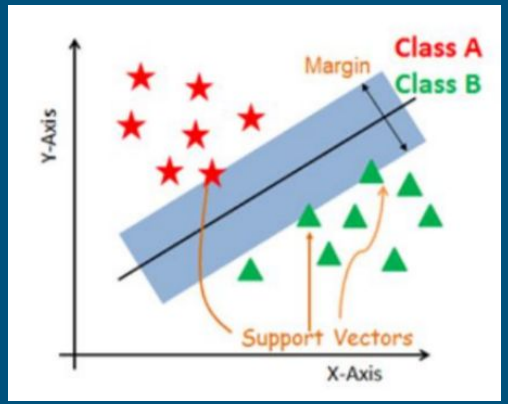
Support Vector Machines (SVM)

Greta Pataki, 2021.11.10.

Support Vector Machines (SVM)

- Supervised learning model
- Classification and regression problems
- Generates an optimal hyperplane in multidimensional space to separate classes
  - Optimal: hyperplane that maximizes the margin (best divides the dataset)
- Instances closest to the hyperplane are the support vectors
  - Define the hyperplane
  - Other instances are not required
- Linear classifier:

$$f(x) = \underline{w}^T \underline{x} + b$$



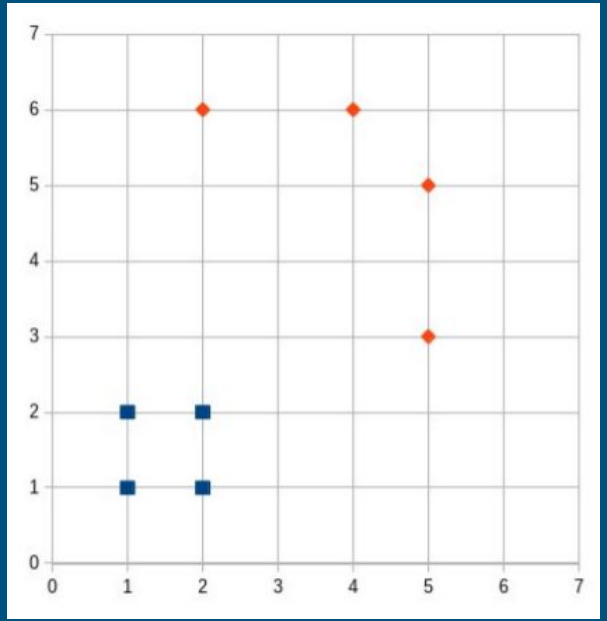
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- find the hyperplane!



Support Vector Machines (SVM)

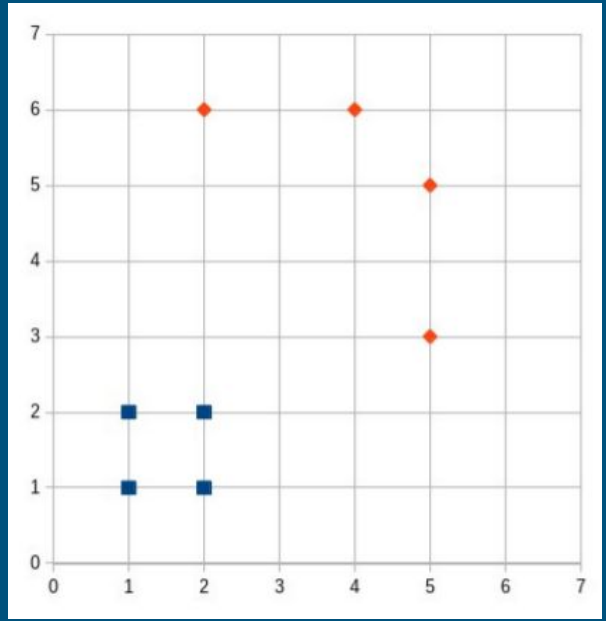
- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- 2 points of the hyperplane:

$$\begin{bmatrix} x \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ y \end{bmatrix}$$



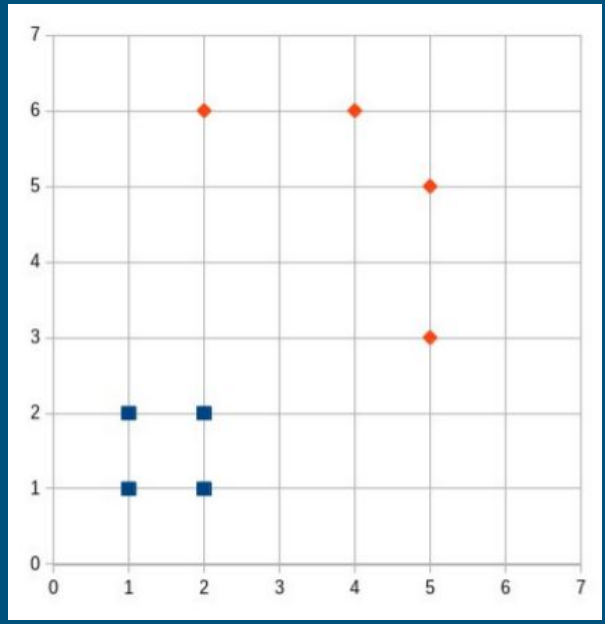
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- 2 points of the hyperplane



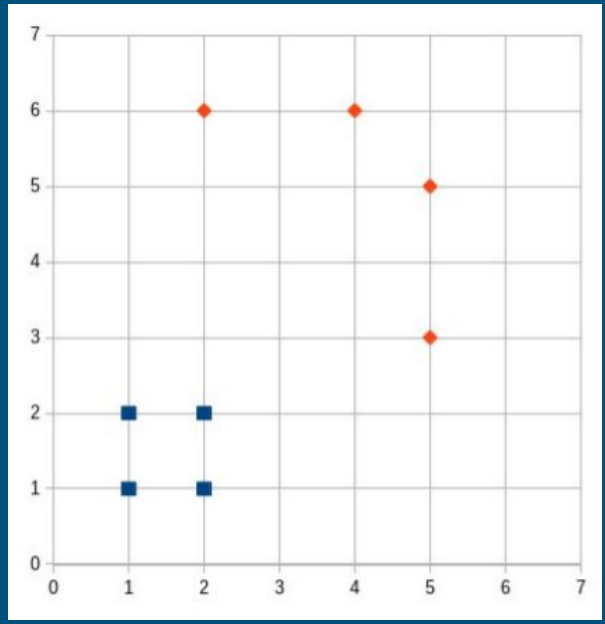
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0



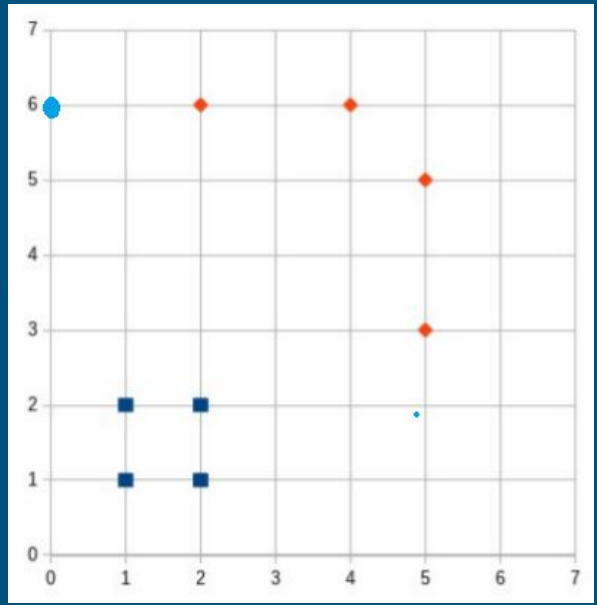
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0



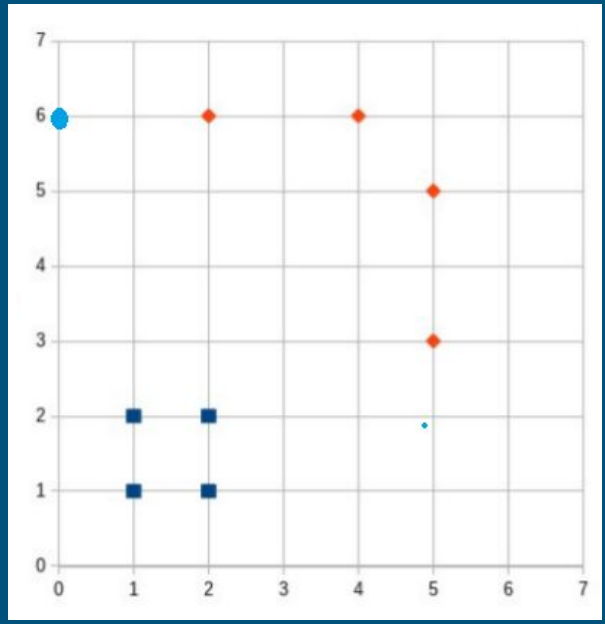
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6



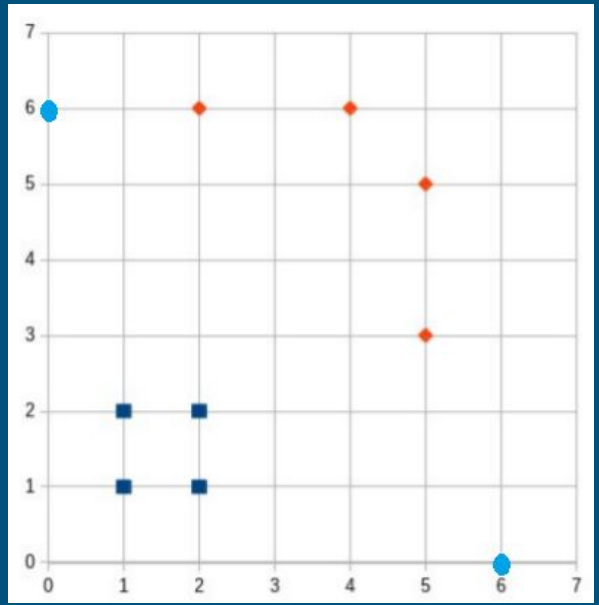
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6



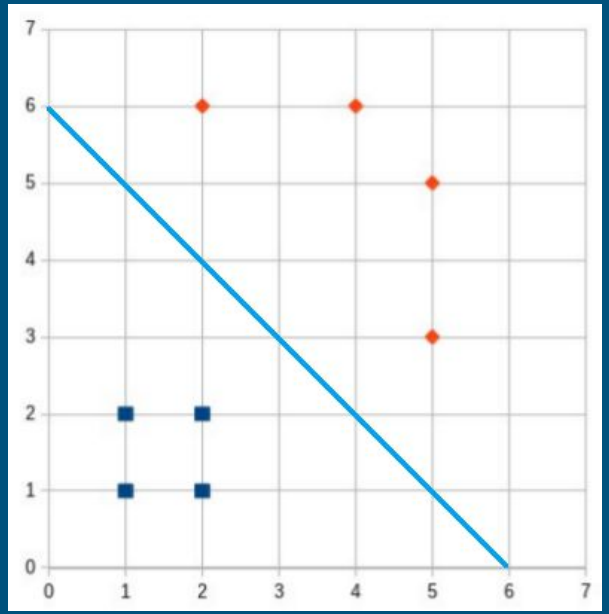
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6



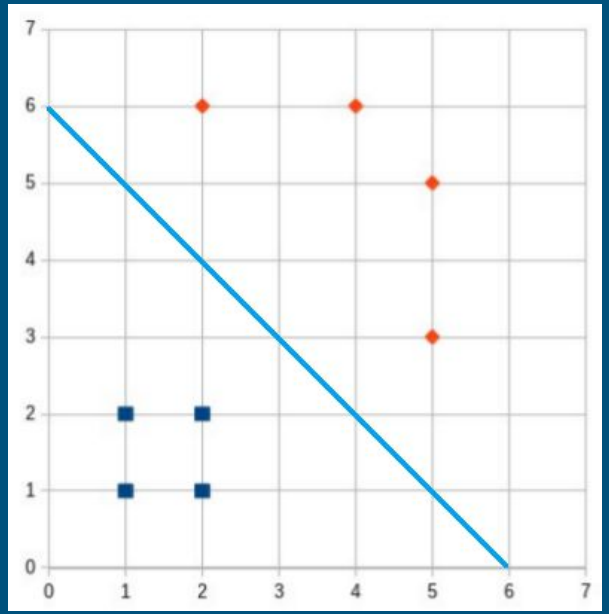
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- mark and list the support vectors!



Support Vector Machines (SVM)

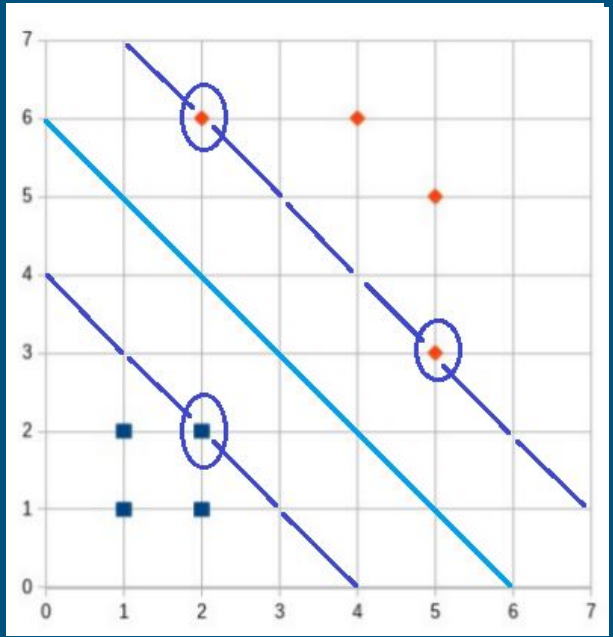
- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- mark and list the support vectors!

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix}$$



Support Vector Machines (SVM)

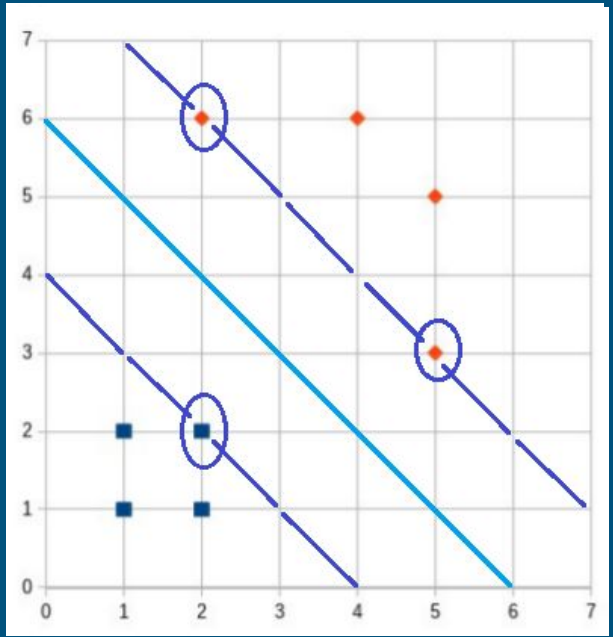
- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- mark and list the support vectors!

$$\ominus \quad \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$



Support Vector Machines (SVM)

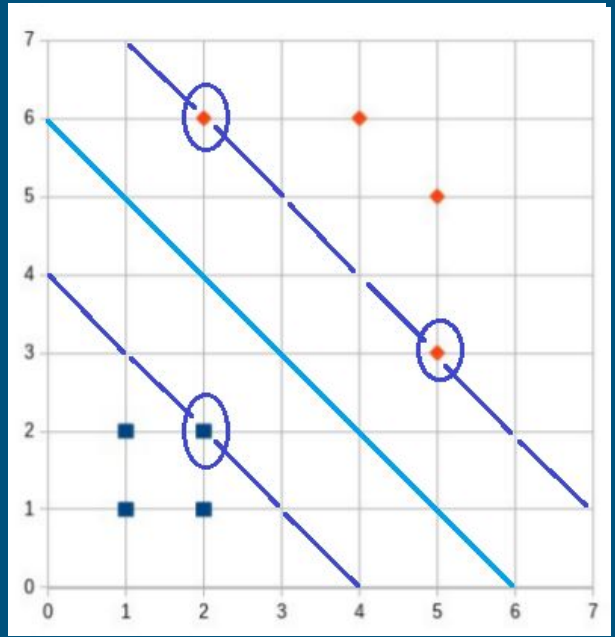
- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- mark and list the support vectors!

$\ominus \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$



Support Vector Machines (SVM)

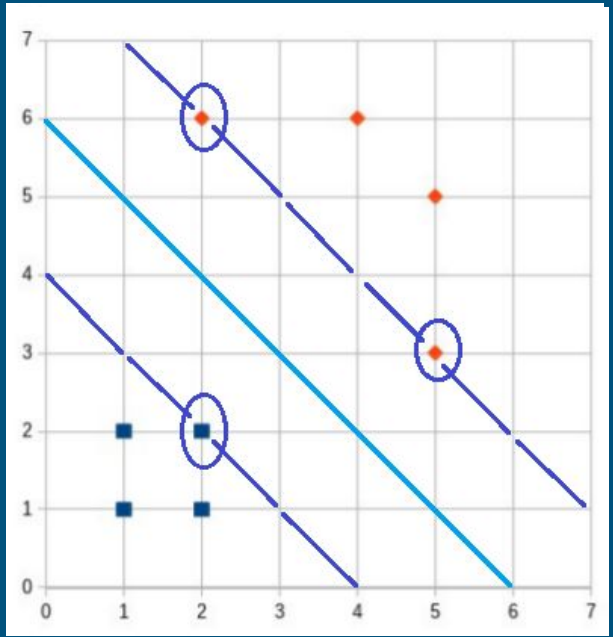
- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- mark and list the support vectors!

$\ominus \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ $\oplus \begin{bmatrix} 2 \\ 6 \end{bmatrix}$



Support Vector Machines (SVM)

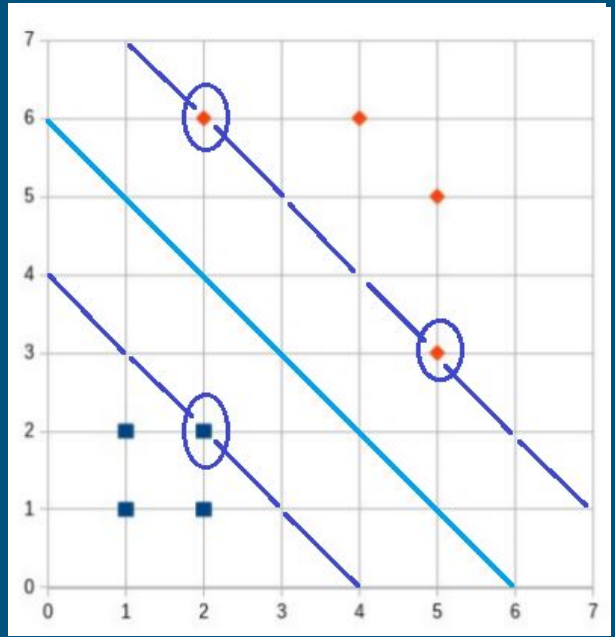
- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- mark and list the support vectors!

$\ominus \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ $\oplus \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ $\begin{bmatrix} 5 \\ 3 \end{bmatrix}$



Support Vector Machines (SVM)

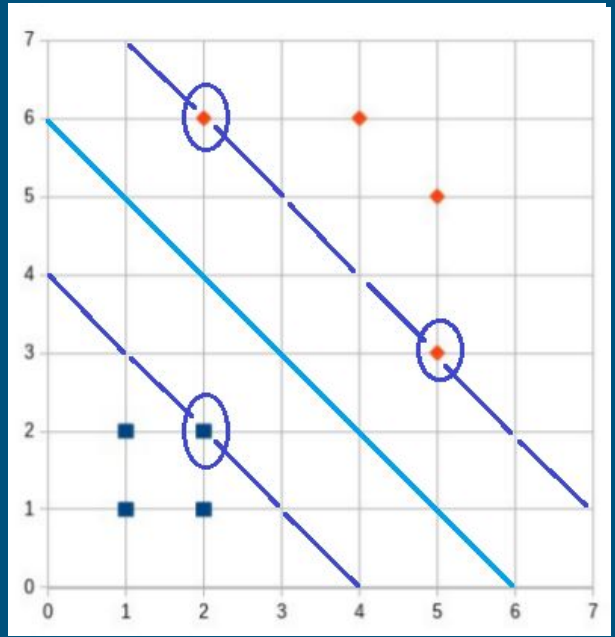
- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- mark and list the support vectors!

$\ominus \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ $\oplus \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ $\oplus \begin{bmatrix} 5 \\ 3 \end{bmatrix}$



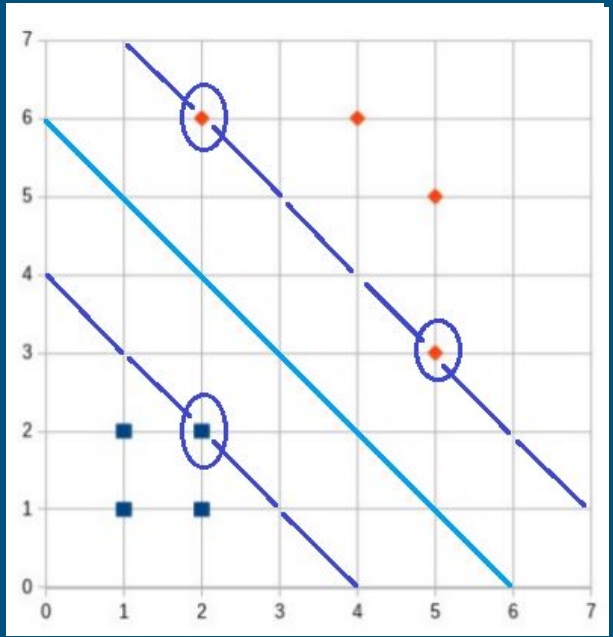
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- why is this hyperplane special?



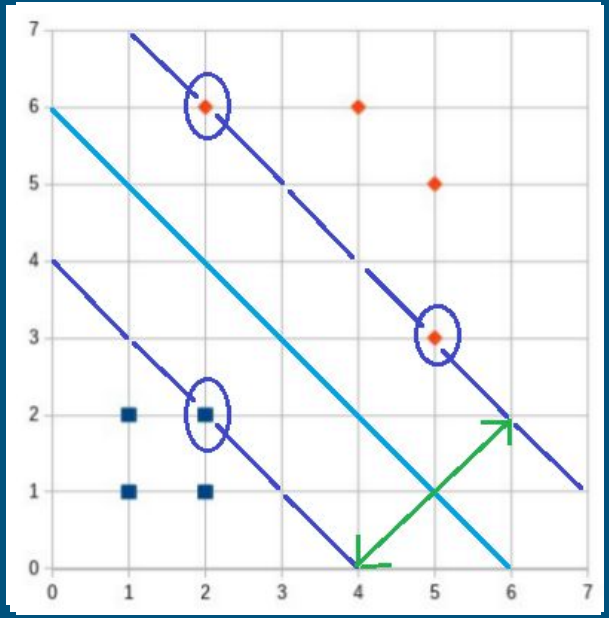
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- why is this hyperplane special?
  - maximizes the margin to the closest instances



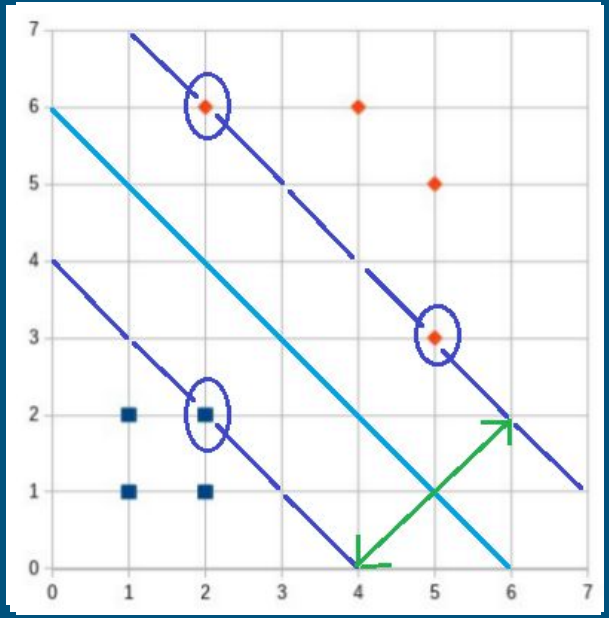
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- which class label belongs to which class?



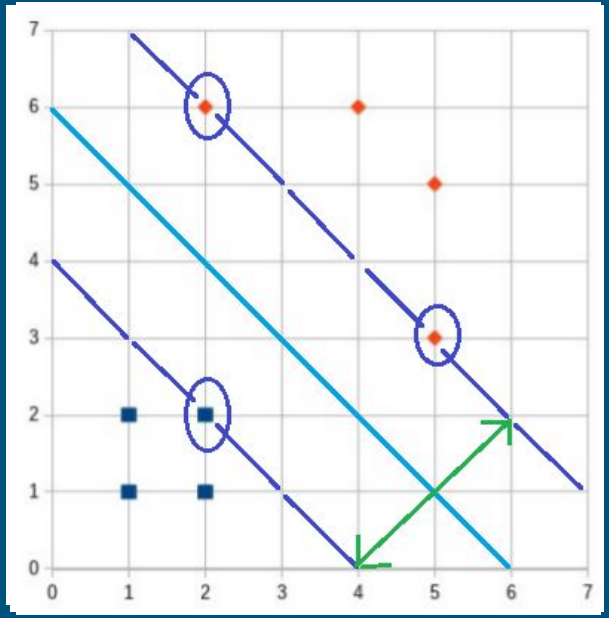
Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$

- which class label belongs to which class?
  - negative
  - positive

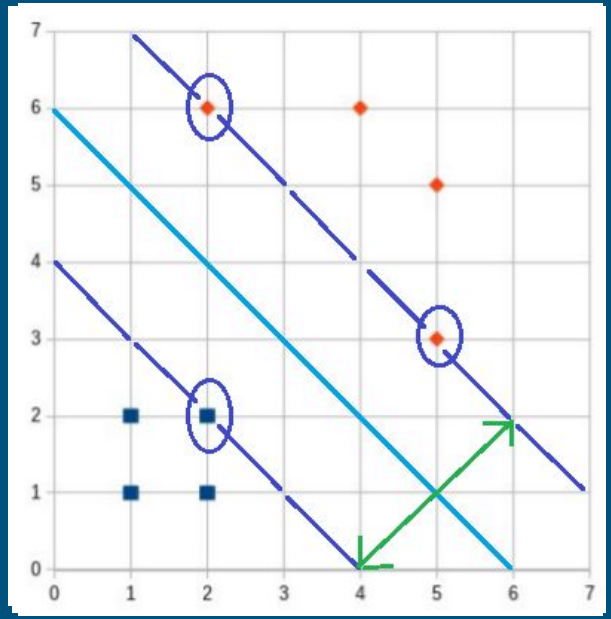


Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- predictions for the following instances:

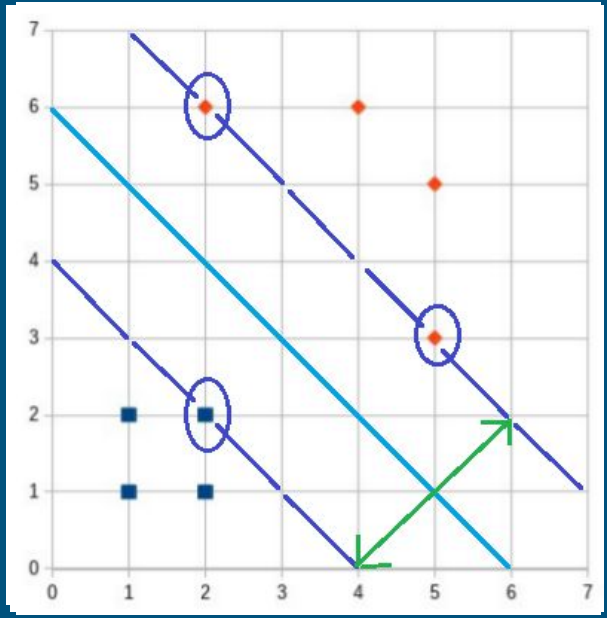
$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 60 \\ 40 \end{bmatrix} \quad \begin{bmatrix} 1000 \\ -998 \end{bmatrix}$$

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

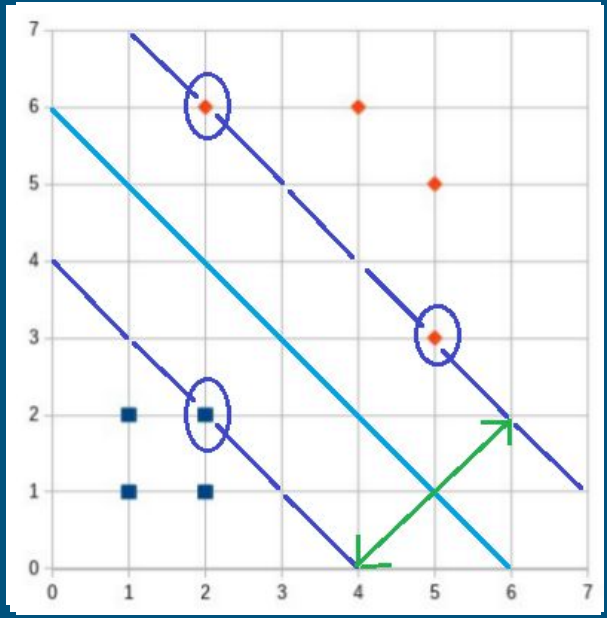
$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} =$$

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

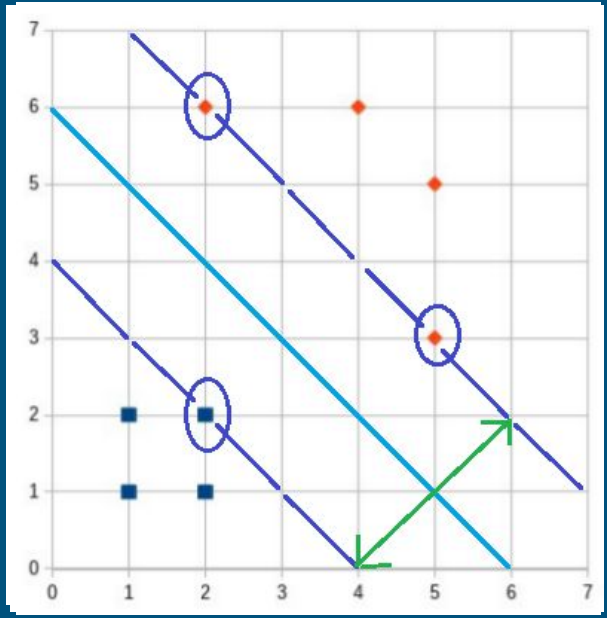
$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$ = 0.5

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

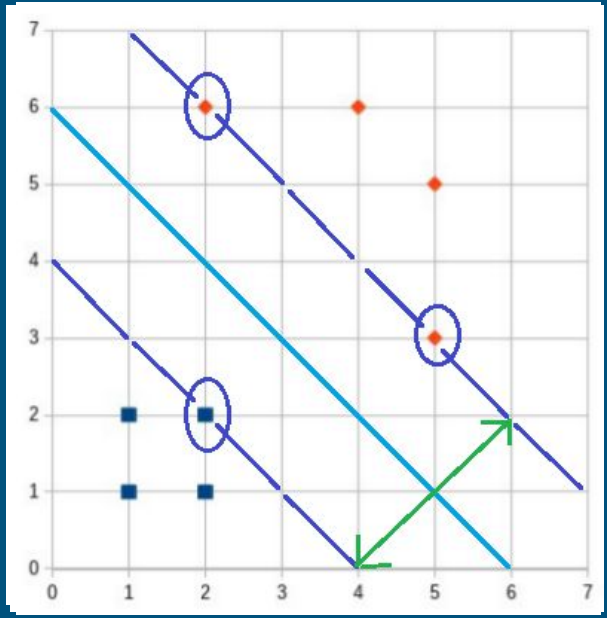
= 0.5 * 4

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

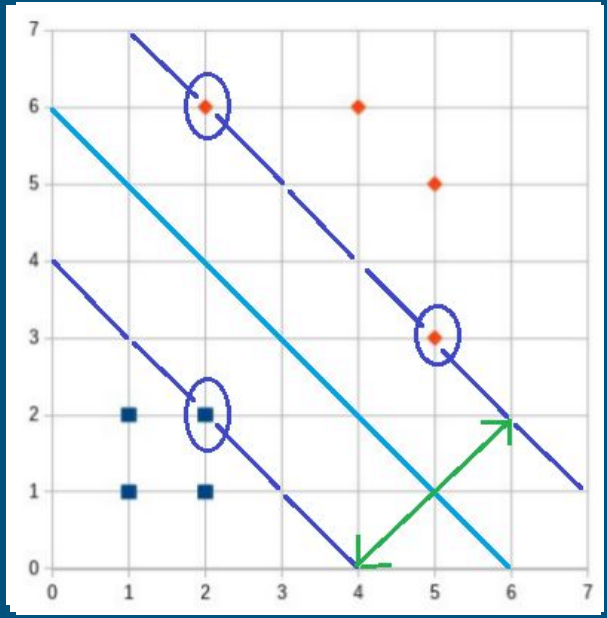
= 0.5 * 4 +

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

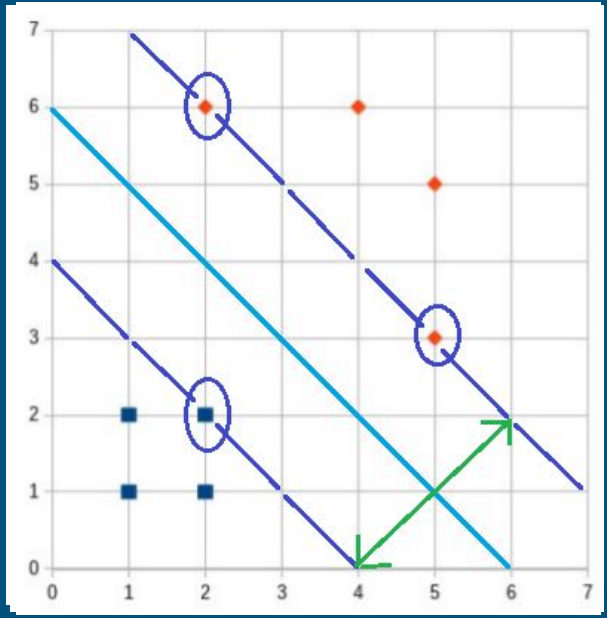
= 0.5 * 4 + 0.5

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

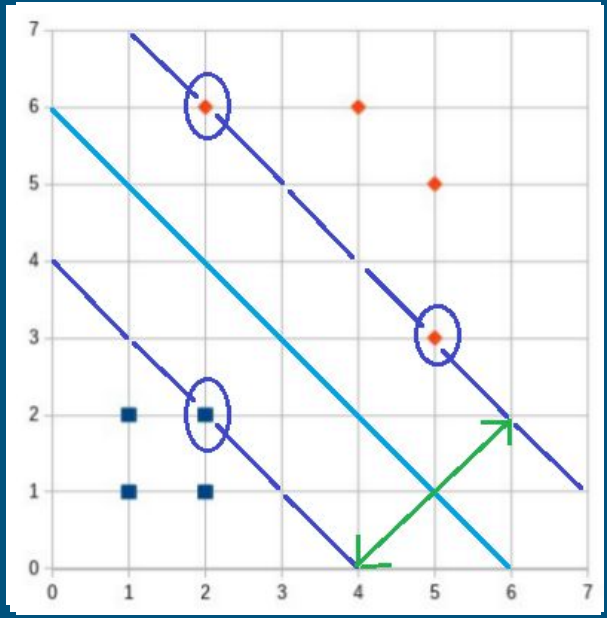
= 0.5 * 4 + 0.5 * 1

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

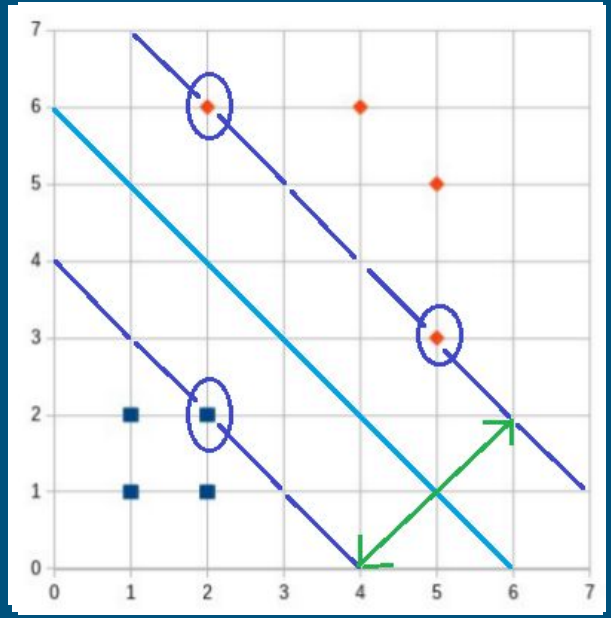
= 0.5 * 4 + 0.5 * 1 - 3 = - 0.5

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

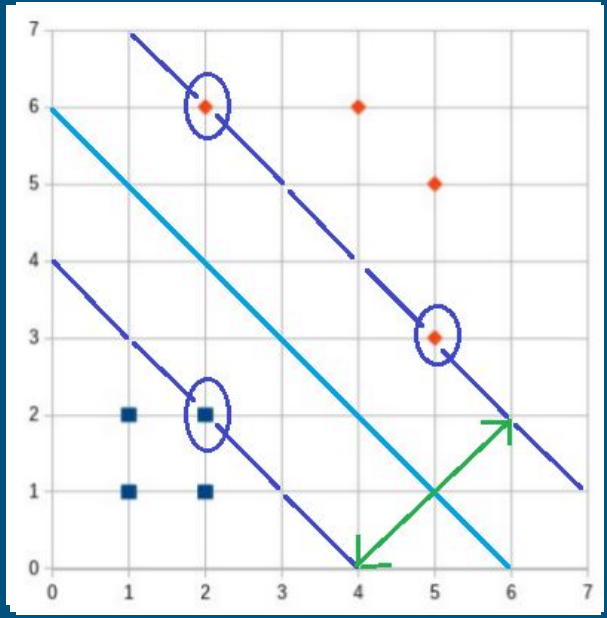
= 0.5 * 4 + 0.5 * 1 - 3 = - 0.5 => negative

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

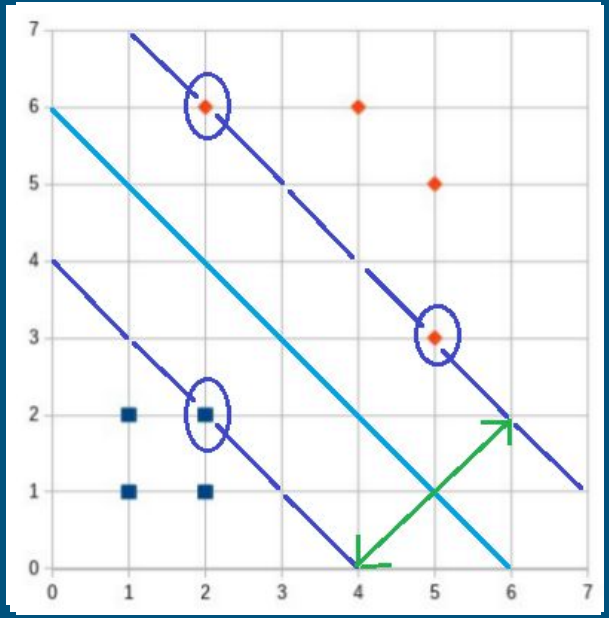
$$\begin{bmatrix} 60 \\ 40 \end{bmatrix} =$$

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 60 \\ 40 \end{bmatrix}$$

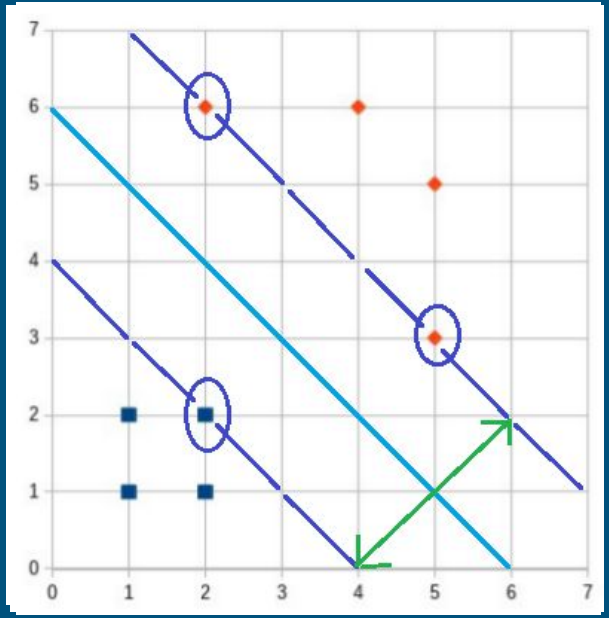
= 0.5 * 60 + 0.5 * 40 - 3 = 47

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 60 \\ 40 \end{bmatrix}$$

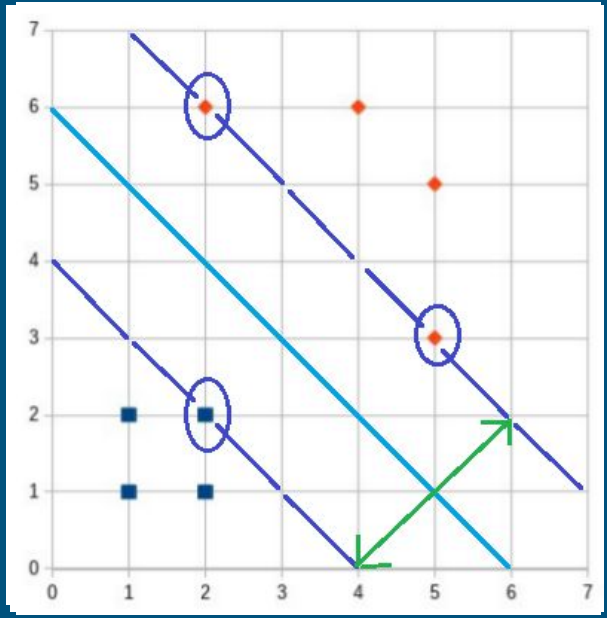
= 0.5 * 60 + 0.5 * 40 - 3 = 47 => positive

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

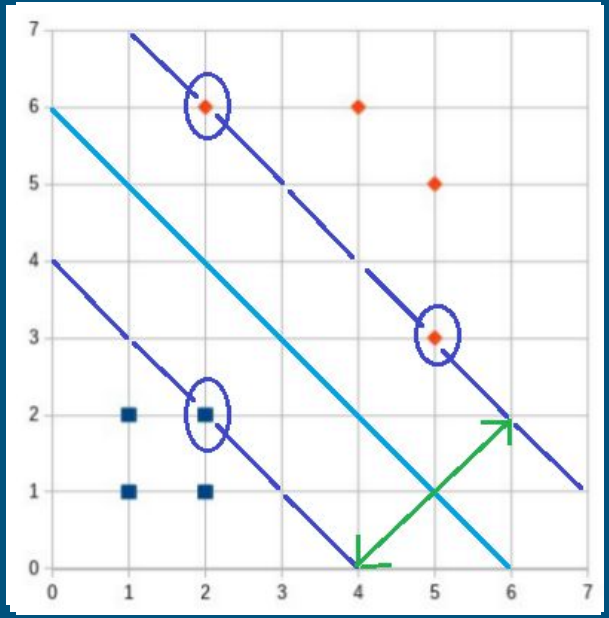
$$\begin{bmatrix} 1000 \\ -998 \end{bmatrix} =$$

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

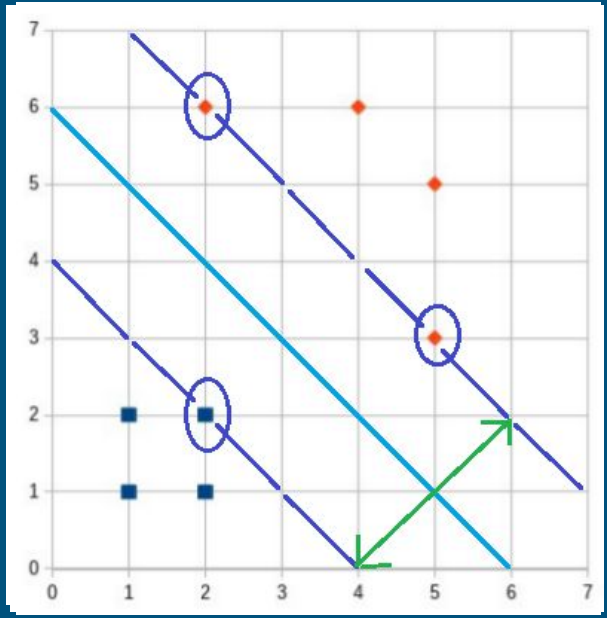
$$\begin{bmatrix} 1000 \\ -998 \end{bmatrix}$$ = 0.5 * 1000 + 0.5 * -998 - 3 = -2

Support Vector Machines (SVM)

- 8 instances
- 2 classes
- Weka:

$$\underline{w} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, b = -3.$$

$$f(x) = \underline{w}^T \underline{x} + b$$



- 2 points of the hyperplane
  - 0.5 x + 0.5 * 0 + b = 0 => x = 6 y = 0
  - 0.5 y + 0.5 * 0 + b = 0 => x = 0 y = 6
- predictions for the following instances:

$$\begin{bmatrix} 1000 \\ -998 \end{bmatrix}$$ = 0.5 * 1000 + 0.5 * -998 - 3 = -2 => negative

Support Vector Machines (SVM)

- sklearn: svm.SVC
- https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html
- parameters:
  - C (float, default=1.0): regularization parameter. The strength of the regularization is inversely proportional to C. Must be strictly positive.
  - kernel {'linear', 'poly', 'rbf', 'sigmoid', 'precomputed', callable}, default='rbf'
    - specifies the kernel type to be used in the algorithm
  - gamma {'scale', 'auto'} or float, default='scale': Kernel coefficient for 'rbf', 'poly' and 'sigmoid'.
- attributes:
  - coef_: weights assigned to the features when kernel="linear"
  - intercept_: constants in decision function.
  - support_vectors_
  - n_support_: number of support vectors for each class