

Representation Learning

Loïc Albarracin - Nicolas Rahmat

Abstract— Fait-on un abstract ?

I. INTRODUCTION

The main goal of machine learning is to build intelligent machines, the so-called AI. We want that machines take "good decisions". In order to do that, they need knowledge. How can knowledge be given to the machines ? Knowledge, our human's one for example, can neither be transmitted verbally nor write in a program. Something do exist, which is very useful to allow machines to get this knowledge : observations ! The machines will also learn things from observations. What is then, machine learning ? It doesn't mean learning by heart, but it means generalizing from examples, where probability mass concentrates. Here, the "guesswork" is to know which new structure or strategy makes sense. In fact, most of the time, data is contained in a high-dimensionality manifold. By fighting this dimensionality problem, machines allow to discover underlying causes and factors, to explain data. Naturally, there is another huge problem : machines will never learn enough examples to cover all data in the world (e.g. functions). This problem is known as the curse of dimensionality problem. To solve it, bins are defined with probability of appearance. For example, with images, on each point, it is possible to apply a geometrical transformation, like a rotation or a translation, with the aim to get new images ! It means that a manifold could be created from all of these images. To improve machine learning, data has to be sorted. Here, human's work is to prepare the input features, which are essential for successful machine learning, which represents about 90% of total effort. This is also an opposition between handcrafting features and learning features. The goal of representation learning is to guess the features, the factors and the causes to finally get a good representation. In regards to deep representation learning, algorithms attached to it attempt to learn multiple levels of representation of increasing complexity and abstraction. When the number of levels can be data-selected, this is called Deep Learning, using several neuronal layers of computation. But a main problem has to be raised : how many particles must we have to define a cluster ? We form them by several methods such as nearest-neighbors, RBF SVMs, local non param-density, estimation and prediction, decision trees, ... There are parameters for each distinguishable region. And how can we make clusters without seeing the data ? Moreover there is the fact that features can be simple, linear or separated hence we employ layers and Deep Learning. It is at these questions that we will answer in this article.