$x_i \in$
R^n
x_i
?
y
x
$p(y|x)$
?
y
k
$f:$
$R \longrightarrow$
$\{1, ..., k\}$
$y =$
$f(x)$
x
y
?
?
k
$k >$
2
$k =$
?
?
f
D
x
$p(y|x, D)$
$p(y =$
$1|x, D)+$
$p(y =$
$0|x, D) =$
1
$\hat{y}$
??
?

$$\hat{y} = argmax_{c=1}^k \, p(y = c|x, D) \qquad (1)$$

?
?
?
??
$3/curves.png[Relationship between the training and test errors. Adapted from [?].] Relationship between the training and test$
?
set_1
set_2
?
?
?
f^*
x
f
y
?
$f^{(1)}$
$f^{(2)}$
$f^{(3)}$
$f(x) =$
$f^{(3)}(f^{(2)}(f^{(1)}(x)))$
?
$f(x)$
$f^*(x)$
$f^*(x)$
x
y
f^*
?
?
x
w
$s(t)$
??

$$s(t) = (x * w)(t) \qquad (2)$$

?
?
?
vs
??
$3/Conv_sep.png[3d separable convolution applied to an input with 3 feature maps.] 3d separable convolution applied to an input w$
c_i
c_o
h
w
k
$k^2 \times$
$c_i \times$
c_o
$k^2 \times$
$c_i \times$