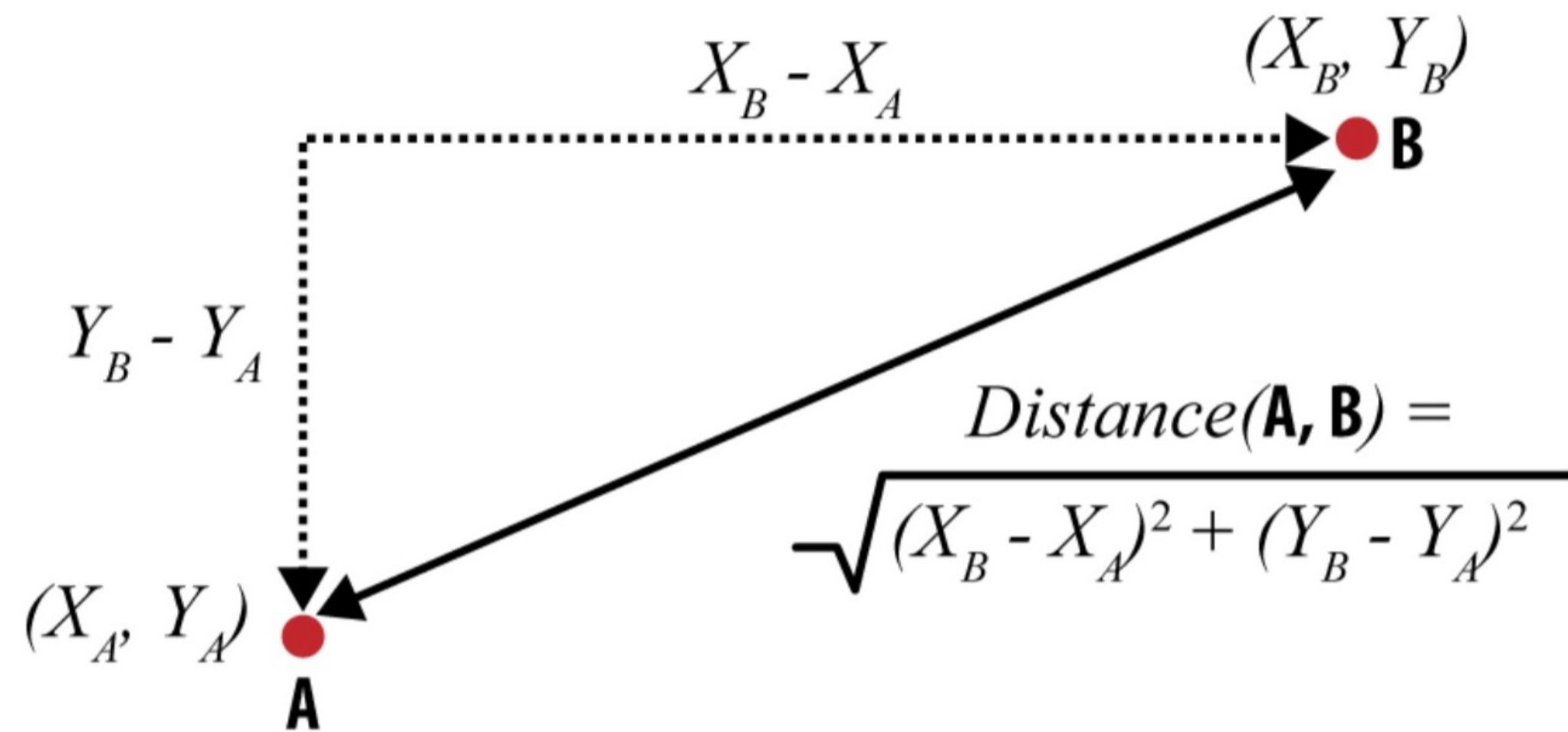


11. UNSUPERVISED LEARNING

LEV KIWI

DISTANCE



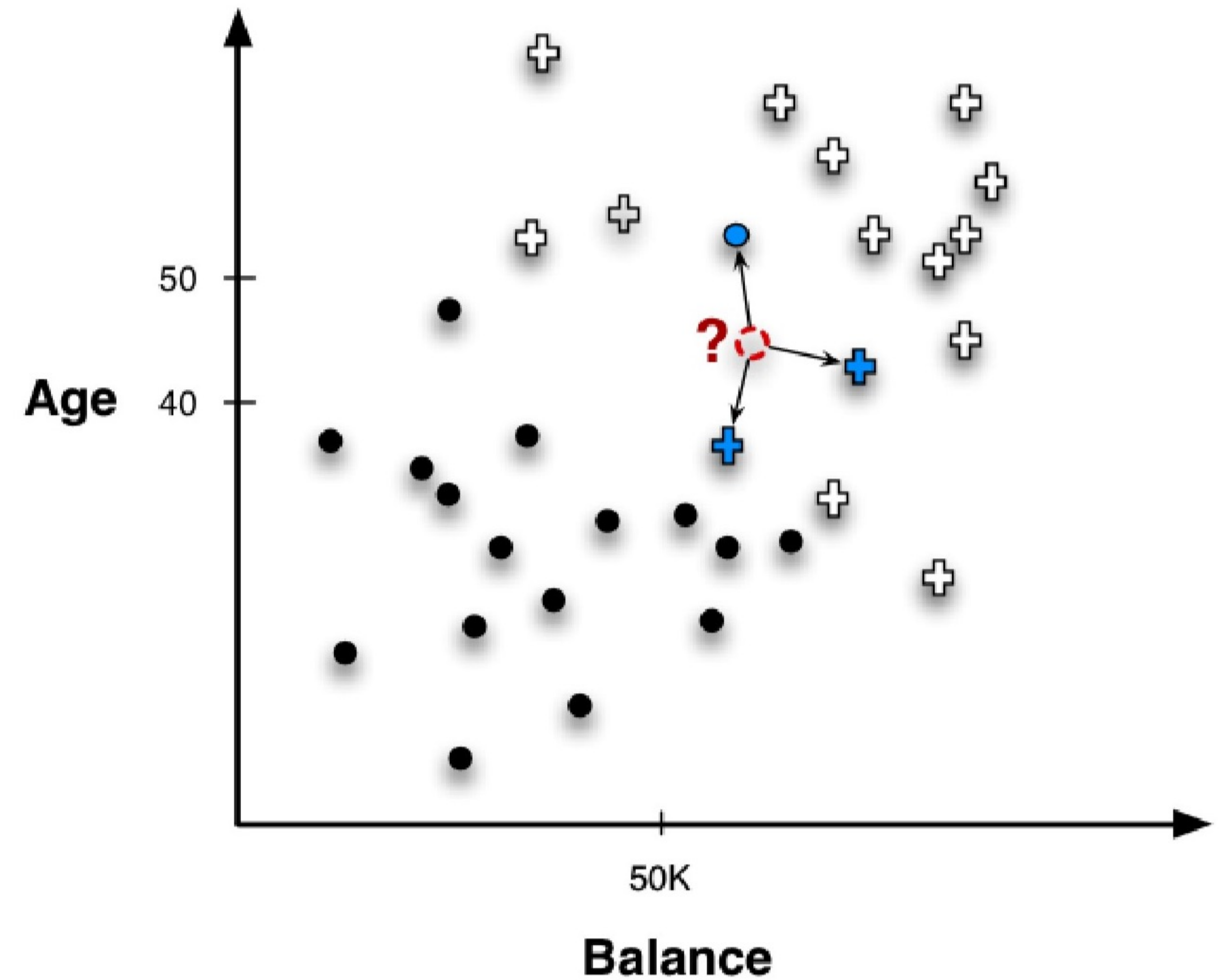
Equation 6-1. General Euclidean distance

$$\sqrt{(d_{1,A} - d_{1,B})^2 + (d_{2,A} - d_{2,B})^2 + \dots + (d_{n,A} - d_{n,B})^2}$$

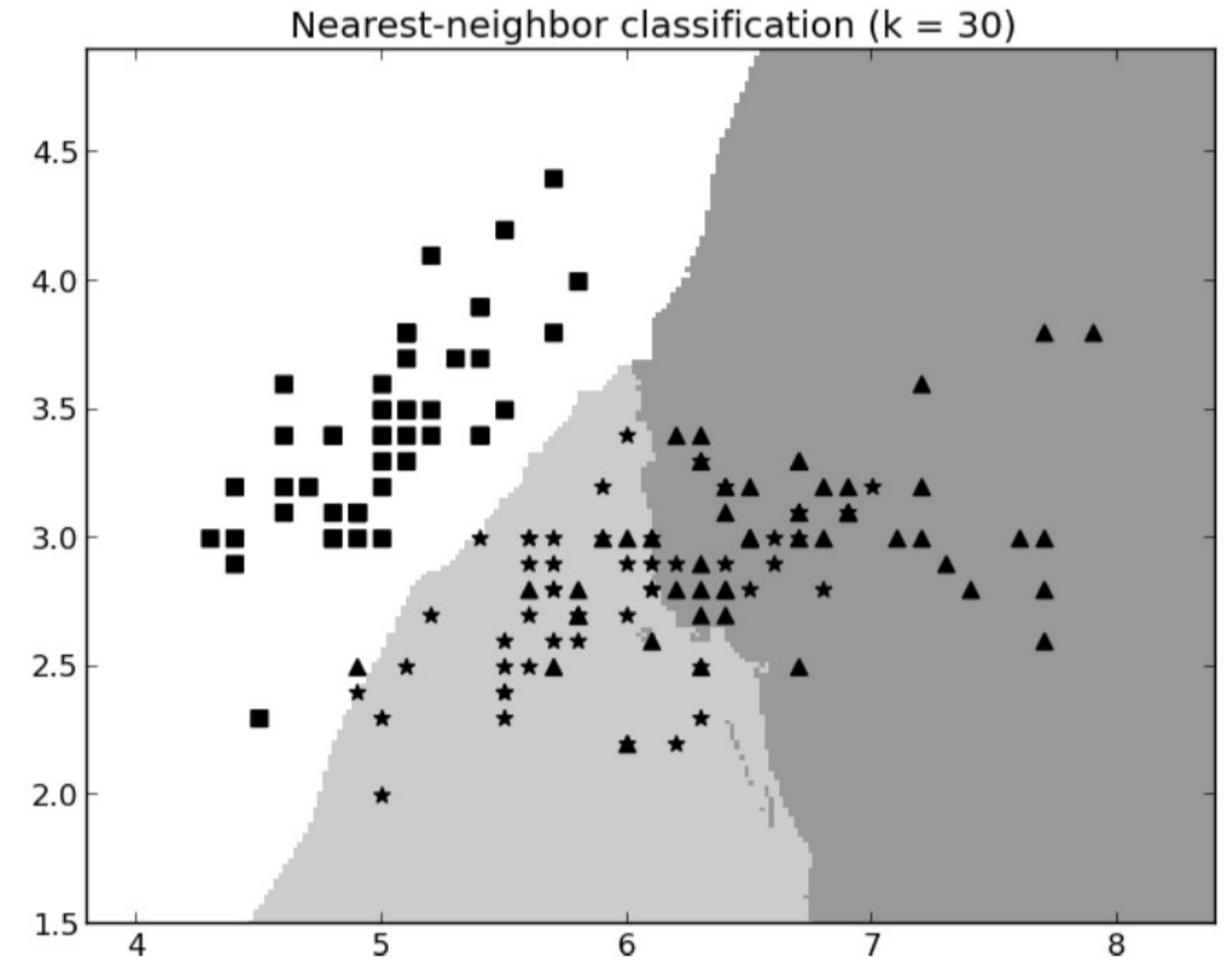
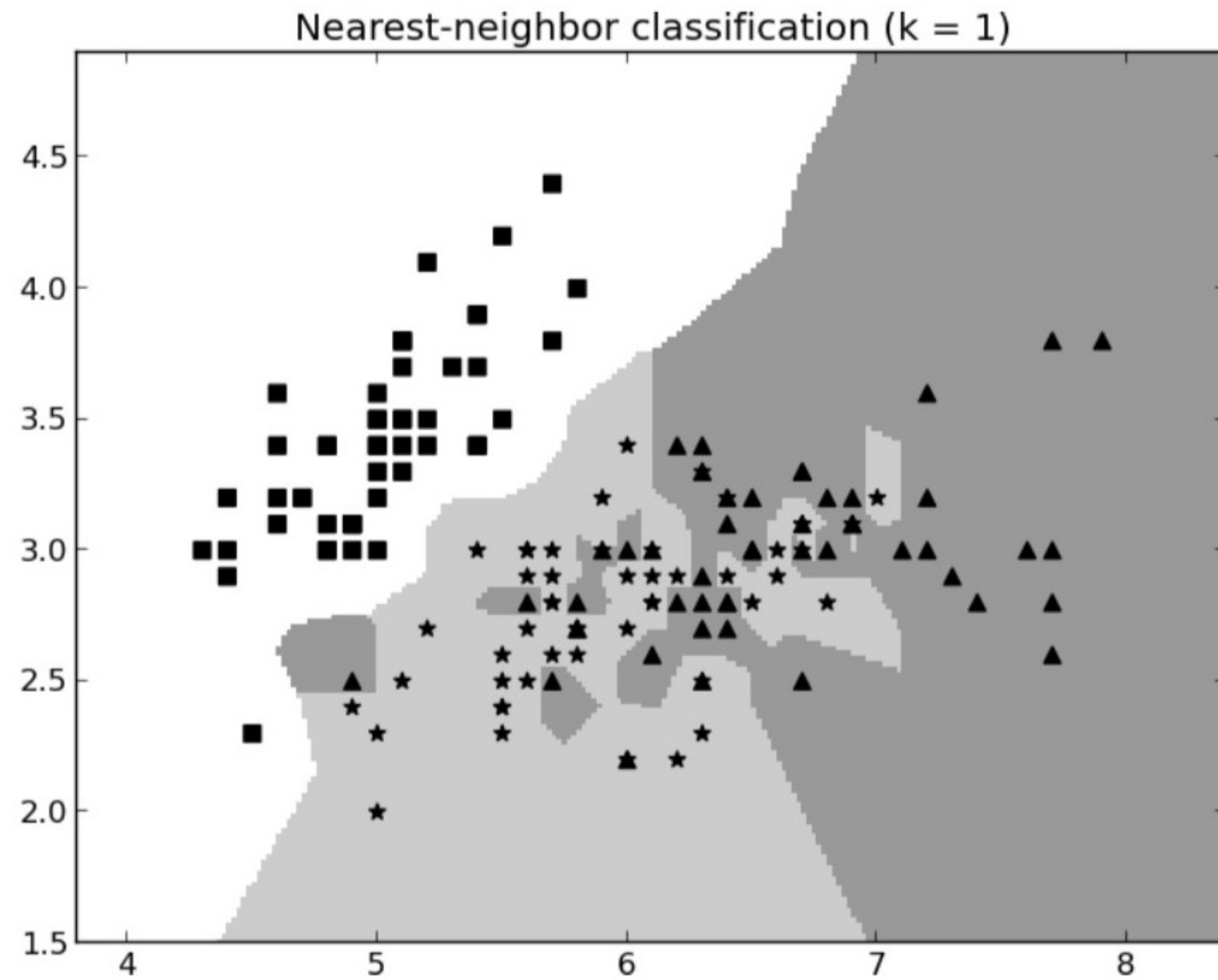
K-NEAREST NEIGHBORS I

Procédure

1. Le dataset d'entraînement est enregistré en mémoire
2. La prédiction est calculée comme étant la tendance centrale des k -voisins les plus proches



K-NEAREST NEIGHBORS II



TACHES DE MACHINE LEARNING

par catégories de d'algorithmes...

- Supervisé.
- Non-Supervisé.
- par Renforcement.

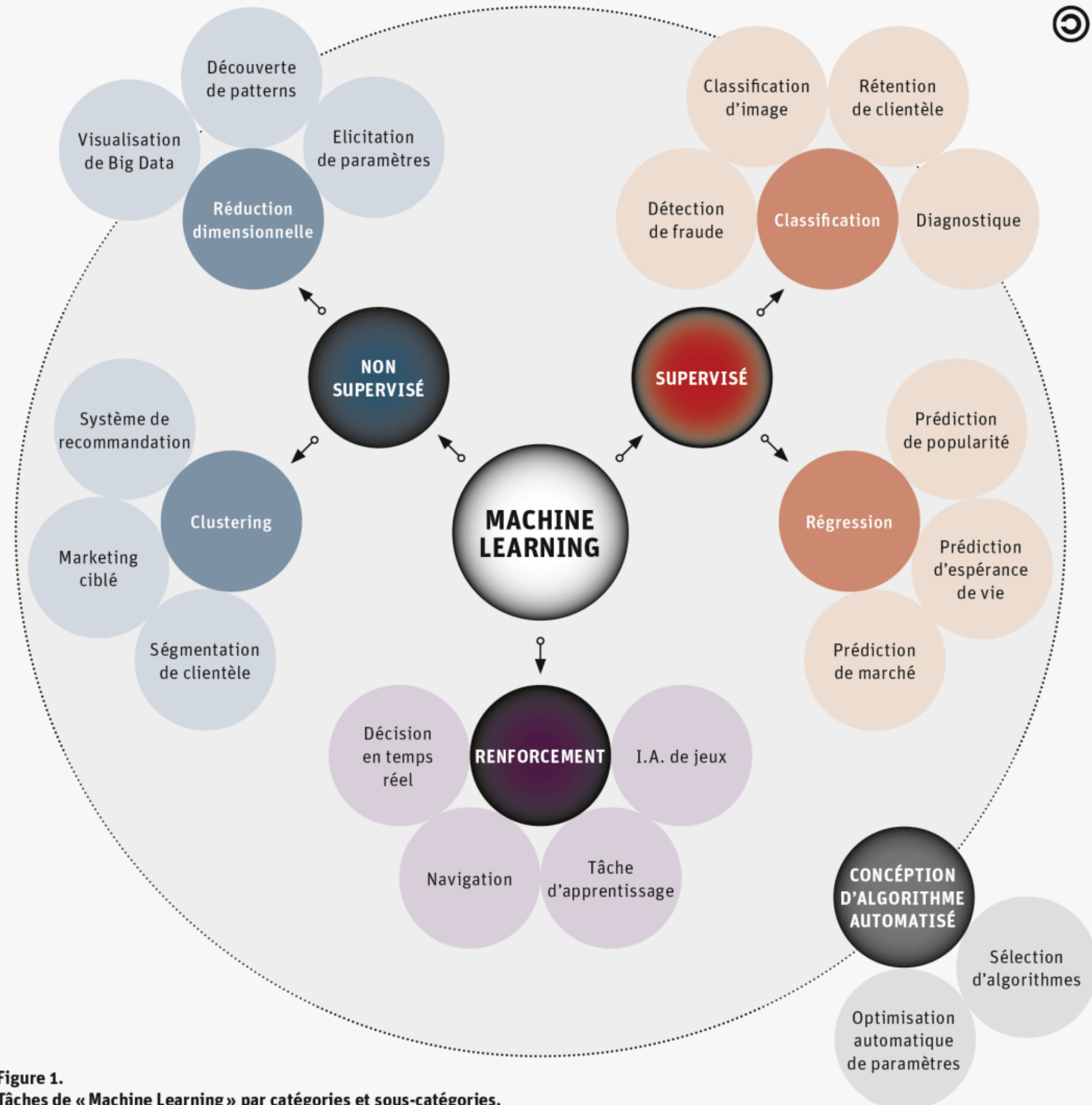
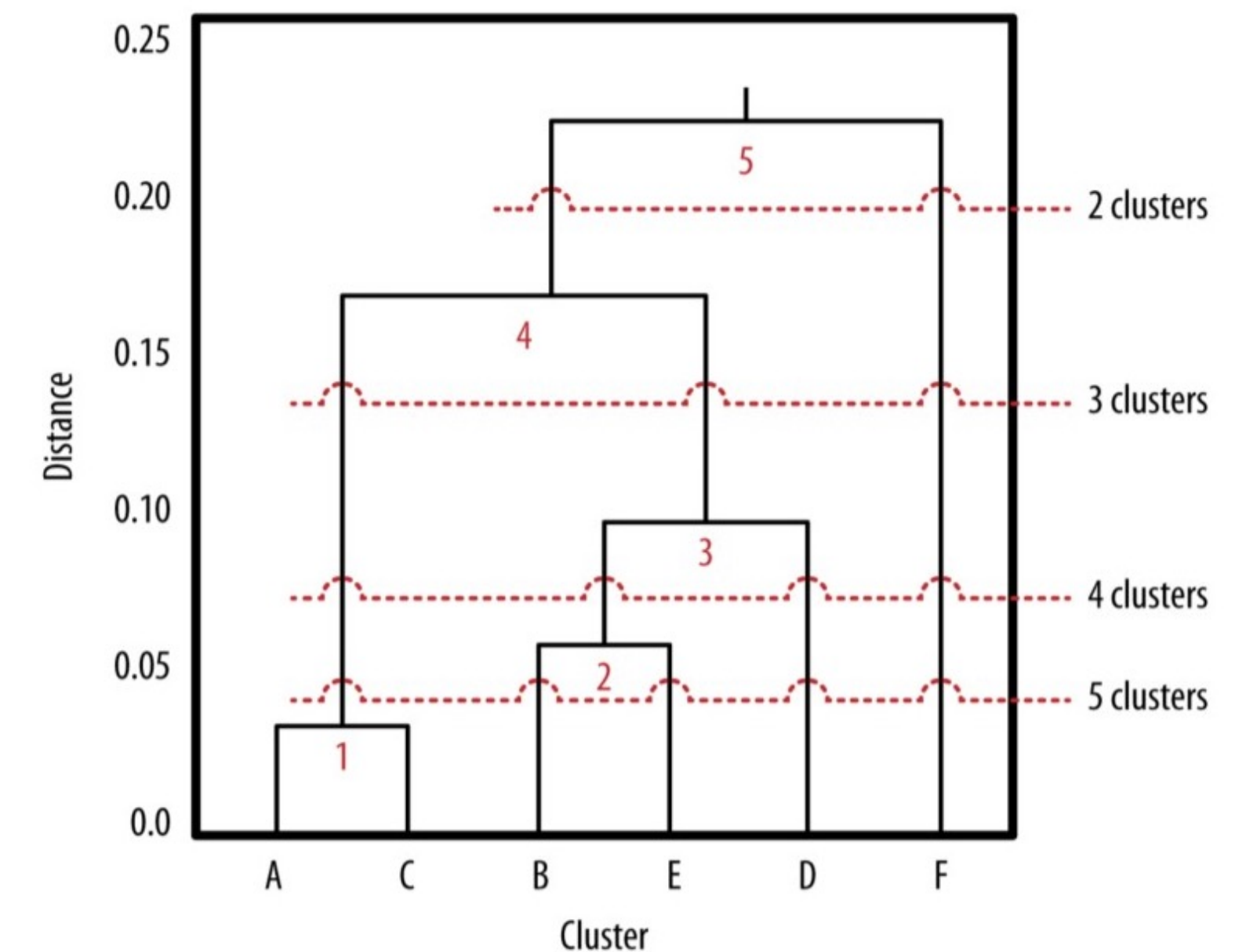
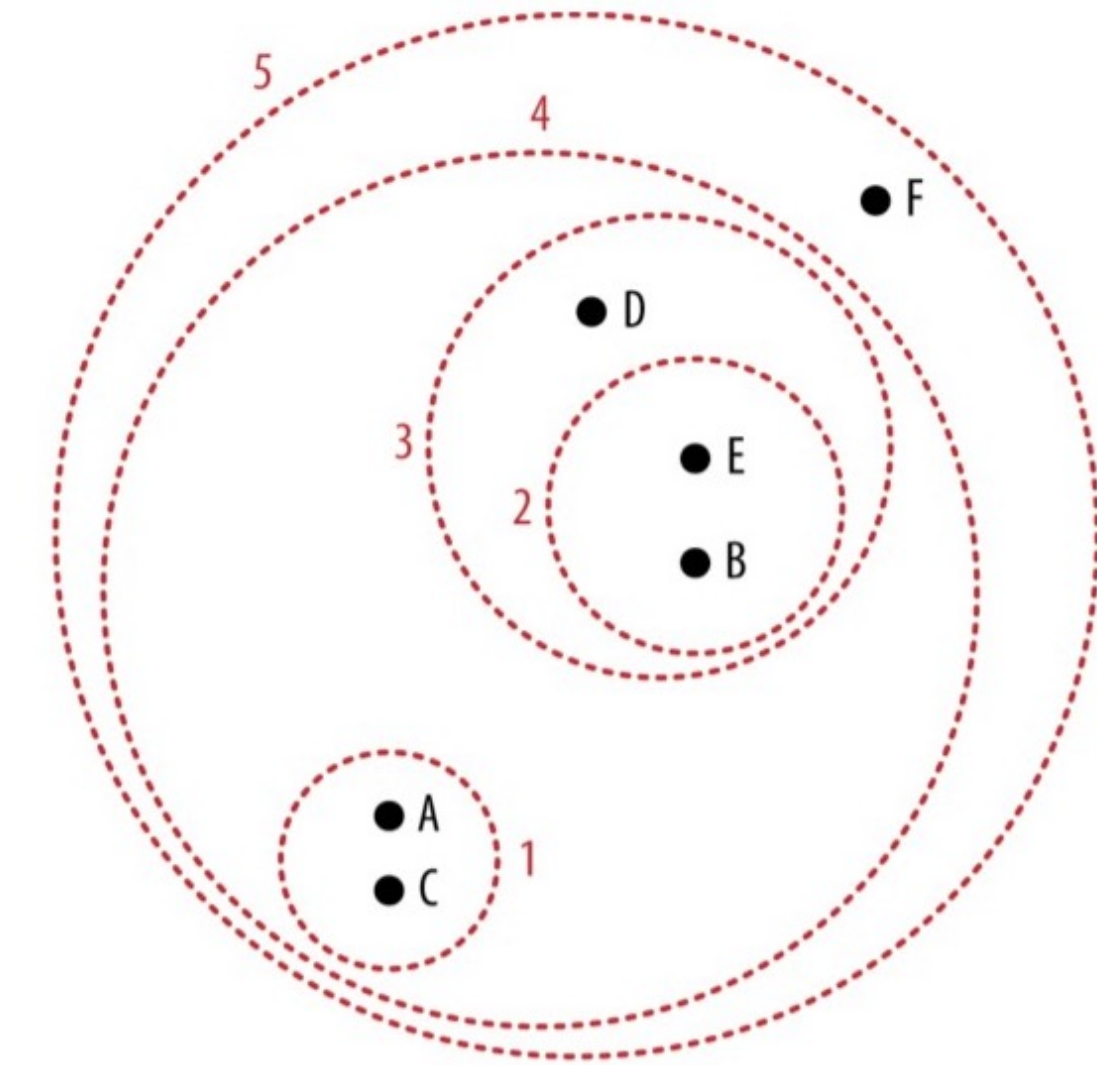


Figure 1.
Tâches de « Machine Learning » par catégories et sous-catégories.

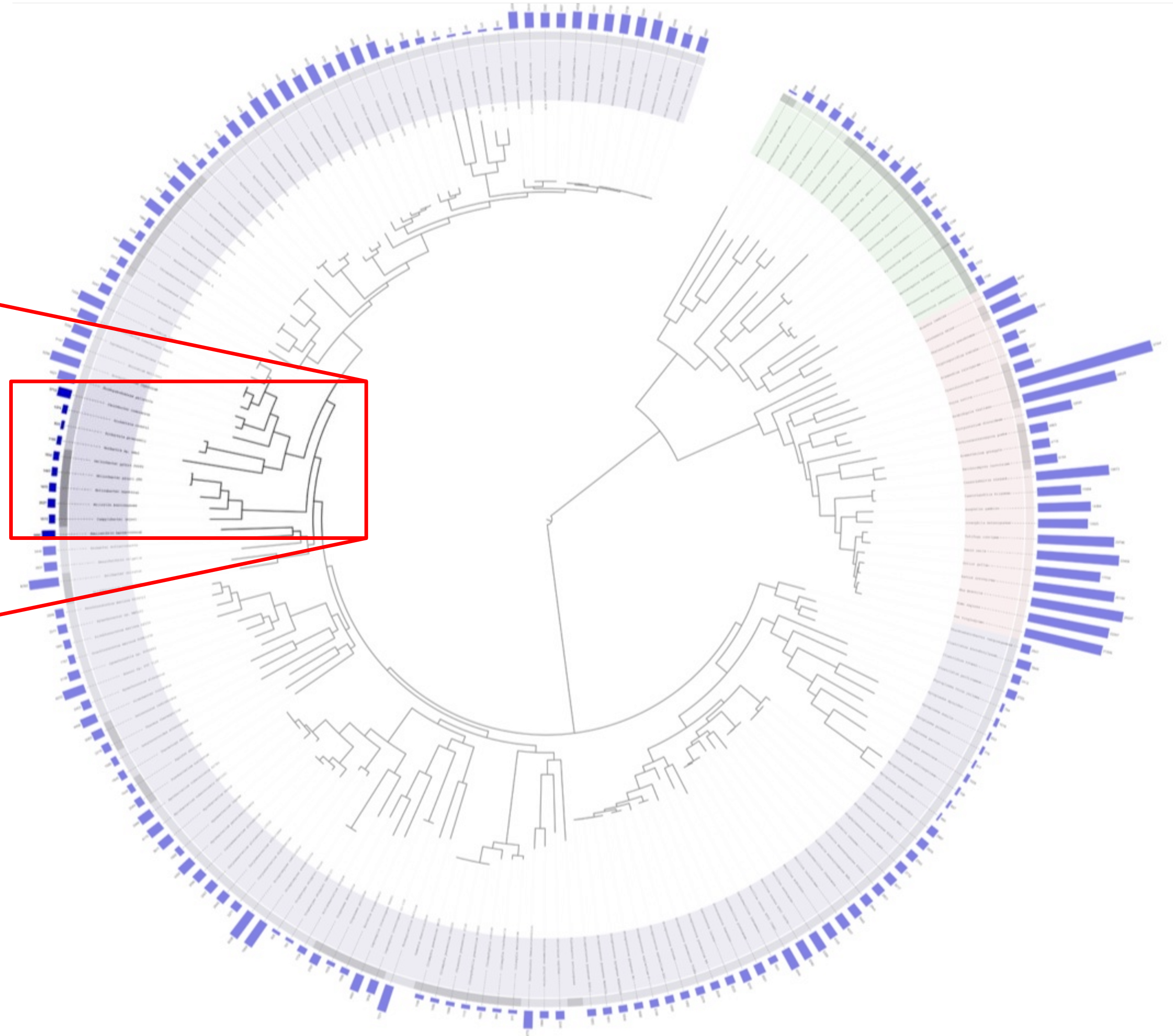
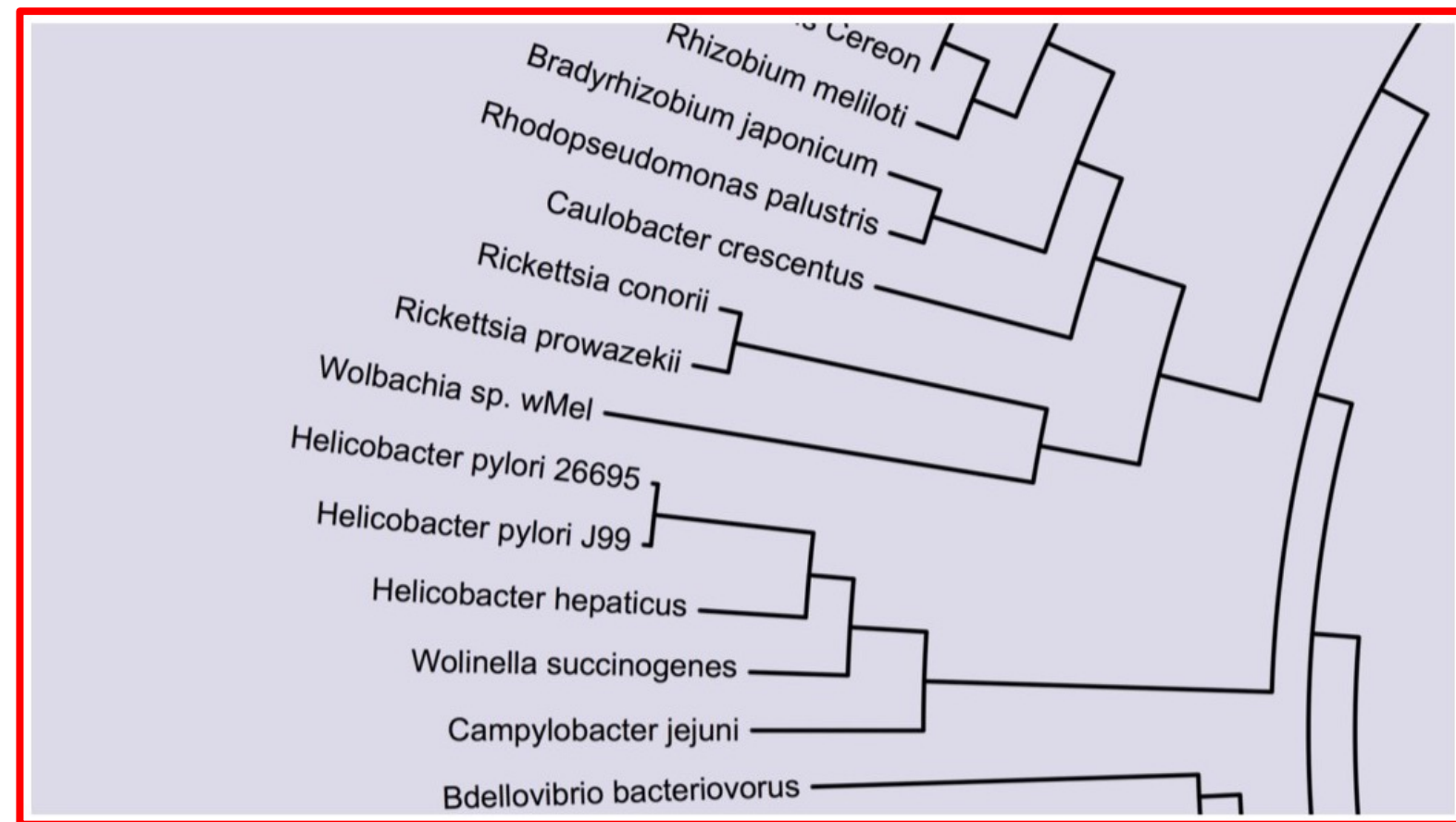
CLUSTERING HIÉRARCHIQUE

Procédé

1. On calcul les distances entre tous les points deux à deux.
2. Les deux points les plus proches sont fusionnés à leurs **barycentre**.
3. Tant qu'on a plus que deux points on reviens à l'étape 1.
4. On décide d'une distance de clustering.



DENDROGRAMME

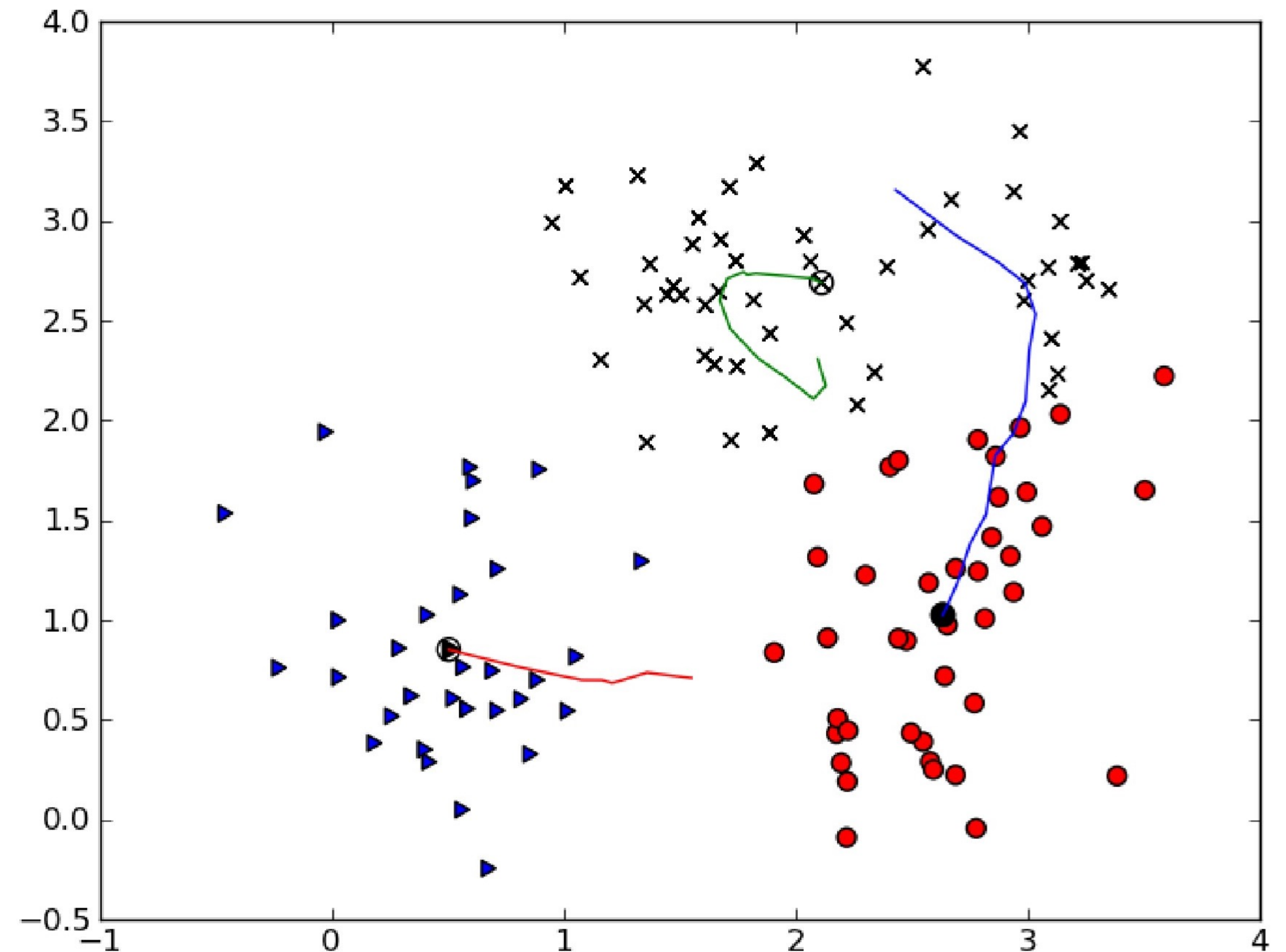
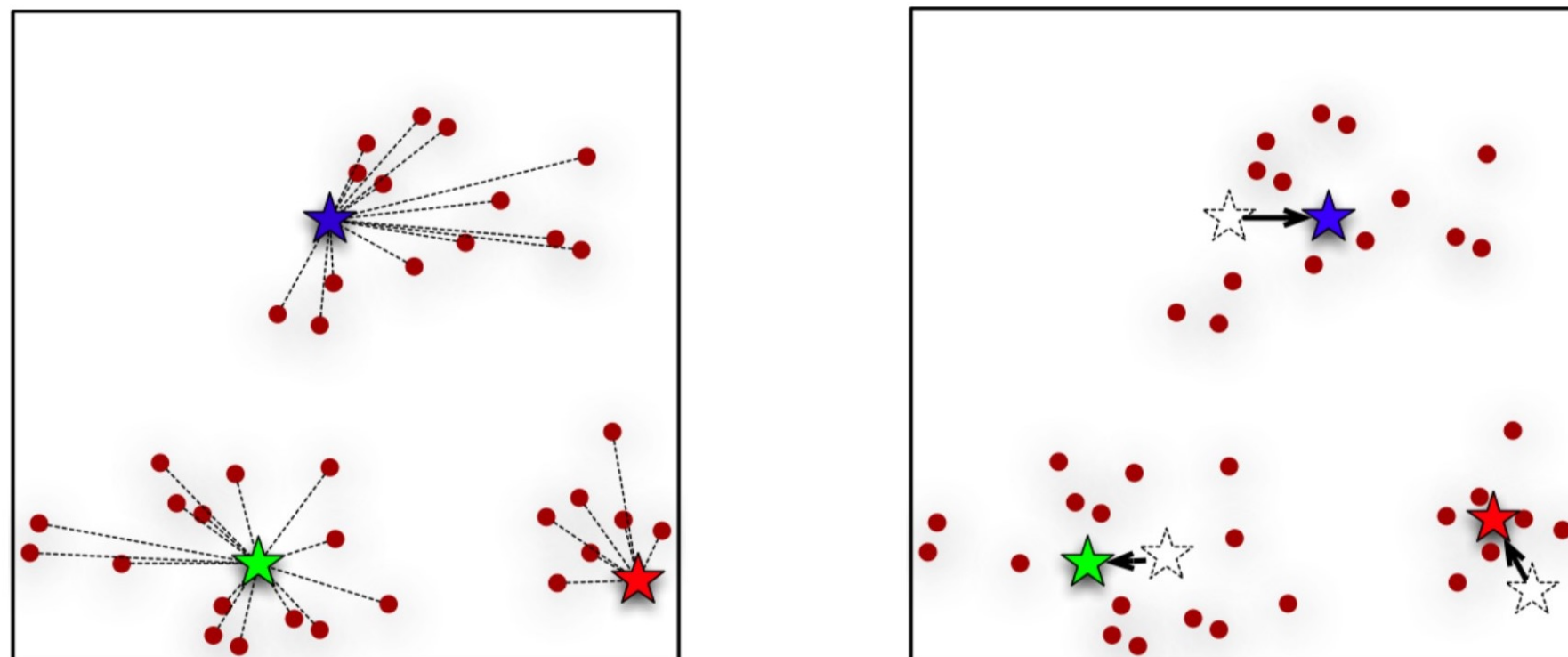


K-MEANS

Procédé

On pose k -points (**k-centres**) aléatoirement dans l'espace

1. Pour chaque point du dataset, on assigne le k -centre le plus proche.
2. On calcule le barycentre de tous les points du même k -centre.
3. Le **barycentre** devient le nouveau k -centre.
4. Si un k -centre a bougé, on retourne à l'étape 1.



SCIKIT-LEARN CLUSTERING

