

## C15 Algorithme des $k$ plus proches voisins

### Principe de l'algorithme

L'algorithme de  $k$  plus proches voisins est un algorithme de classification des données.

## C15 Algorithme des $k$ plus proches voisins

### Principe de l'algorithme

L'algorithme de  $k$  plus proches voisins est un algorithme de classification des données.

On dispose d'un jeu de données qui associe chaque donnée à une classe.

## C15 Algorithme des $k$ plus proches voisins

### Principe de l'algorithme

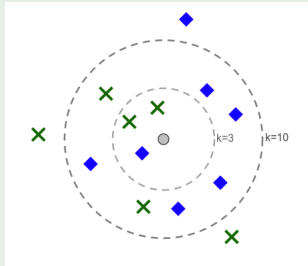
L'algorithme de  $k$  plus proches voisins est un algorithme de classification des données.

On dispose d'un jeu de données qui associe chaque donnée à une classe.

L'algorithme attribut à une nouvelle donnée  $d$  non classée la classe majoritaire de ses  $k$  plus proches voisins.

## C15 Algorithme des $k$ plus proches voisins

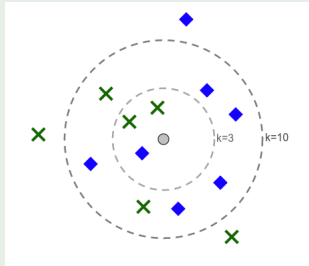
### Exemple



Le point gris central est la donnée à classer. Quel sera le résultat de l'algorithme :

## C15 Algorithme des $k$ plus proches voisins

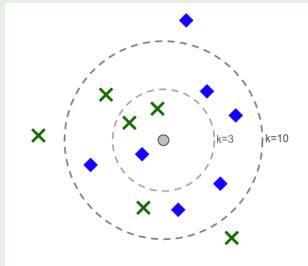
### Exemple



Le point gris central est la donnée à classer. Quel sera le résultat de l'algorithme :  
Pour  $k = 3$  ?

## C15 Algorithme des $k$ plus proches voisins

### Exemple

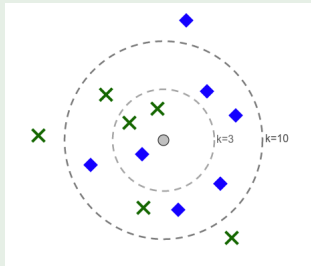


Le point gris central est la donnée à classer. Quel sera le résultat de l'algorithme :

Pour  $k = 3$ ? Il y a 2 croix et un losange dans les 3 plus prochains voisins, la classe majoritaire est donc la croix et l'algorithme classe la donnée comme une croix.

## C15 Algorithme des $k$ plus proches voisins

### Exemple



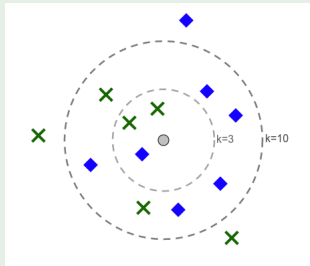
Le point gris central est la donnée à classer. Quel sera le résultat de l'algorithme :

Pour  $k = 3$ ? Il y a 2 croix et un losange dans les 3 plus prochains voisins, la classe majoritaire est donc la croix et l'algorithme classe la donnée comme une croix.

Pour  $k = 10$ ? Cette fois il y a 6 losanges et 4 croix parmi les 10 plus proches voisins, la donnée est donc classée parmi les losanges.

## C15 Algorithme des $k$ plus proches voisins

### Exemple



Le point gris central est la donnée à classer. Quel sera le résultat de l'algorithme :

Pour  $k = 3$ ? Il y a 2 croix et un losange dans les 3 plus prochains voisins, la classe majoritaire est donc la croix et l'algorithme classe la donnée comme une croix.

Pour  $k = 10$ ? Cette fois il y a 6 losanges et 4 croix parmi les 10 plus proches voisins, la donnée est donc classée parmi les losanges.



## C15 Algorithme des $k$ plus proches voisins

### Remarques

Sur l'exemple précédent, on a utilisé la distance euclidienne dans le plan.  
D'autres distances sont envisageables.

## C15 Algorithme des $k$ plus proches voisins

### Remarques

Sur l'exemple précédent, on a utilisé la distance euclidienne dans le plan.  
D'autres distances sont envisageables.

Le nombre de  $k$  de voisins considéré influence la prédiction de l'algorithme  
(voir exemple précédent)