Machine Learning

V. Adamchik CSCI 567 Spring 2019

Discussion 6 University of Southern California

CNN

Kernel Methods

Problem 1

Suppose a convolution layer takes a $4\times6\times3$ image as input and outputs a $3\times4\times6$ tensor. Which of the following is a possible configuration of this layer?

(A) One $2\times3\times6$ filter, stride 1, no zero-padding.
(B) Six $2\times3\times3$ filters, stride 1, no zero-padding.
(C) Six $3\times4\times3$ filters, stride 2, no zero-padding.
(D) Six $3\times4\times3$ filters, stride 1, 1 zero-padding.

Problem 2

Consider the following CNN. An 8×8×3 image input, followed by a convolution layer with 2 filters of size 2×2 (stride 1, no zero-padding), then another convolution layer with 4 filters of size 3×3 (stride 2, no zero-padding), and finally a max-pooling layer with a 2×2 filter (stride 1, no zero-padding).

(1) How many parameters do we need to learn for this network?

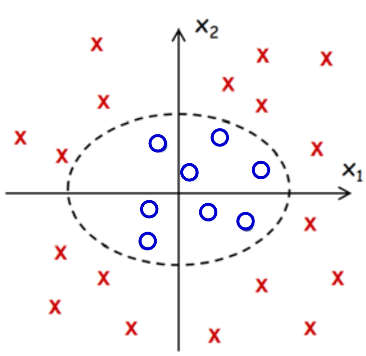
(2) What is the picture final dimension?

Problem 3

What does the following map do?

$$(x_1, x_2) \to \phi(x) = (x_1^2, x_2^2, \sqrt{2}x_1x_2)$$

Original space



Problem 4

Prove that

$$K(x,z) = \varphi(x)^{\mathrm{T}}\,\varphi(z) = \|x - z\|^2$$

is not a valid kernel by the Mercer theorem (K is a kernel if and only if the Gram matrix is positive semidefinite)

Problem 5

Every positive semidefinite matrix has eigenvalues ≥ 0

Problem 6

What is the corresponding $\phi(x)$ for the polynomial kernel of 2-dimensional vectors:

$$K(x, z) = (x^T z + c)^d$$

where c=1 and d=2.

Problem 7

Consider a polynomial kernel defined as

$$K(x, z) = (x^T z)^{50}$$

where x and z are 4-dimensional vectors.

What is a dimension of a correspondent $\phi(x)$ vector?