

Université Sciences et Technologie HOUARI BOUMEDIENE

Faculté d'informatique

Département d'Intelligence Artificielle et Science des données



Apprentissage Automatique et Réseaux de neurones

MASTER 1 SYSTÈME INFORMATIQUE INTELLIGENT

---

TP7 : Les SVMs

---

Réalisé par:

- Merrouche Amira
- Djeghali ikram
- Ait oubelli syphax
- Samy moussi

## Introduction

Dans ce rapport de TP nous explorons le classifieur machine à vecteurs de support. Nous essayerons d'utiliser plusieurs noyaux et voir la différence entre leur capacité de séparation.

### 1. Machine à vecteurs de support

Les **machines à vecteurs de support** (SVM) sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination, c'est-à-dire décider à quelle classe appartient un échantillon, ou de régression, c'est-à-dire prédire la valeur numérique d'une variable.

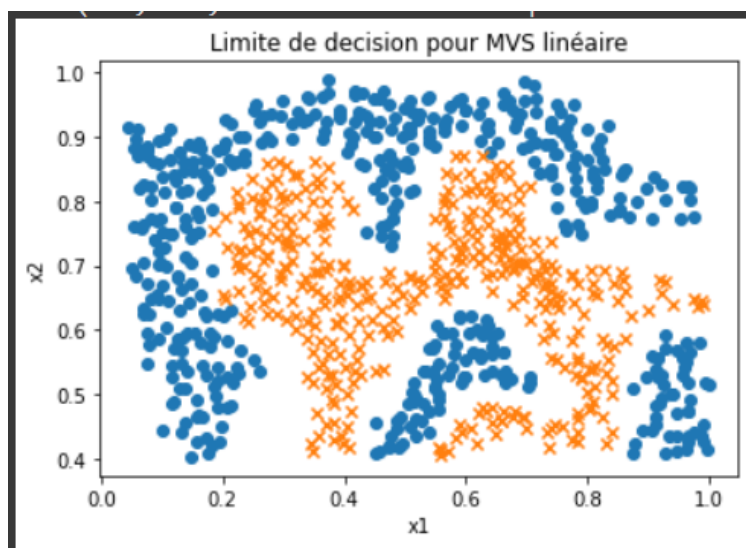
Le principe des SVM consiste à ramener un problème de classification ou de discrimination à un hyperplan (feature space) dans lequel les données sont séparées en plusieurs classes dont la frontière est la plus éloignée possible des points de données (ou "marge maximale").

Dans ce TP nous avons un ensemble de données auxquelles nous allons essayer d'appliquer et comparer plusieurs noyaux (Linéaire, Polynomial ainsi que RBF (noyau Gaussien) ).

### 2. Les Noyaux Utilisés

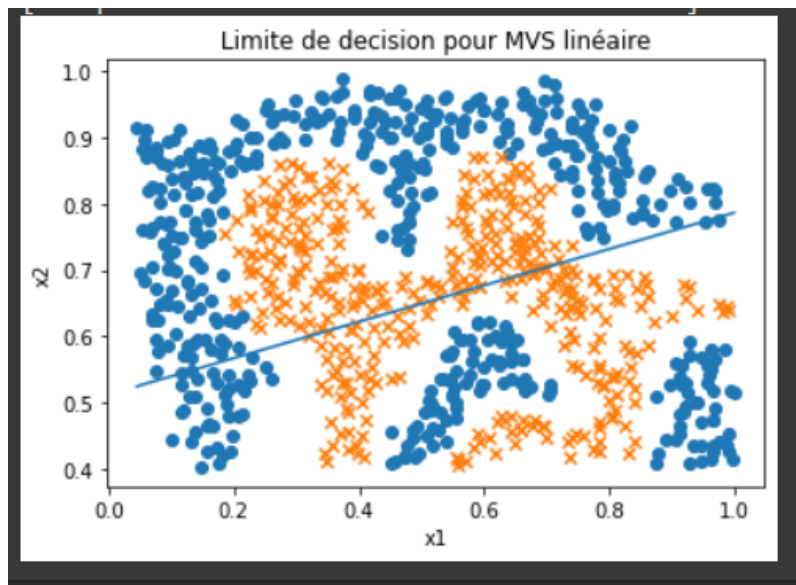
Nous commençons d'abord par afficher nos données

- **Visualisation Des Données**



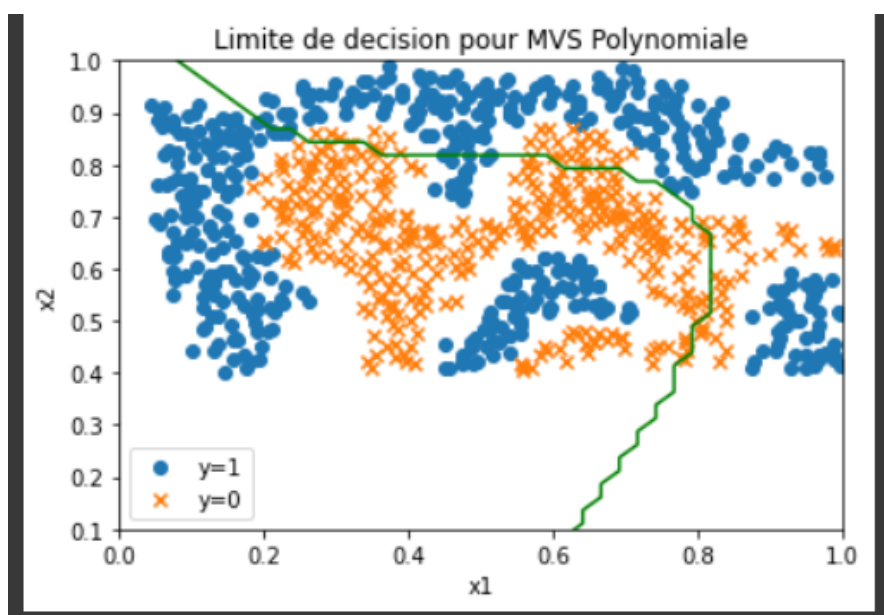
D'après la distribution des données nous pouvons déjà imaginer que les séparateurs linéaires et polynomial nous donneront de mauvais résultats à cause de la distribution des données.

- **Application du noyau linéaire**



Comme nous pouvons le constater le MVS linéaire nous donne un mauvais résultat et ceci est dû au fait que nos données ne sont pas linéairement séparables.

- **Application du noyau polynomial**



- **Application du noyau RBF (noyau Gaussien)**

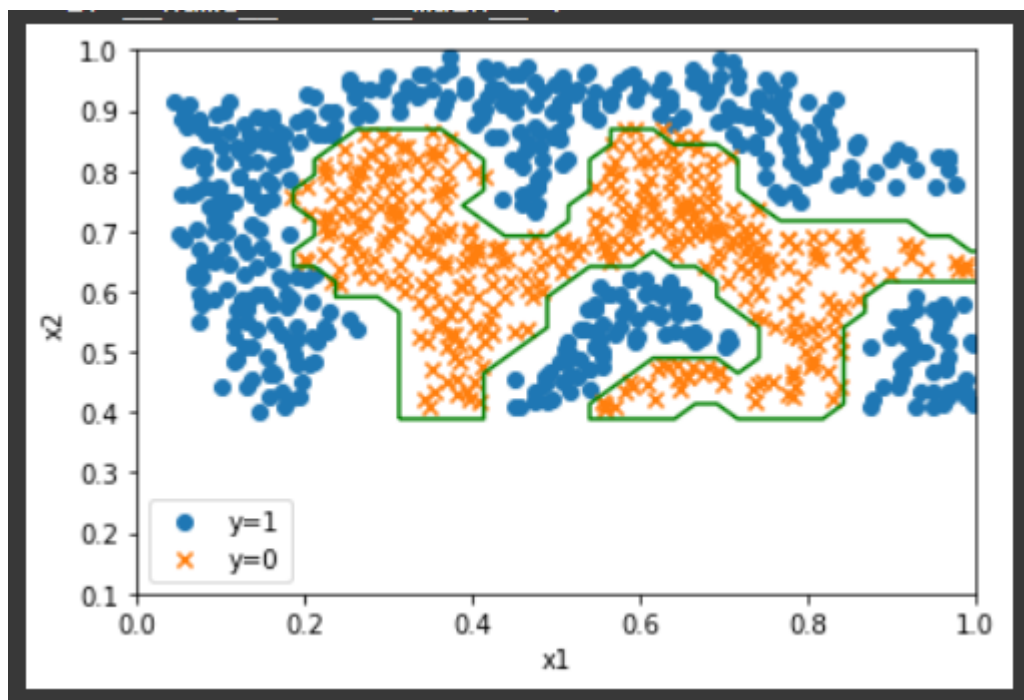
Pour ce noyau nous avons cherché quels seraient les paramètres optimaux à savoir *C* et *Gamma* qui nous permettraient d'avoir le meilleur résultat possible.

```
recherche des meilleurs parametres pour notre classificateur SVM
```

```
1 C_range = 10. ** np.arange(-3, 8)
2 gamma_range = 10. ** np.arange(-5, 4)
3
4
5 param_grid = dict(gamma=gamma_range, C=C_range)
6
7
8 grid = GridSearchCV(SVC(), param_grid=param_grid)
9
10 grid.fit(X, y)
11 print("Les meilleurs parametres pour notre svc sont: ", grid.best_estimator_)
```

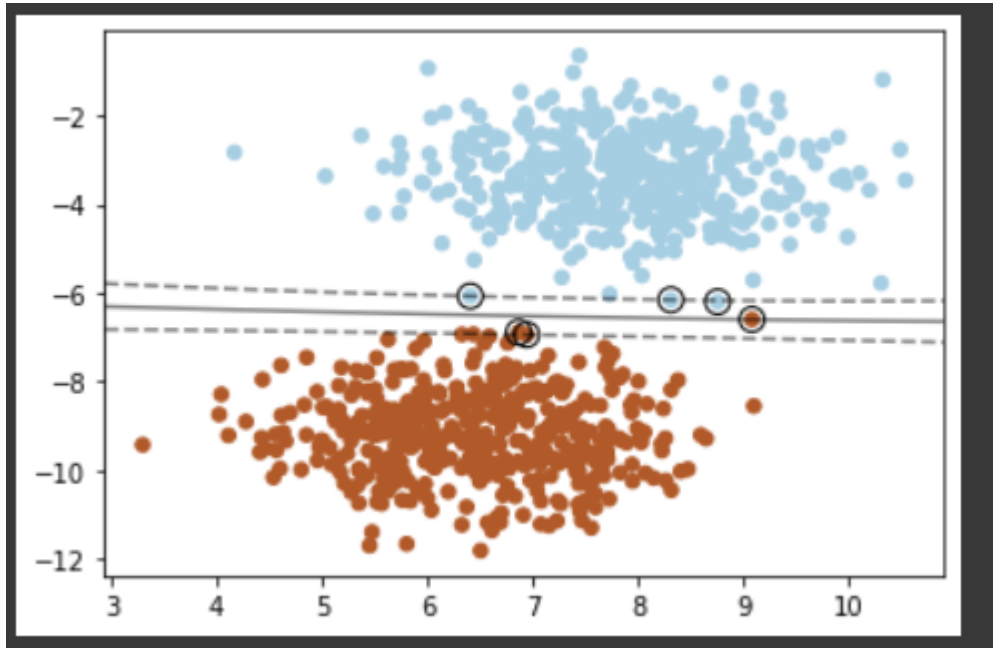
Les meilleurs parametres pour notre svc sont: SVC(C=10.0, gamma=1000.0)

Après avoir testé différentes valeurs de *C* et *Gamma* nous avons obtenu que pour ces données les meilleurs valeurs seraient :  
*Gamma* =1000 et *C* =10.



- **Application du noyau RBF (noyau Gaussien) sur d'autres données**

Nous avons testé notre noyau Gaussien sur d'autres données en gardant les paramètres optimaux pour les données précédentes



Nous pouvons remarquer que le résultat obtenu est assez bon.

### 3. Conclusion

En conclusion, l'efficacité d'un SVM dépend de la fonction du noyau et des paramètres du noyau car ces derniers déterminent l'équation de l'hyperplan qui représente le classifieur des données et par conséquent c'est ce qui détermine la précision des prédictions de notre SVM.