SGN-41007 Pattern Recognition and Machine Learning
Exam 1.3.2019
Heikki Huttunen

> Use of calculator is allowed.
> Use of other materials is not allowed.
> The exam questions need not be returned after the exam.
> You may answer in English or Finnish.

1. Are the following statements true or false? No need to justify your answer, just T or F. Correct answer: 1 pts, wrong answer: $-\frac{1}{2}$ pts, no answer 0 pts.

   (a) The Receiver Operating Characteristics curve plots the probability of detection versus the probability of false alarm for all thresholds.

   (b) Random forest has a linear decision boundary.

   (c) Least squares estimator minimizes the squared distance between the data and the model.

   (d) The ReLU (rectified linear unit) activation function is defined as

   $$f(x) = \frac{1}{1 + \exp(-x)}.$$

   (e) Maxpooling returns the average within each disjoint block of neighboring samples.

   (f) Cross-validation is used for model accuracy evaluation.

2. *Maximum likelihood estimation.* Consider the model

   $$x[n] = A \exp(-n) \sin(\theta n) + w[n], \qquad n = 0, 1, \ldots, N - 1,$$

   where $w[n] \sim \mathcal{N}(0, \sigma^2)$ and θ is a known real number. In other words, we assume that our measurement is a damped sinusoid at known frequency and phase and want to estimate the amplitude A. Derive the maximum likelihood estimator of A.
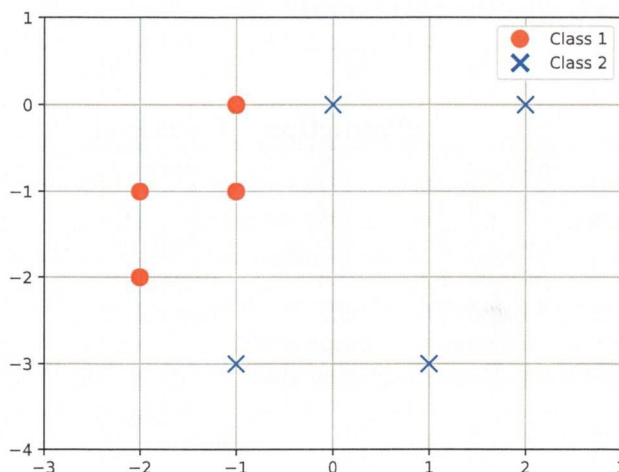

Figure 1: Training sample of question 3.

3. A dataset consists of two classes, containing four samples each. The samples are shown in Figure 1. The classes are linearly separable, and there are many linear decision boundaries that classify the training set with 100 % accuracy.

(a) (4p) Find the Linear Discriminant Analysis (LDA) classifier for this data. The covariances of the classes are

$$C_1 = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad C_2 = \frac{1}{3}\begin{pmatrix} 5 & 6 \\ 6 & 9 \end{pmatrix}.$$

Present the decision rule for sample $\mathbf{x} \in \mathbb{R}^2$ in the following format:

$$\text{Class}(\mathbf{x}) = \begin{cases} 1, & \text{if } \boxed{\textit{something}} \\ 2, & \text{otherwise} \end{cases}$$

(b) (2p) Find the threshold c at the center of the projected class means.

4. (6 pts) Consider the Keras model defined in Listing 1. Inputs are 96×96 color images from 10 categories.

(a) Draw a diagram of the network.

(b) Compute the number of parameters for each layer, and their total number over all layers.

5. A random forest classifier is trained on training data set and the `predict_proba` method is applied on the test data of five samples. The predictions and true labels are in Table 1.

In the exercises we drew a receiver operating characteristic (ROC) curve. In this question, however, you are requested to draw a precision-recall-curve (PRC) instead, which is a curve drawn on precision-recall axes by sliding the detection threshold over all reasonable values. See also the supplementary material regarding the definitions of precision and recall.

(a) Draw the precision-recall curve.

(b) What is the Area Under the PR Curve?

	Prediction	True label
Sample 1	0.8	positive
Sample 2	0.3	positive
Sample 3	0.6	positive
Sample 4	0.4	negative
Sample 5	0.2	negative

Table 1: Results on test data for question 5.