

Enclosing Machine Learning

Xunkai Wei

University of Air Force Engineering, China

Yinghong Li

University of Air Force Engineering, China

Yufei Li

University of Air Force Engineering, China

INTRODUCTION

As known to us, the cognition process is the instinct learning ability of the human being. This process is perhaps one of the most complex human behaviors. It is a highly efficient and intelligent information processing process. For a cognition process of the natural world, humans always transfer the feature information to the brain through their perception, and then the brain processes the feature information and remembers it. Due to the invention of computers, scientists are now working toward improving its artificial intelligence, and they hope that one day the computer could have its intelligent “brain” as human does. However, it is still a long way for us to go in order to let a computer truly “think” by itself.

Currently, artificial intelligence is an important and active research topic. It imitates the human brain using the idea of function equivalence. Traditionally, the neural computing and neural networks families are the majority parts of the direction (Haykin, 1994). By imitating the working mechanism of the human-brain neuron, scientists have built the neural networks theory following experimental research such as perception neurons and spiking neurons (Gerstner & Kistler, 2002) in order to understand the working mechanism of neurons.

Neural-computing and neural networks (NN) families (Bishop, 1995) have made great achievements in various aspects. Recently, statistical learning and support vector machines (SVM) (Vapnik, 1995) have drawn extensive attention and shown better performances in various areas (Li, Wei & Liu, 2004) than NN, which implies that artificial intelligence can also be made via advanced statistical computing theory. Nowadays, these two methods tend to merge under the statistical learning theory framework.

BACKGROUND

It should be noted that, for NN and SVM, the function imitation happens at microscopic view. Both of them utilize the mathematic model of neuron working mechanism. However, the whole cognition process can also be summarized as two basic principles from the macroscopical point of view (Li, Wei & Liu, 2004): the first is that **humans always cognize things of the same kind**, and the second is that **humans recognize and accept things of a new kind easily**.

In order to clarify the idea, the function imitation explanation of NN and SVM is first analyzed. The function imitation of human cognition process for pattern classification (Li & Wei, 2005) via NN and SVM can be explained as follows:

The training process of these machines actually imitates the learning processes of human being, which is called “cognizing process”. While the testing process of an unknown sample actually imitates the recognizing process of human being, which is called “recognizing process”.

From a mathematical point of view, the feature space is divided into many partitions according to the selected training principles (Haykin, 1994). Each feature space partition is then linked with a corresponding class. Given an unknown sample, NN or SVM detects its position and then assigns the indicator. For more details, the reader should refer to (Haykin, 1994 & Vapnik, 1995).

Now, suppose a sample database is given, if a totally unknown new sample comes, both SVM and NN will not naturally recognize it correctly and will consider it to the closest known in the learned classes (Li & Wei, 2005).

The root cause of this phenomenon lies in the fact that the learning principle of the NN or SVM is based on

feature space partition. This kind of learning principle may amplify each class's distribution region especially when the samples of different kinds are small due to incompleteness. Thus it is impossible for NN or SVM to detect the unknown new samples successfully.

However, this phenomenon is quite easy for humans to handle. Suppose that we have learned some things of the same kind before, and then if given similar things we can easily recognize them. And if we have never encountered with them, we can also easily tell that they are fresh things. Then we can remember their features in the brain. Sure, this process will not affect other learned things. This point surely makes our new learning paradigm different from NN or SVM.

MAIN FOCUS

Humans generally cognize things of one kind and recognize totally unknown things of a new kind easily. So why not let the learning machine “cognize” or “recognize” like humans (Li, Wei & Liu, 2004). Thus our intention is only focused on learning each single class instead of separating them. To learn each class, we can **let each class be cognized or described by a cognitive learner** to imitate the “cognizing process”. Therefore, the false alarm zone is now greatly cut down for small samples case. After training, each minimum volume bounding cognitive learner scatters in the feature space. And all learners' boundaries consist of the whole knowledge to the learned classes. If given an unknown sample, the cognitive recognizer then detects **whether the unknown sample is located inside a cognitive learner's boundary** to imitate the “recognizing process”. If the sample is totally new (i.e., none of the trained cognitive learner contains the sample), it can be again described by a new cognitive learner and the new obtained learner can be added to the feature space without affecting others. This concludes the basic process of our proposed enclosing machine learning paradigm (Wei, Li & Li, 2007A).

Mathematic Modeling

In order to make previously mentioned ideas practical, they have to be linked with concrete mathematical models (Wei, Li & Li, 2007A). Actually the first principle can be modeled as a minimum volume

enclosing problem. The second principle can be ideally modeled as a point detection problem. Generally, the minimum volume enclosing problem is quite hard to handle samples generated in arbitrary distribution and especially if such the distribution shape might be rather complex to be calculated directly. Therefore, an alternative is to use regular shapes such as sphere, ellipsoid and so on to enclose all samples of the same class with the minimum volume objective (Wei, Löfberg, Feng, Li & Li, 2007). Moreover, the approximation method can be easily formulated and efficiently solved as a convex optimization problem.

Enclosing Machine Learning Concepts

The new learning methodology has three aspects. The first is to learn each class respectively, and it is called cognitive learning. The second is to detect the unknown samples' location and determine their indicator, and it is called cognitive classification. While the third is to conduct a new cognitive learning process, and it is called feedback self-learning. The third is to imitate learning samples from an unknown new kind. The whole process is depicted in Figure 1.

Cognitive Learner: A cognitive learner is defined as the bounding boundary of a minimum volume set which encloses all the given samples. The cognitive learner can be either a sphere or an ellipsoid or their combinations. Figure 2 and Figure 3 depict the examples of sphere learner, ellipsoid learner, and combinational ellipsoid learner in 2D.

We conclude that the basic geometric shapes are the best choices, because they have many commendable features: (1) regular to make calculation easier. (2) convex bodies, which guarantee the optimality. (3) fault tolerance to assure generalization performance. And thus operations like intersection, union or complement of them can be implemented easily within convex optimization framework. Then the volume of them can be minimized to enclose the given samples. This is the most important reason why we call it enclosing machine learning.

Remarks: As for the illustrated three type learner, obviously the sphere learner has the biggest volume. Next is single ellipsoid learner. The combinational ellipsoid learner has the smallest volume.

6 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/enclosing-machine-learning/10903

Related Content

Data Transformation for Normalization

Amitava Mitra (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 566-571).
www.irma-international.org/chapter/data-transformation-normalization/10877

Mining Generalized Association Rules in an Evolving Environment

Wen-Yang Lin and Ming-Cheng Tseng (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1268-1274).
www.irma-international.org/chapter/mining-generalized-association-rules-evolving/10985

Fuzzy Methods in Data Mining

Eyke Hüllermeier (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 907-912).
www.irma-international.org/chapter/fuzzy-methods-data-mining/10928

Classification and Regression Trees

Johannes Gehrke (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 192-195).
www.irma-international.org/chapter/classification-regression-trees/10819

Inexact Field Learning Approach for Data Mining

Honghua Dai (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1019-1022).
www.irma-international.org/chapter/inexact-field-learning-approach-data/10946