Machine Learning - Lecture 18
Knowledge test

Chris Thornton

December 1, 2011

Under what circumstances will delta-rule learning produce a maximum-margin separating hyperplane?

Separating hyperplanes are often only found after datapoints have been mapped into a higher-dimensional space. This is because the mapping has the effect of (a) increasing the bias of the modeling method, (b) increasing the ways in which differently classified datapoints can be distinguished, (c) decreasing the number of differently classified datapoints, (d) increasing the distance between differently classified datapoints?

A feature space based on a kernel function has (a) one dimension for each datapoint, (b) one dimension for each feature, (c) one dimension for each pair of datapoints, (d) one dimension for each pair of features?

Let's say we want to use non-linear support-vector machines to learn a classification rule for images of politicians. We need an appropriate kernel function. What does this function need to do? How could it do it?
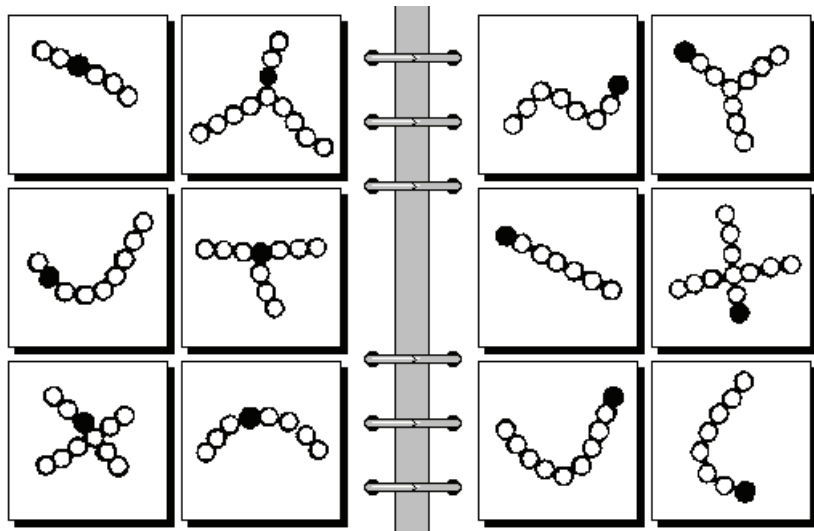
Solve this Bongard problem (#91)
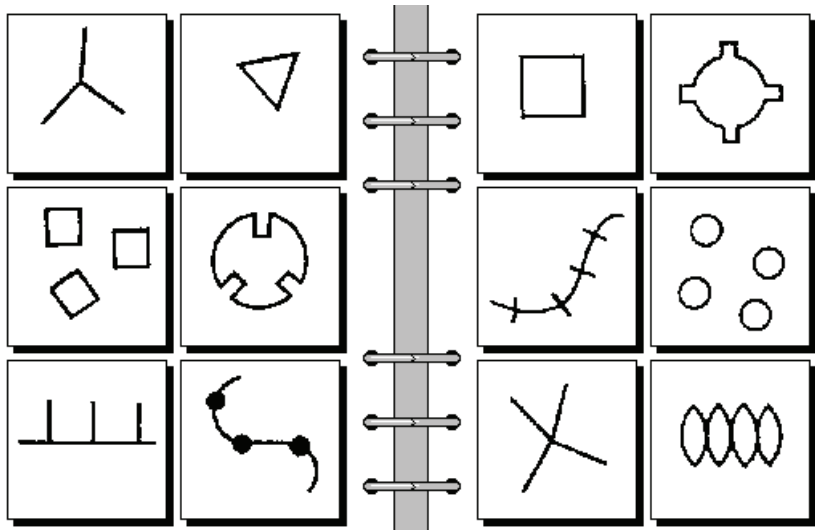
Question

Solve this Bongard problem (#94)

Solve these letter-analogy problems

Solve these letter-analogy problems

- 'abc' goes to 'abd' as 'ijklmnop' goes to what?

Solve these letter-analogy problems

- 'abc' goes to 'abd' as 'ijklmnop' goes to what?
- 'abc' goes to 'abq' as 'ijk' goes to what?

Solve these letter-analogy problems

- 'abc' goes to 'abd' as 'ijklmnop' goes to what?
- 'abc' goes to 'abq' as 'ijk' goes to what?
- 'abc' goes to 'abd' as 'mrrjjj' goes to what?

Solve these letter-analogy problems

- 'abc' goes to 'abd' as 'ijklmnop' goes to what?
- 'abc' goes to 'abq' as 'ijk' goes to what?
- 'abc' goes to 'abd' as 'mrrjjj' goes to what?

In the classical Ptolemaic system, the planets were thought to orbit the earth, and it was necessary to introduce a vast structure of epicycles to explain the observed motions. Under the heliocentric alternative (first proposed by Copernicus) all the planets are deemed to orbit the sun. It becomes much easier to explain their observed motions.

Consider a certain set of planetary observations that existed prior to the time of Copernicus. What effect did introduction of the heliocentric theory have on the (estimated) Kolmogorov complexity of the data?

How is generalization performance likely to be affected where a SVM produces a high degree of data-space distortion? Why?

The NFL argument seems to suggest there can be no completely general approach (i.e., bias) in supervised learning? How can we then explain the apparent generality of supervised learning in humans and animals?

Does the NBC model data in terms of shapes or areas in the data space? Is so, what is the form of these?

Imaging we have a dataset which records the purchasing decisions made by the customers of a certain supermarket. Can we expect the variables of the dataset to be independent? Explain.