

#Jenny



Finally I get this ebook, thanks for all these I can get now!

#Rio



Cool! I'am really happy

#Markus Jensen



I did not think that this would work, my best friend showed me this website, and it does! I get my most wanted eBook

#Hun Tsu



wtf this great ebook for free?!

#Che Salsa



My friends are so mad that they do not know how I have all the high quality ebook which they do not!

#Diego Butler



so many fake sites. this is the first one which worked! Many thanks

## Basics of Statistical Decision Theory

We want to attack the supervised learning problem from the viewpoint of probability and statistics. Thus, let's consider  $X$  and  $Y$  as random variables with a joint probability distribution  $\Pr(X, Y)$ .

Assume a loss function  $L(Y, f(X))$ , such as a squared loss ( $L_2$ ):

$$L(Y, f(X)) = (Y - f(X))^2$$

Choose  $f$  so that the expected prediction error is minimized: Known as regression function

$$EPE(f) = E[L(Y, f(X))] = \int (y - f(x))^2 \Pr(dy, dx)$$

Also known as (non-linear) filtering in signal processing

The minimizer is the conditional expectation:  $\hat{f}(x) = E(Y | X = x)$

So, if we knew  $\Pr(Y|X)$ , we would readily estimate the output

See [HTF] for the derivation of this minimization.  
Also see [Bishop] for a different derivation—know about a heavier machinery called calculus of variations

3

[Download PDF version of :](#)  
**Probability And Statistical Decision Theory**