

Support Vector Machines (SVM)

Angelo Parker

MATH 3220

Data Mining Methods

Table of Contents

- A Brief History of SVMs
- What are SVMs?
- Examples of SVMs
- Pros of SVMs
- Cons of SVMs

A Brief History

- The original SVM algorithm was created in 1963 by Vapnik and Chervonenkis for binary classification. Only able to conduct linear separation.
- In 1992, Vapnik, Boser, and Guyon added a method to create non-linear separation.
- In 1995, the most referred method, was finalized by Vapnik and Cortes.

What are SVMs

- SVMs are binary graphical classification models that use regression lines to separate and push data points closer to each other into more distinct groups.

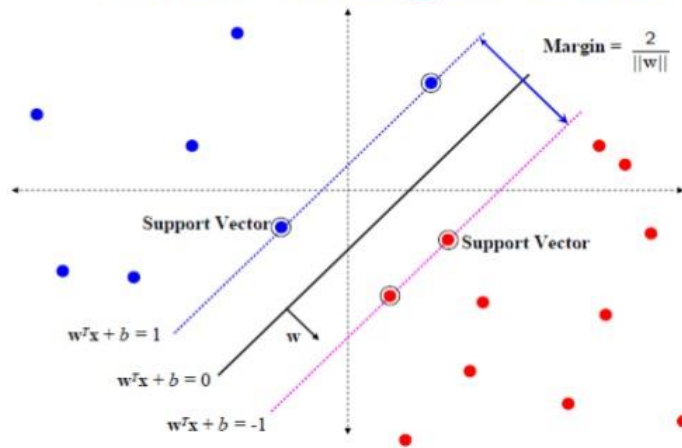
$$\vec{w} \cdot \vec{x} - b = 0,$$

$$\vec{w} \cdot \vec{x} - b = -1.$$

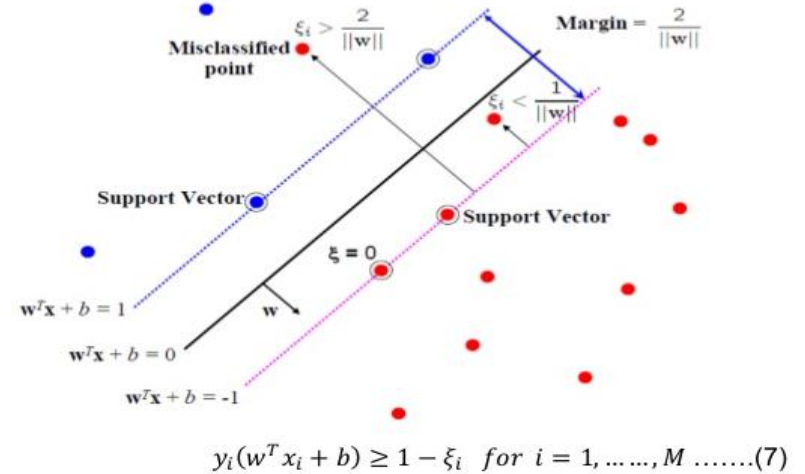
$$\vec{w} \cdot \vec{x} - b = 1$$

- SVMs are in the e1071 package in R.

Hard-margin SVM



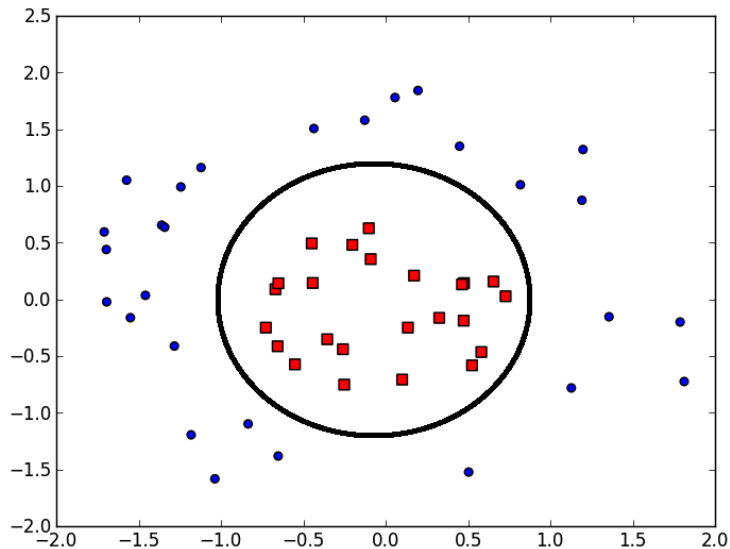
Soft-margin SVM



$$y_i(\vec{w} \cdot \vec{x}_i - b) \geq 1, \quad \text{for all } 1 \leq i \leq n.$$

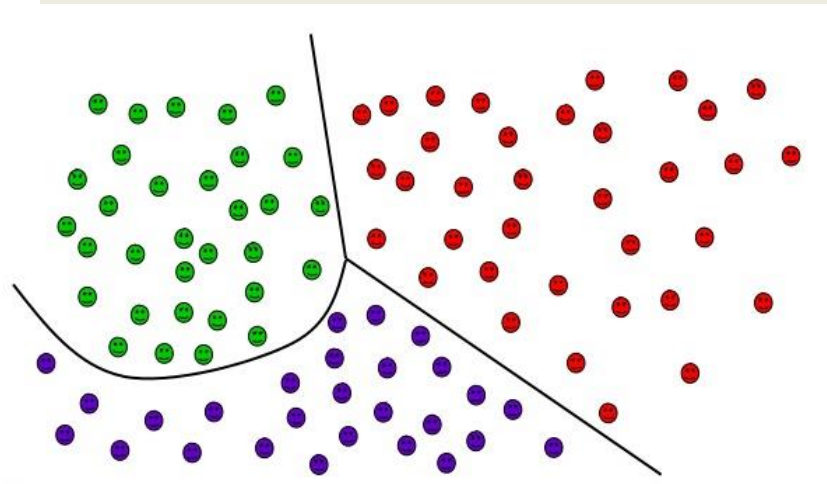
$$\max(0, 1 - y_i(\vec{w} \cdot \vec{x}_i - b)).$$

$$\left[\frac{1}{n} \sum_{i=1}^n \max(0, 1 - y_i(\vec{w} \cdot \vec{x}_i - b)) \right] + \lambda \|\vec{w}\|^2,$$



$$k(\vec{x}_i, \vec{x}_j) = \exp(-\gamma \|\vec{x}_i - \vec{x}_j\|^2)$$

$$\gamma > 0$$



1. $y_i (\vec{w} \cdot \vec{x}_i - b) \geq 1$, for all $1 \leq i \leq n$.
2. $y_i (\vec{w} \cdot \vec{x}_i - b) \geq 1$, for all $1 \leq i \leq n$.
3. $y_i (\vec{w} \cdot \vec{x}_i - b) \geq 1$, for all $1 \leq i \leq n$.

Pros of SVMs

- Can designate outliers.
- Can be applied to text mining including hand- written text.
- Can be used on multivariate data. Will create planes or hyper planes to adjust.
- Can create confusion matrices to showcase accuracy or “fitness”.

Cons of SVMs

- All data must be already classes/labeled. Support Vector Clustering can be used for unsupervised learning to set the stage for SVMs.
- SVMs only work on a binary level. Multiple SVMs will have to be created then plotted or another algorithm must be used to condense a problem for SVM to work.
- Does not give probability of data or new data.

Review

- A Brief History of SVMs
- What are SVMs?
- Examples of SVMs
- Pros of SVMs
- Cons of SVMs

Sources

- Guduru, N. (2006). *Text mining with support vector machines and non-negative matrix factorization algorithms* (Unpublished master's thesis).
- Meyer, D. (2017). *Support Vector Machines*. CRAN.
- Cortes, C.; Vapnik, V. (1995). "Support-vector networks". *Machine Learning*. **20** (3): 273–297. doi:10.1007/BF00994018.