ML Exam 1 3.3.23

Q1 Multiple choice (4×5 points)

I don't recall the given "options", but the four questions went something like that:

i) What is the Bayes error.

ii) Something about the fischer-discriminant

iii) When do you use a biased estimator.

iv) What is the k-means algorithm

Q2 Probabilities. (4×5 points)

Consider the geometric probability distribution given by

$$p(x|\theta) = \theta(1-\theta)^{x-1} \quad \text{with } x \in \mathbb{Z}^+ = \{1, \dots, \}$$

~~i) Given the Dataset D=~~

i) Given some Dataset $D = \{x_1, \dots, x_N\}$, write down $p(D|\theta)$

ii) Given the Dataset $D = \{3, 4, 5\}$ find the parameter θ using maximum likelihood

iii) We take a Bayesian view of the problem. Calculate $P(\theta|D)$ with $D=\{2\}$ (Dataset with just one element) with the prior $\begin{cases} p(\theta)=1, & 0\le\theta\le 1 \\ p(\theta)=0 & \text{else} \end{cases}$

iv) Evaluate the probability $P(x>1|D)$

$$= \int P(x>1|\theta)\,p(\theta|D)\,d\theta$$

Hint: $\int_0^1 \theta(1-\theta)^A d\theta = \frac{1}{(A+1)(A+2)}$

↑ This was given

Q3| Kernels. (10 + 10 points)

i) Let k be a positive definite kernel, show that

$$k_2(x,x') = k(x,x') - k(x,z) - k(z,x') + k(z,z)$$

is again a positive definite kernel.

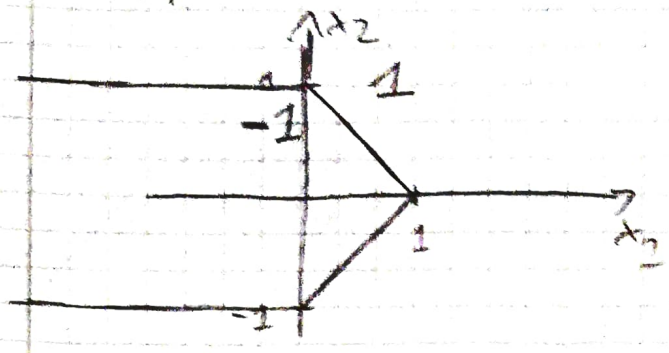
ii) Show that the feature map $\phi(x) = Wx + b$

ii) Consider the kernel $k(x,x') = \langle Wx+b, Wx'+b \rangle$

Show that the feature map $x \mapsto \phi(x-z)$ induces k_z (with $k(x,x') = \phi(x) \cdot \phi(x')$).

(Not quite sure if this question was this way around, but it should lead to a similar argumentation)

Q4| Neural network (15+5 points)



i) Build a neural network that classifies inputs (x_1, x_2) as outlined by the graph shown above using neurons of the form

$$a_j = \text{Sign}\left(\sum_i w_{ij} a_i + b_j\right)$$

Write down the weights of the neurons

ii) ~~Calculated~~ Calculate the output of the network for the input $(-2, 2)$.

Q5| Kernel ridge regression (5+10+5 points)

Consider the kernel $k(x, x') = \frac{1}{a + \|x - x'\|^2}$

i) Implement the kernel using numpy/Scipy library (They hinted at a documentation sheet)

def kernel(X_1, X_2, a)

:

return kernel.

ii) Implement the kernel ridge-regression for a given lambda

$$f(x) = K(x, X) \cdot (K + \lambda I)^{-1} \cdot y$$

def krr(X_{train}, y_{train}, X_{test}, lambda, a)

:

return y_{test}.

iii) Compute the mean squared-error of the kernel-ridge regression when testing it on the training data

The wording was different, but that's how I interpreted the question...