

Supervised Learning

1. *Imagine our hypothesis is not one rectangle but a union of two (or $m > 1$) rectangles. What is the advantage of such a hypothesis class? Show that any class can be represented by such a hypothesis class with large enough m .*

In the case when there is a single rectangle, all the positive instances should form one single group; with two rectangles, for example (see figure 2.2), the positive instances can form two, possibly disjoint clusters in the input space. Note that each rectangle corresponds to a conjunction on the two input attributes, and having multiple rectangles corresponds to a disjunction. Any logical formula can be written

as a disjunction of conjunctions. In the worst case ($m = N$), we have a separate rectangle for each positive instance.

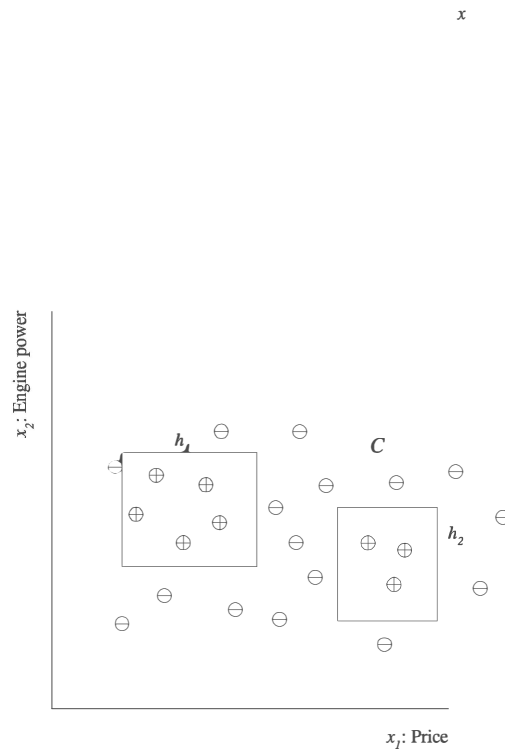


Figure 2.2 Hypothesis class is a union of two rectangles.

3. In many applications, wrong decisions—namely, false positives and false

negatives—have a monetary cost, and these two costs may be different. What is the relationship between the positioning of h between S and G and the relative costs of these?

We can see that S makes no false positives, but only false negatives; similarly, G makes no false negatives, only false positives. So if false positives and false negatives are equally bad, we want our h to be halfway; if false positives are costlier, we want h to be closer to S ; if false negatives are costlier, h should be closer to G .