

Programmation d'une Intelligence Artificielle :
arbres ou apprentissage automatique
Apprentissage Semi Supervisé.

Paul Mangold

L3 MIASHS

9 avril 2021

Machine Learning Semi Supervisé

On a considéré deux cas :

- données avec des labels : **machine learning supervisé** ;
- données sans labels : **machine learning non supervisé**.

On a considéré deux cas :

- données avec des labels : **machine learning supervisé** ;
- données sans labels : **machine learning non supervisé**.

Et entre les deux :

- on connaît quelques labels ;
- on a beaucoup de points sans labels.

On a considéré deux cas :

- données avec des labels : **machine learning supervisé** ;
- données sans labels : **machine learning non supervisé**.

Et entre les deux :

- on connaît quelques labels ;
- on a beaucoup de points sans labels.

Est-ce qu'on arrive à exploiter les points sans labels ?

Machine Learning Semi Supervisé

Encore (et toujours) des fruits...

Par exemple :

Couleur	Diamètre	Noyau	Fruit
Rouge	5	Non	Pomme
Vert	5	Non	Pomme
Jaune	3	Non	Banane
Jaune	5	Non	Citron
Rouge	2	Oui	Cerise

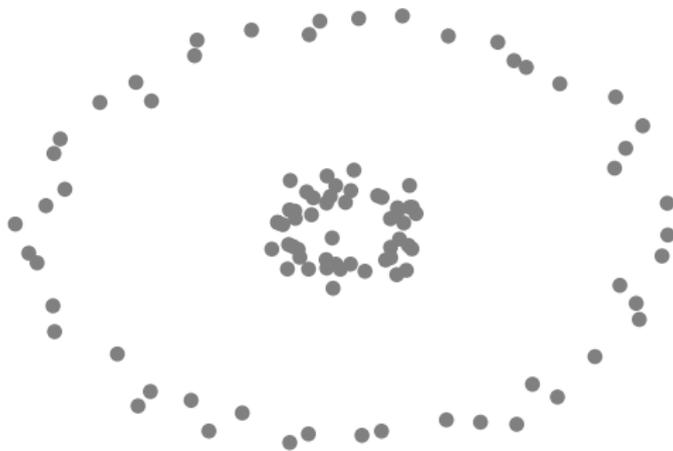
Avec Labels



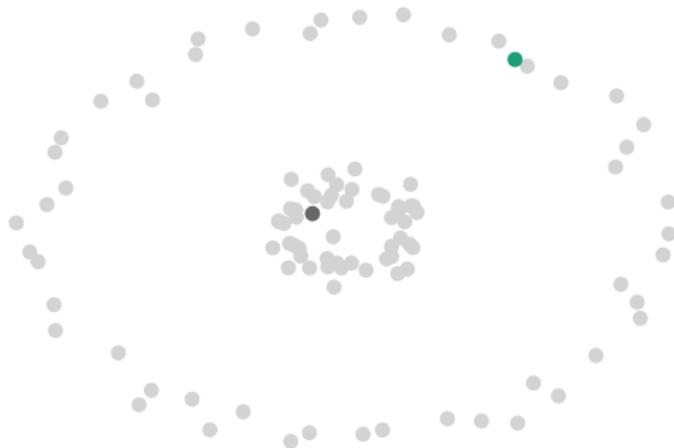
Sans Labels

On a seulement 5 points avec labels, mais beaucoup sans.

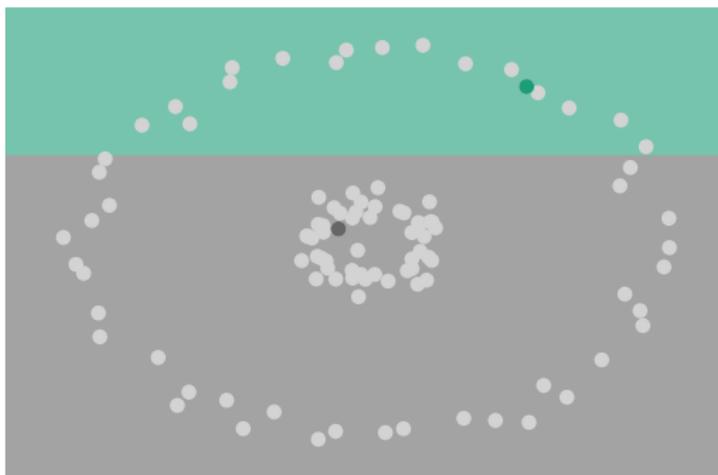
Prenons ces données là :



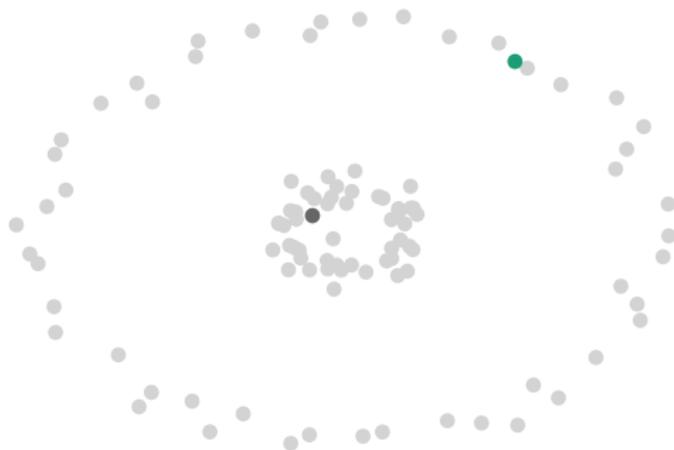
On n'a les labels que de deux points...



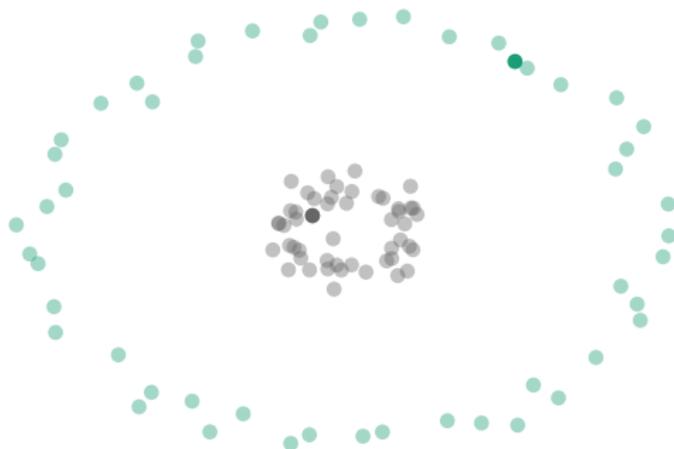
Si on n'utilise que ces points, on apprend ce classifieur :



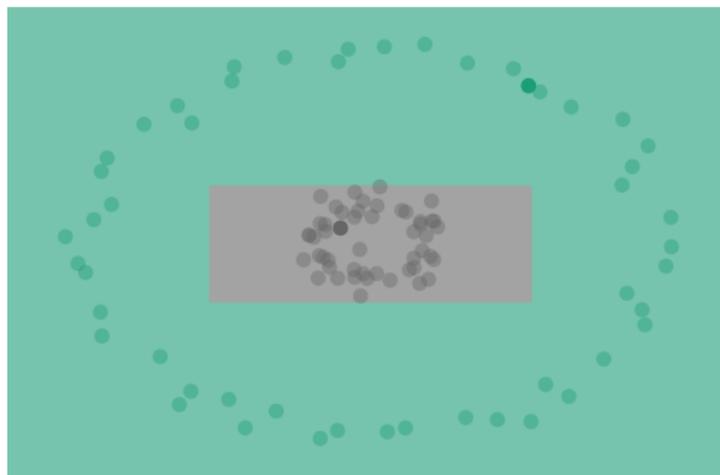
Mais on dirait qu'il y a une forme de structure dans notre dataset :



On peut propager les labels :



Et le classifieur obtenu est bien meilleur !



Comment Apprendre Des Nouveaux Labels ?

Comment Apprendre Des Nouveaux Labels ?

Inattendu : des graphes !

Bien sûr on va utiliser...

Comment Apprendre Des Nouveaux Labels ?

Inattendu : des graphes !

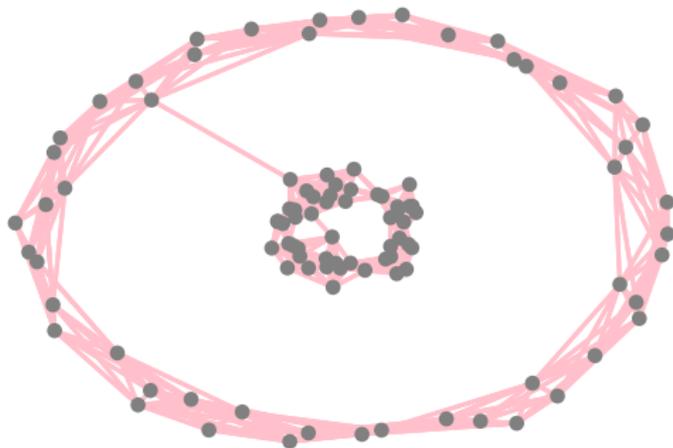
Bien sûr on va utiliser... des **graphes** !

(vous commencez à vous en douter...)

Comment Apprendre Des Nouveaux Labels ?

Inattendu : des graphes !

Graphe des plus proches voisins :



On va propager les labels que l'on connaît sur ce graphe !

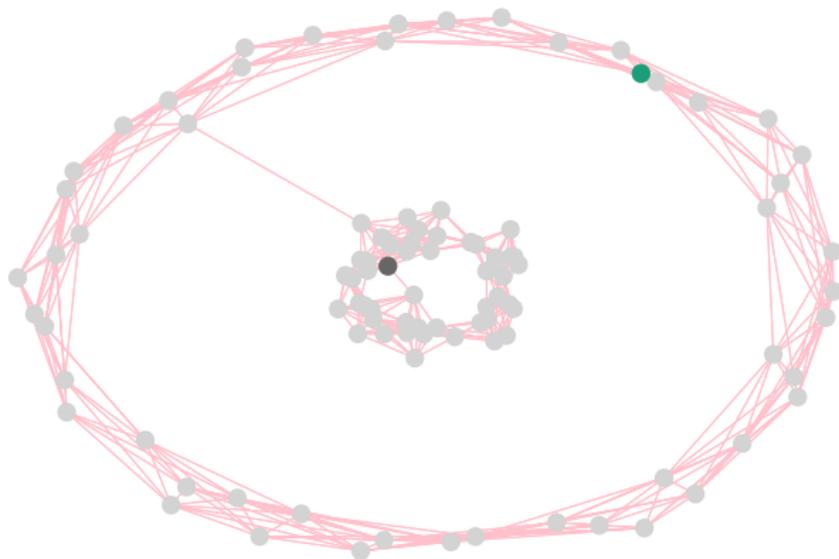
Un algorithme de **propagation des labels** :

- au début, on connaît quelques labels y ;
- en boucle jusqu'à convergence :
 - pour chacun des points :
 - si les voisins du points ont des labels :
→ attribuer le label le plus fréquent des voisins au point
 - sinon ne rien faire.

Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

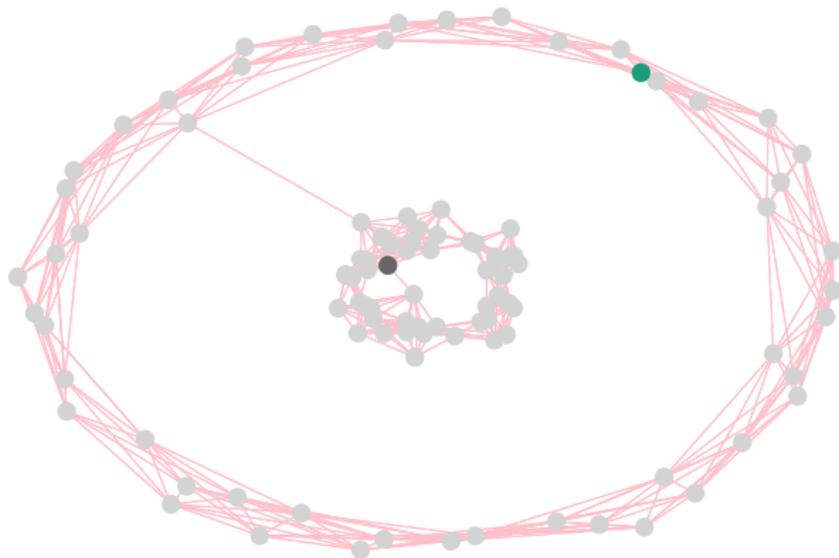
Au début, on n'a que deux labels :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

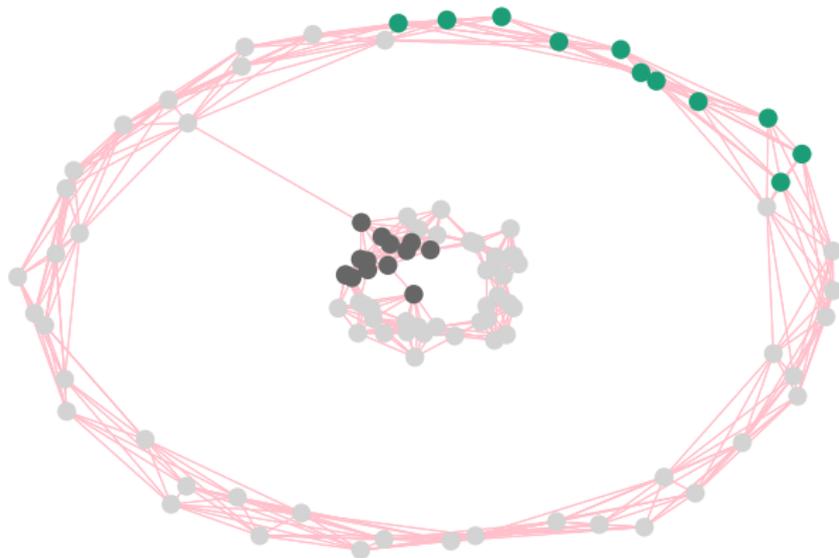
On complète avec le label le plus fréquent chez les voisins :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

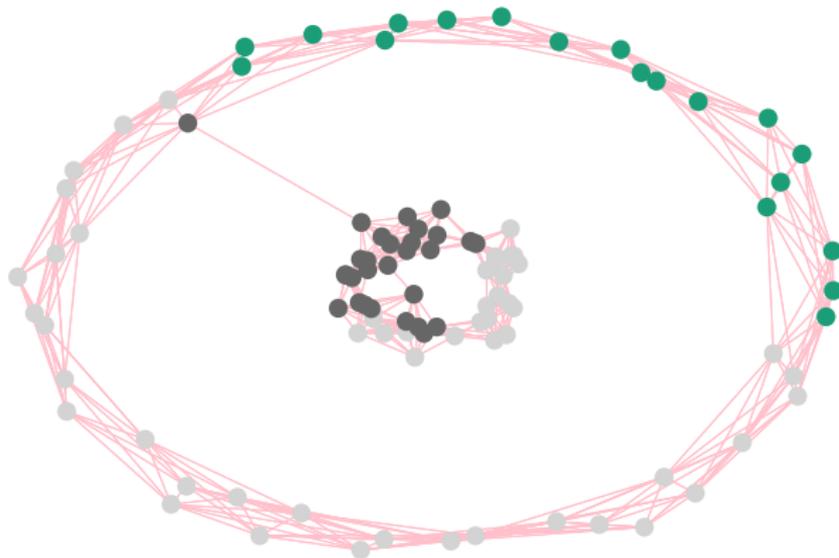
On complète avec le label le plus fréquent chez les voisins :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

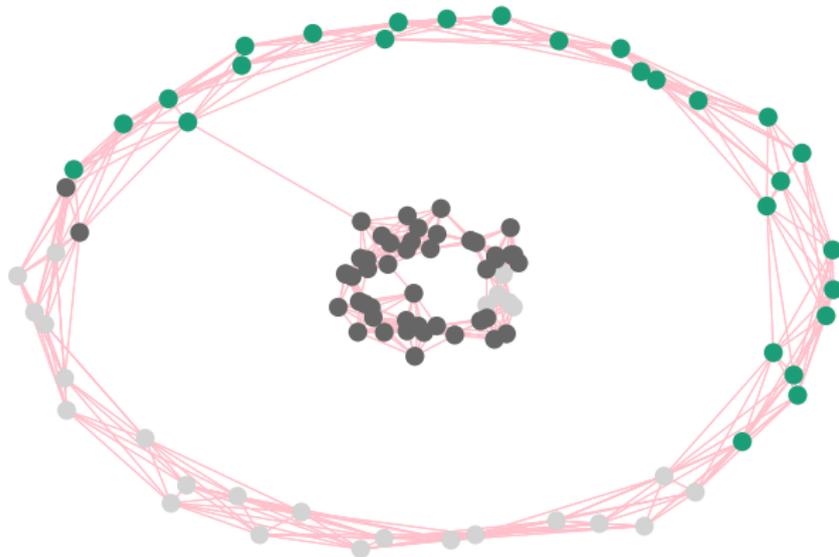
On complète avec le label le plus fréquent chez les voisins :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

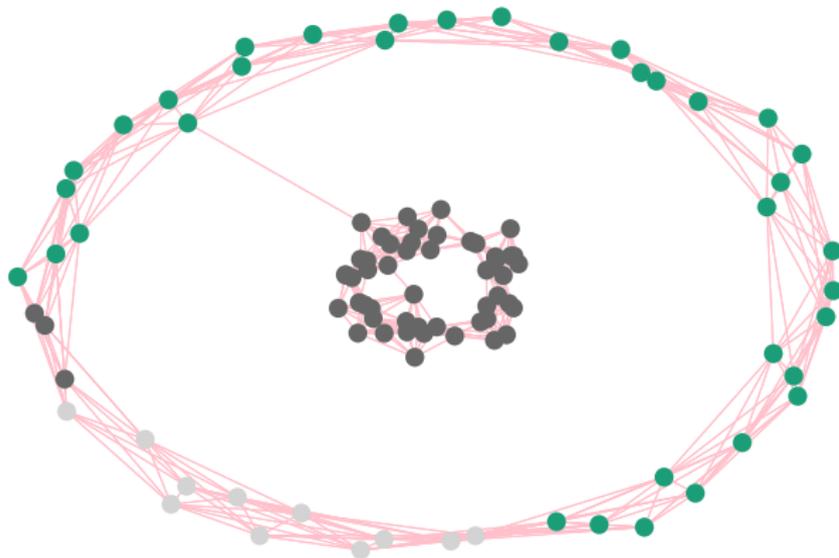
On complète avec le label le plus fréquent chez les voisins :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

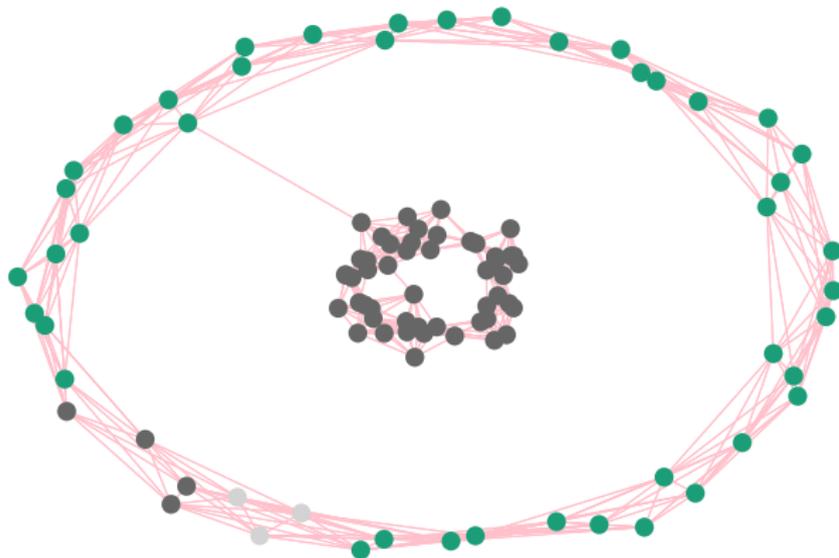
On complète avec le label le plus fréquent chez les voisins :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

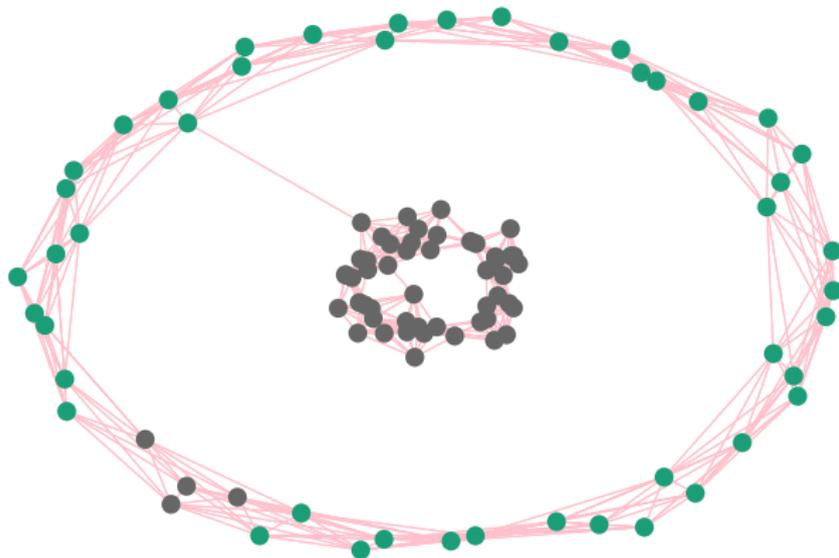
On complète avec le label le plus fréquent chez les voisins :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

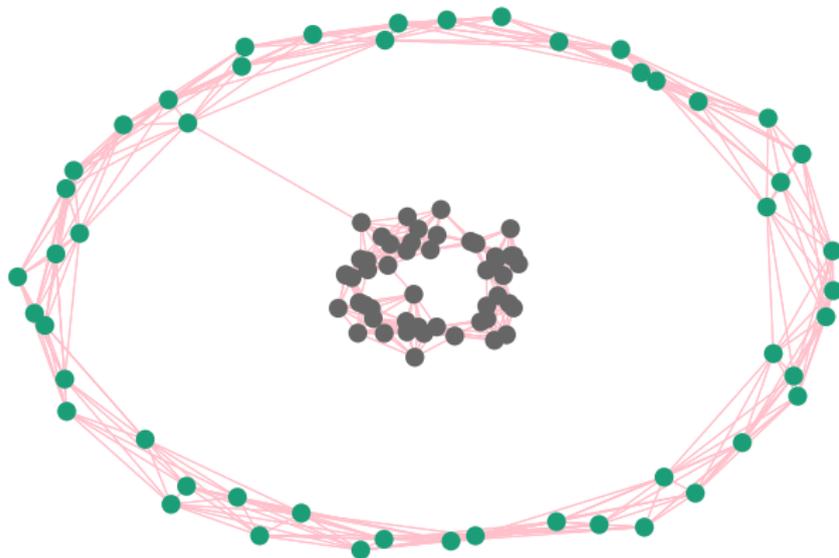
On complète avec le label le plus fréquent chez les voisins :



Comment Apprendre Des Nouveaux Labels ?

Propagation des labels

Jusqu'à convergence !



TP !

Dans le TP d'aujourd'hui on va :

- regarder comment se comporte l'algorithme de propagation des labels dans différents cas ;
- regarder si cela permet effectivement d'obtenir des meilleurs classifieurs.