Knowledge-driven Support Vector Machines

Yingjie Tian
Research on Fictitious Economy & Data Science
Chinese Academy of Sciences, Beijing 100190, China

Support Vector Machine (SVM) is a state-of-the-art classification technique in data mining and has been shown to have excellent generalization performance in a variety of applications. It is able to capture non-linearities in the data. However, this strength is also its main weakness. For instance, when an unlabeled example is classified by the linear decision function of SVM as positive or negative, the only explanation that can be provided is that some linear weighted sum of the variables of the example are lower (higher) than some threshold; such an explanation is completely non-intuitive to human experts. More complicate non-linear SVM models are typically regarded as incomprehensible black-box models. The opaqueness of SVM models can be remedied through the use of rule extraction techniques. If the SVM is regarded as the current state-of-the-art, SVM rule extraction can be the state-of-the-art of the (near) future. By extracting rules that mimic the black box SVM as closely as possible, we can provide some insight into the logics of the SVM model. This explanation capability is of crucial importance in any domain where the model needs to be validated before being implemented, such as in credit scoring (loan default prediction) and medical diagnosis.

On the other hand, the incorporation of prior knowledge into SVMs is the key element that allows to increase the performance in many applications. Support vector machines aim at learning an unknown decision function based only on a set of N input-output pairs (xi, yi). Nonetheless, in real world applications, a certain amount of information on the problem is usually known beforehand. For instance, in character recognition, if an image is slightly translated or rotated it still represents the same character. This prior knowledge indicates that one should incorporate invariance to translations and rotations into the classifier. The different forms of prior knowledge mainly considered are presented hierarchically and divided into two main groups: class-invariance and knowledge on the data. The first one includes invariances to transformations, to permutations and in domains of input space, whereas the second one contains knowledge on unlabeled data, the imbalance of the training set or the quality of the data. In the last decade, authors considered the introduction of prior knowledge into SVMs and some reviews can be found in the literature.

We call the above two knowledge incurred Support Vector Machines as knowledge-driven Support Vector Machines. We want to combine the two aspects closely by the knowledge and construct a system in which the knowledge incorporated---knowledge extracted—knowledge amended—knowledge integrated---and so on.