-e-learning-using-soft-computing-techniques/213042