# Project3: Support Vector Machine

**Group Members:**
Annine Duclaire Kenne
Idriss Nguepi Nguefack
Mohamed Elabbas
Mouhamadou Mansour Sow

African Institute for Mathematical Sciences, AIMS-Senegal

Supervised by: Tutors

April 22, 2022

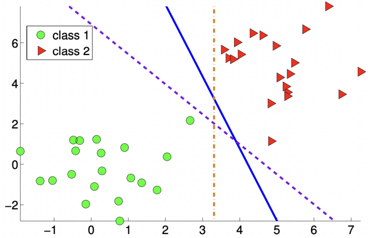
# Presentation outline

1. Introduction
   - Motivation
   - Problem
   - Objective
2. Methodology
3. Results and discussion
4. Conclusion

# Motivation

Support Vector Machine (SVM) is a discriminant algorithm used to classify data points in different classes. In our case, we will use it in case of sentiment analysis. In other words, we will use SVM to classify whether a sentiment related to a movie is positive or negative based on reviews that we have in our dataset.

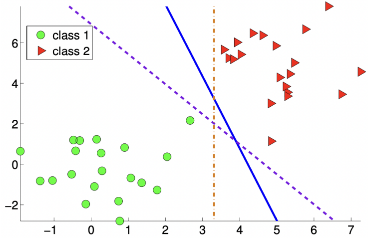
# Problem




**Details**

- Find the right support vector
- Compute the margin using the supports
- classify each data point

# Problem




## Details

- Find the right support vector
- Compute the margin using the supports
- classify each data point

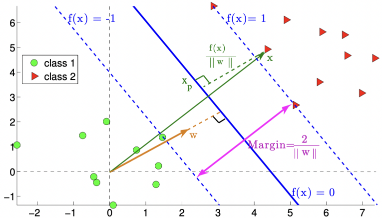
# Objective

## Objective

Implementing the optimized form of Support Vector Machine (SVM).

# Methodology

Let $\mathcal{D} = \{(x_i, y_i) \in \mathcal{X} \times \{-1, 1\}\}$ the set of labeled points

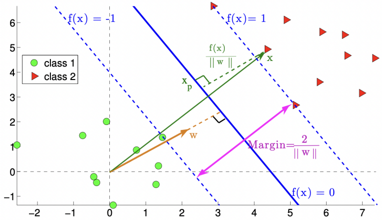



**Details**

- $f(x) = w^T.x + b$
- The margin is equal to $M = \frac{2}{\|w\|}$

# Methodology

Let $\mathcal{D} = \{(x_i, y_i) \in \mathcal{X} \times \{-1, 1\}\}$ the set of labeled points




**Details**

- $f(x) = w^T.x + b$
- The margin is equal to $M = \frac{2}{\|w\|}$

# Methodology

We are going to maximize the margin $M = \frac{2}{||w||}$ to be large as possible.

## Primal problem

$$\begin{cases} min_{w,b}\frac{1}{2}||w||^2 \\ s.t\ y_i(w^Tx_i + b) \geq 1\ \forall i \end{cases} \quad (Primal\ \ problem) \tag{2.1}$$

Let derive the dual problem from (2.1)
The Lagrangian associated to the primal problem (2.1) is given by:

## Lagrangian

$$L(w, b, \alpha) = \frac{1}{2}||w||^2 - \sum_{i}^{n} \alpha_i \left[y_i(w^Tx_i + b) - 1\right] \tag{2.2}$$

# Methodology

## Partial derivative

$$\frac{\partial L}{\partial w} = w - \sum_i^n \alpha_i y_i x_i \quad (2.3)$$

$$\frac{\partial L}{\partial b} = -\sum_i^n \alpha_i y_i \quad (2.4)$$

## Solve $\partial L = 0$

$$\partial L = 0 \Longrightarrow \begin{cases} w = \sum_i^n \alpha_i y_i x_i \ (a) \\ \sum_i^n \alpha_i y_i = 0 \ (b) \end{cases}$$

# Methodology

## Partial derivative

$$\frac{\partial L}{\partial w} = w - \sum_{i}^{n} \alpha_i y_i x_i \tag{2.3}$$

$$\frac{\partial L}{\partial b} = - \sum_{i}^{n} \alpha_i y_i \tag{2.4}$$

## Solve $\partial L = 0$

$$\partial L = 0 \Longrightarrow \begin{cases} w = \sum_i^n \alpha_i y_i x_i \ (a) \\ \sum_i^n \alpha_i y_i = 0 \ (b) \end{cases}$$

AIMS African Institute for Mathematical Sciences SENEGAL

# Methodology

**Dual problem**

(b) - (a) in 2.2 We have

$$L = -\frac{1}{2}\sum_{i}^{n}\sum_{j}^{n}\alpha_j y_j x_j^T y_i x_i \alpha_i + \sum_{i}^{n}\alpha_i \tag{2.5}$$

**Dual problem**

Let $Q = (Q_{ij})$ where $Q_{ij} = y_j y_i x_j^T x_i$

$$\begin{cases} L = -\frac{1}{2}\alpha^T Q\alpha + 1^T\alpha \\ s.t\ y^T\alpha = 0\ and\ \alpha \geq 0 \end{cases} \tag{2.6}$$

# Methodology

## Dual problem

(b) - (a) in 2.2 We have

$$L = -\frac{1}{2}\sum_{i}^{n}\sum_{j}^{n}\alpha_j y_j x_j^T y_i x_i \alpha_i + \sum_{i}^{n}\alpha_i \tag{2.5}$$

## Dual problem

Let $Q = (Q_{ij})$ where $Q_{ij} = y_j y_i x_j^T x_i$

$$\begin{cases} L = -\frac{1}{2}\alpha^T Q\alpha + 1^T\alpha \\ s.t\ y^T\alpha = 0\ and\ \alpha \geq 0 \end{cases} \tag{2.6}$$

# Methodology

## Dual problem

So to find $w$ and $b$ we need first to find the value of $\alpha$, it should be the solution of the optimization problem below.

$$\begin{cases} min_{\alpha} \frac{1}{2}\alpha^T Q\alpha - 1^T\alpha \\ s.t\ y^T\alpha = 0\ and\ \alpha \geq 0\ (Dual\ \ problem) \end{cases} \tag{2.7}$$

# Methodology

Lets calculate b.

## Compute b

$$y_s(wx_s + b) = 1 \quad (2.8)$$

By substituting (a) in (2.8) :

$$y_s(\sum_{m \in S} \alpha_m y_m x_m . x_s + b) = 1 \quad (2.9)$$

# Methodology

**compute b**

$$y_s^2\left(\sum_{m\in S}\alpha_m y_m x_m . x_s + b\right) = y_s \tag{2.10}$$

with $y_s^2 = 1$

$$b = \frac{1}{N_s}\sum_{s\in S}\left(y_s - \sum_{m\in S}\alpha_m y_m x_m . x_s\right) \tag{2.11}$$

# Results and discussion

# Conclusion

AIMS African Institute for Mathematical Sciences SENEGAL

# References

Jurasfky, Daniel and Martin, James H, "An introduction to natural language processing, computational linguistics, and speech recognition", 2000, Pearson Education, Inc

**THANKS FOR YOUR**

**KIND**

**ATTENTION !!!**