Please carefully read and follow the general instructions regarding exercises. Failing to meet the requirements might lead to penalties. https://elearn.uef.fi/mod/page/view.php?id=293750

If you suspect that something is wrong with some exercise question, please contact the lecturer.

If you face persistent issues while working on an exercise, do ask for help, e.g. during a course meeting or by contacting the lecturer via email.

Consider the dataset consisting of twelve training instances (ids 1–12, on the left) and seven test instances (ids 13–19, on the right), with four variables v_1–v_4 and a class label y that can take one of three values, shown in Figure 1.

id	v_1	v_2	v_3	v_4	y
(1)	1	4	−1	0	●
(2)	3	6	−2	0	●
(3)	7	5	−6	0	●
(4)	2	5	1	0	●
(5)	0	4	6	0	●
(6)	4	6	2	0	●
(7)	6	2	−1	0	●
(8)	8	3	−6	0	●
(9)	7	1	1	1	●
(10)	3	2	6	1	●
(11)	5	1	2	1	●
(12)	1	3	2	1	●

id	v_1	v_2	v_3	v_4	y
(13)	7	4	−6	0	●
(14)	3	3	1	0	●
(15)	4	5	−2	0	●
(16)	6	5	6	1	●
(17)	0	3	−1	0	●
(18)	9	2	1	1	●
(19)	1	1	6	0	●

Figure 1: Dataset consisting of twelve training instances (left) and seven test instances (right)

Problem 1 (Predicting class labels with a decision tree).

a) Use the decision tree shown in Figure 2 to predict labels for the seven test instances above.

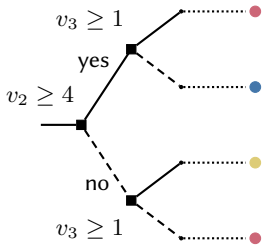

Figure 2: A decision tree

It is then assumed that the blue and yellow classes are equivalent for the purpose of the study and the task of interest is discriminating red instances (positive class) from the rest (negative class).
The output from the task in question *(a)* can be binarized accordingly.

b) Give the binarized predictions.

Problem 2 (Support Vector Machines). We trained a hard-margin linear SVM for the binary classification of separating red instances from blue and yellow ones, after mapping the dataset to two dimensions as follows, where the first dimension is the sum of attributes $v_1 + v_3$, while attribute v_4 stands as the second dimension.

For the training dataset above, after mapping the variables and considering two classes (red vs. combined blue or yellow), we obtained the following vector of Lagrange multipliers:

$$\boldsymbol{a} = \langle 0, 57, 57, 0, 0, 0, 0, 130, 0, 16, 0, 0 \rangle$$

a) Write down the transformed training dataset and highlight the support vectors.
b) Compute the weights and bias defining the separating hyperplane ($\boldsymbol{w}$ and b).
c) Compute the predicted labels for the test dataset.
d) Can you find a non-linear transformation that separates the red instances from the rest in one dimension? Give the corresponding function.