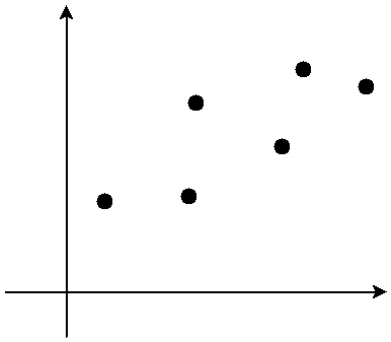
Entrance test - M33SAD

1 Principal Component Analysis (PCA)



A set of points in 2D space is shown in the figure above. This set of points should be projected onto 1D (a line) using PCA. Draw this line into the figure and mark the positions of the projected points on the line.

2 Decision tree

Patient	Temperature	Joint pain	Flu
1	39.0	no	no
2	36.8	yes	yes
3	37.2	no	no
4	38.3	yes	yes
5	36.1	yes	no

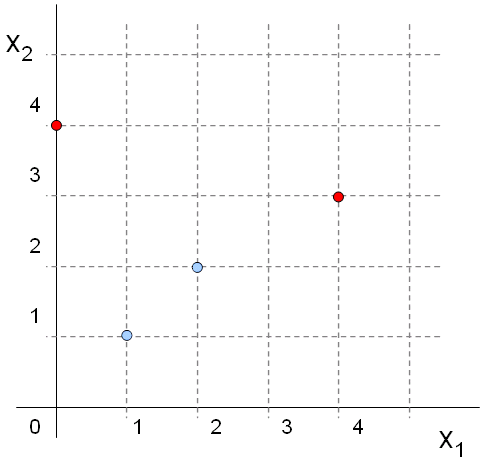
Create a decision tree based on the data listed in the table. The tree should assign value of the dependent variable *flu* based on knowledge of the values of independent variables *temperature* and *joint pain*. Use the principle of minimum entropy to create a binary tree with the least number of tests and zero error on the given data.

3 Perceptron

Solve the problem of classification into two classes. The feature-space is two-dimensional. You have a training set with four examples, two for the class A and two for the class B:

$$T = \{\langle[1,1], A\rangle, \langle[2,2], A\rangle, \langle[0,4], B\rangle, \langle[4,3], B\rangle\} \ .$$

See the figure below and write down any perceptron equation with a minimum classification error on the training set. Draw the corresponding line into the image. What is the training error of the classification?



4 Bayesian Decision Making

Imagine you are a shepherd, whose sheep are constantly attacked by wild beasts. You care about your beloved flock and recall the machine learning course you took at the farming university during childhood.

Three species of predators live in the surrounding forests: bear, wolf and lynx. You know from experience that – from eight cases there are four predator attacks by wolves, three attacks by bears and one by lynx. The wolves are known to attack in pack and usually kill more sheep at once (two or more sheep are killed in 8 out of 10 attacks). Bears are solitary, but may also kill more sheep (two or more sheep are killed in 5 of 10 attacks). Lynx always kills one sheep. Wolves never return to the same place for a new attack next night, a lynx returns only in one out of ten cases, bear returns in half of cases.

Answer the following questions:

(a) What is the probability of an attack, in which a lynx killed more than one sheep?

For the following questions, suppose exactly one of your sheep was killed by a wild animal during the night.

(b) What is the probability that sheep was attacked by bear?

(c) Which predator killed the sheep? Decide merely on the basis of maximum likelihood (with the assumption of uniform a priori probability). Explain the decision.

(d) Which predator killed the sheep? Use the method of maximum a posterior probability. Explain the decision.

(e) What is the probability that a predator will return the next night? Assume the MAP hypothesis about the killer from the previous question and provide a simplified approximation.

(f) What is the probability that a predator will return the next night? Consider full Bayesian learning.