

Présentation Projet SDA

Projet Algorithmique

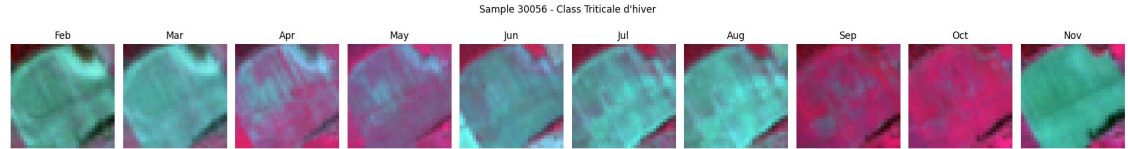
Cerisara Nathan, Lardy Emilie, Gerouille Guilhem

Sommaire

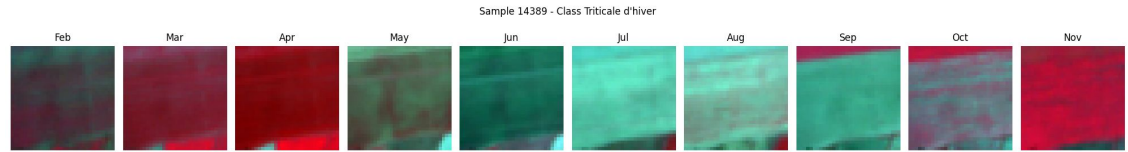
1. Présentation du sujet
2. Equilibrage des données
3. Méthode: ACP + Arbre de décision
4. Méthode: Modèles linéaires
5. Méthode: Resnet
6. Méthode: ACP + SVM
7. Conclusion

1. Présentation du sujet

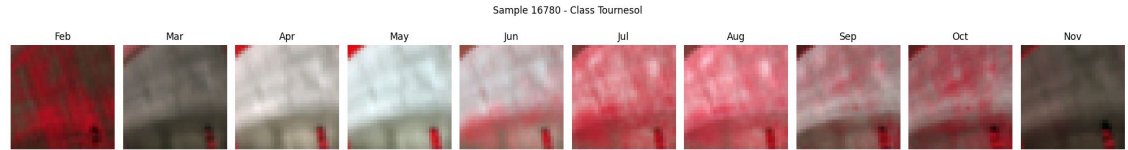
- Classification de champs



- 10 images (3 canaux)
 - > 1 par mois



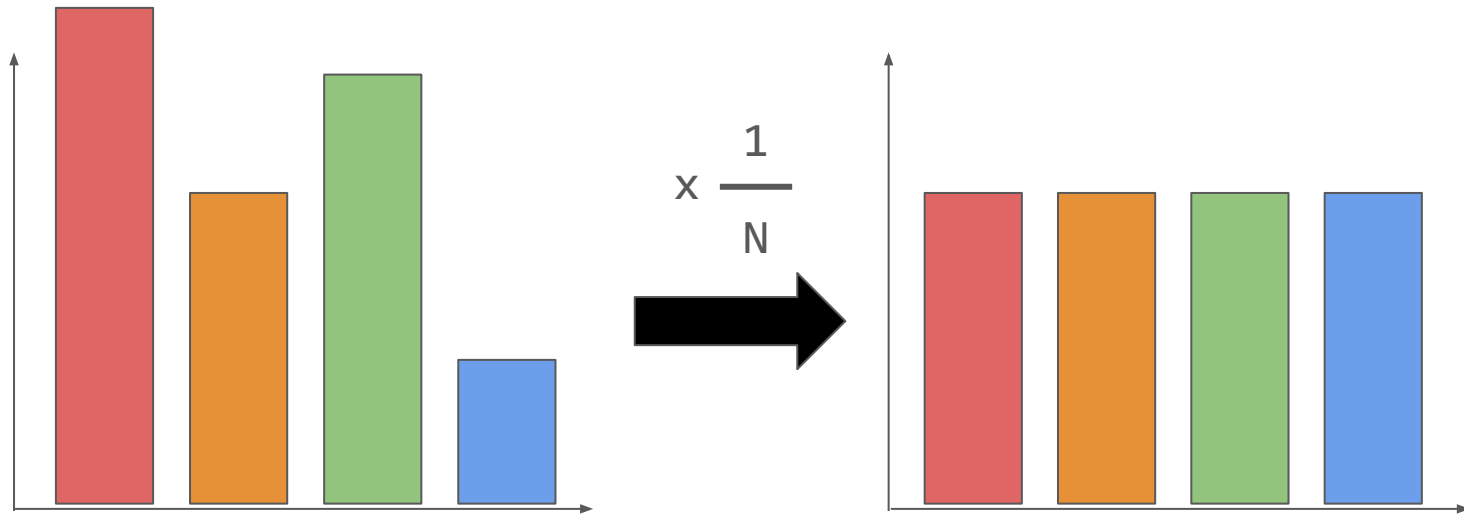
- 20 classes de cultures



2. Equilibrage des données

Méthode 1 :

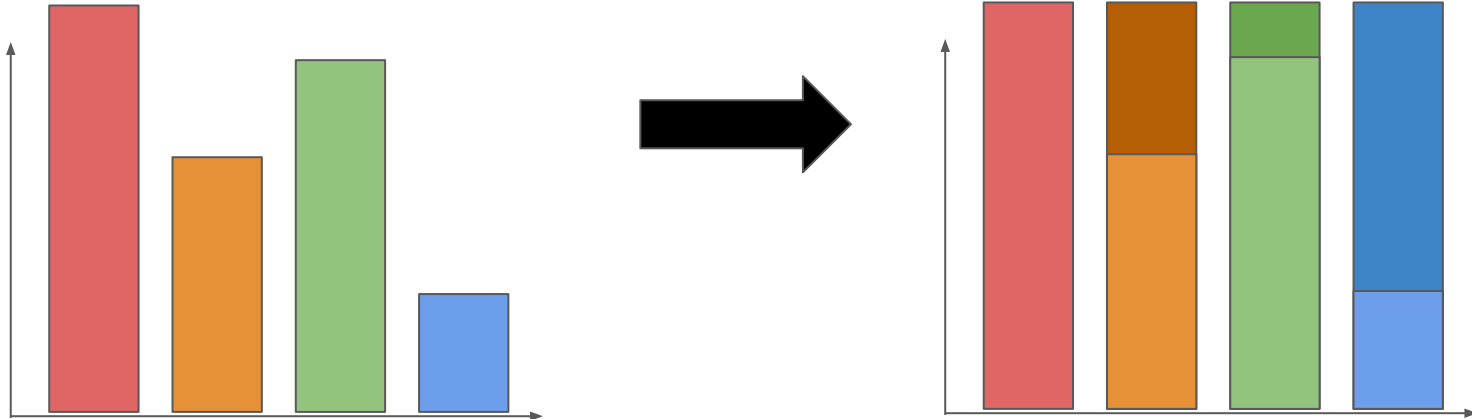
- Utilisation paramètre équilibrage poids automatiques



2. Equilibrage des données

Méthode 2 :

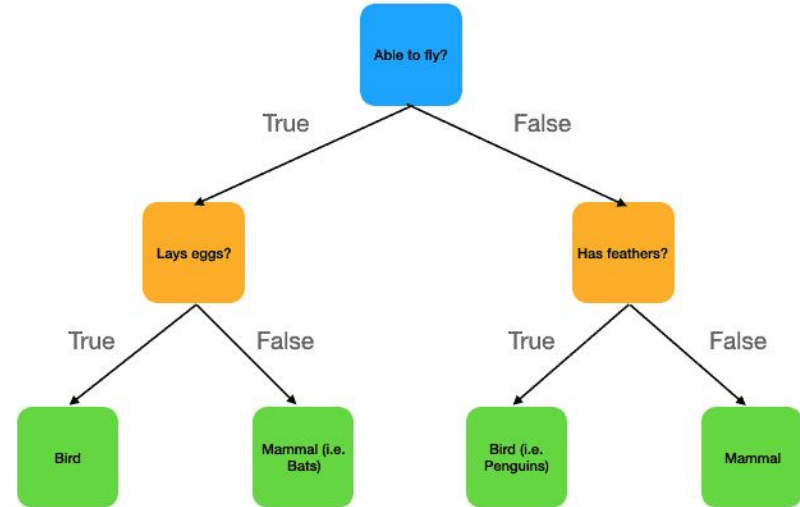
- Resnet / modèles linéaires -> Équilibrage manuel
 - Avec N le nombre max d'elt dans une classe
 - On ajoute à toutes les classes $N - N_i$ elts comme suit:
 - On prend une image de base de la classe, auquel on ajoute un bruit gaussien



3. Méthode: ACP + Arbre de décision

Arbre de décision :

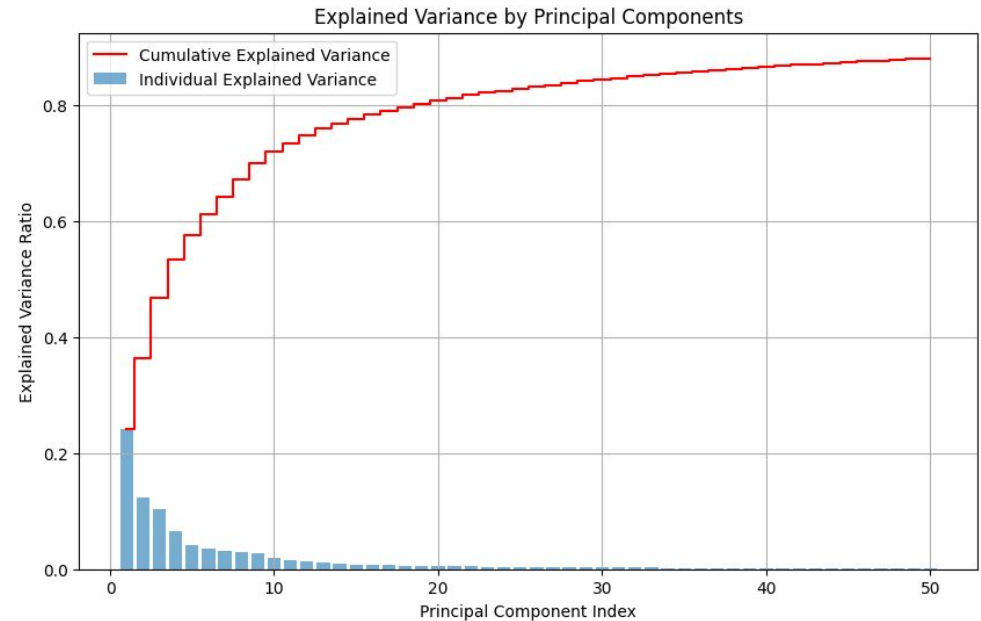
- Puissants pour la classification
- Facilement interprétables
- Problème : dimension très élevée
- Solution : ACP



3. Méthode: ACP + Arbre de décision

ACP :

- Réduction à 10 dimensions
- Calcul plus rapide



3. Méthode: ACP + Arbre de décision

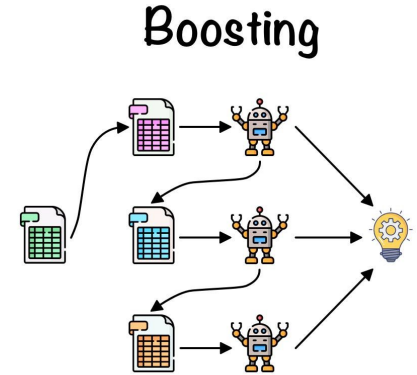
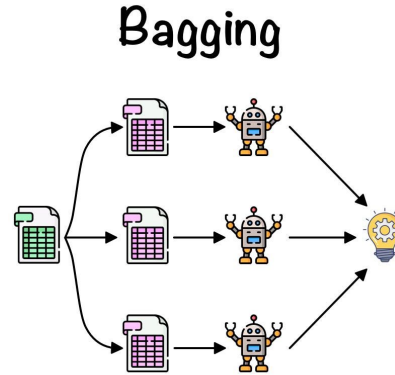
Résultats :

| Précision | Sans équilibrage | Avec équilibrage |
|--------------------|------------------|------------------|
| Jeu d'entraînement | 75% | 50% |
| Jeu de test | 68% | 40% |

3. Méthode: ACP + Arbre de décision

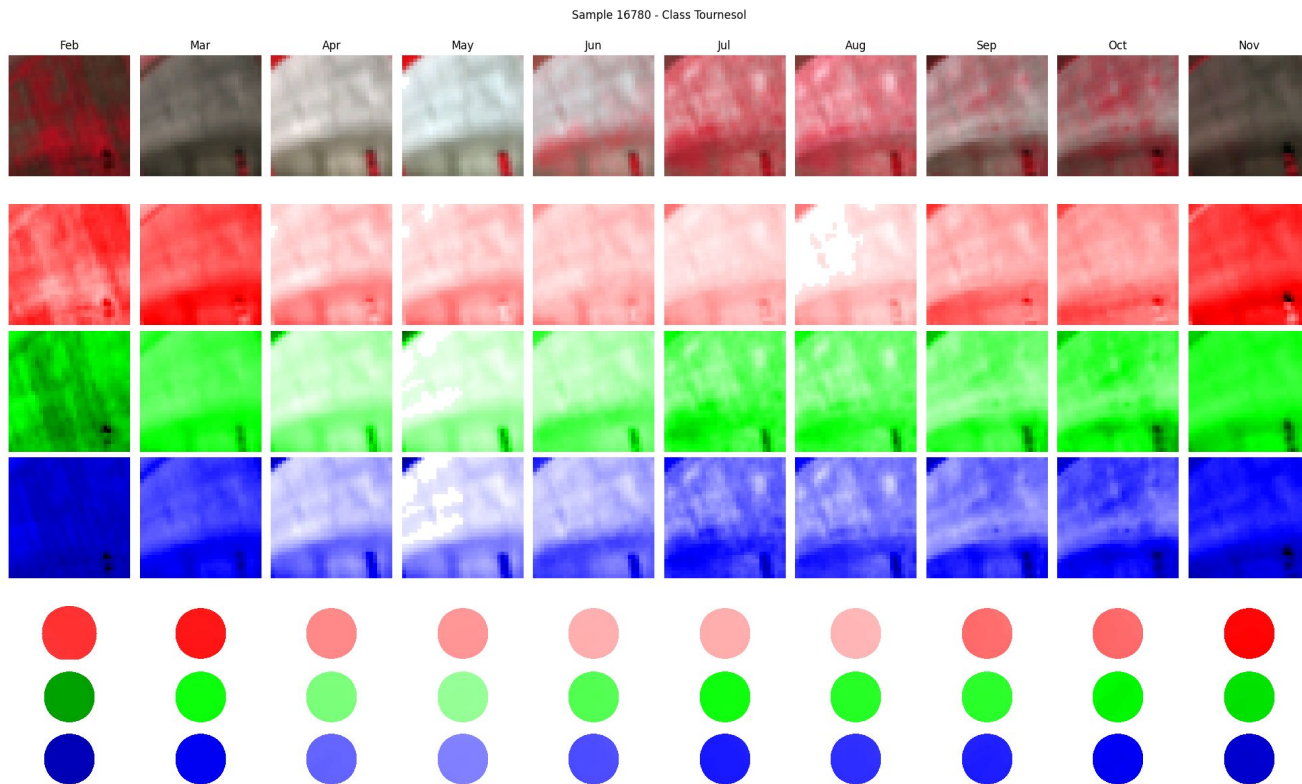
Améliorations possibles :

- Optimisation par bagging
- Optimisation par boosting



4. Méthode: Modèles linéaires

Moyenne des couleurs:

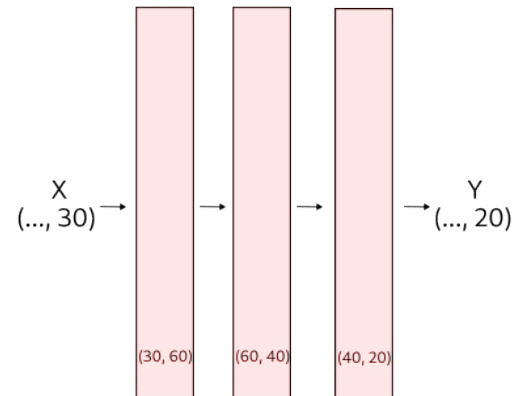
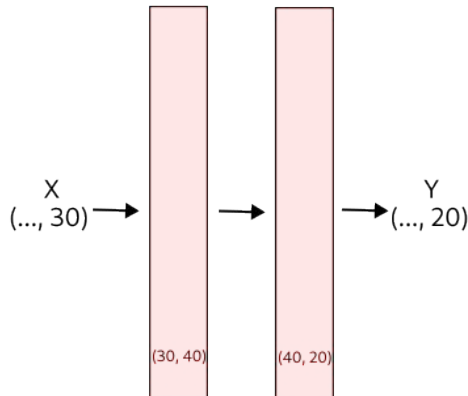
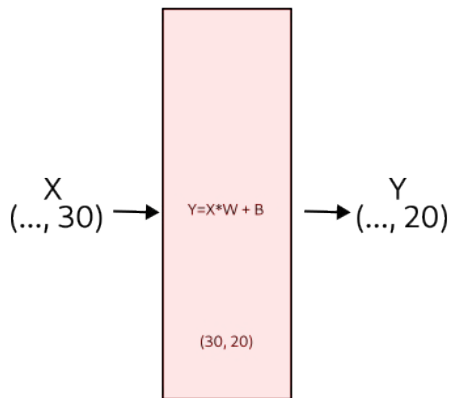
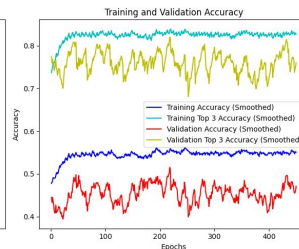
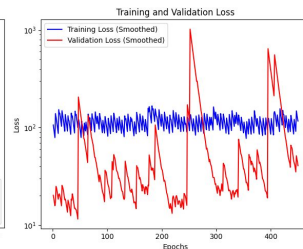
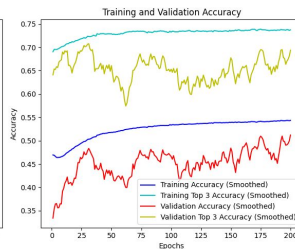
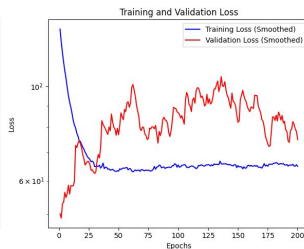
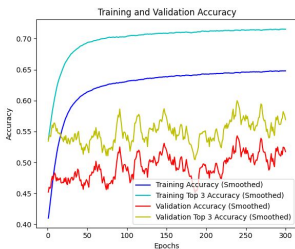
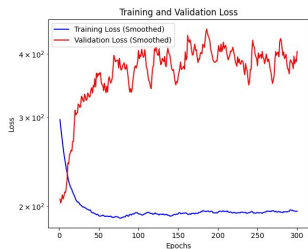


4.b. Modèles linéaires

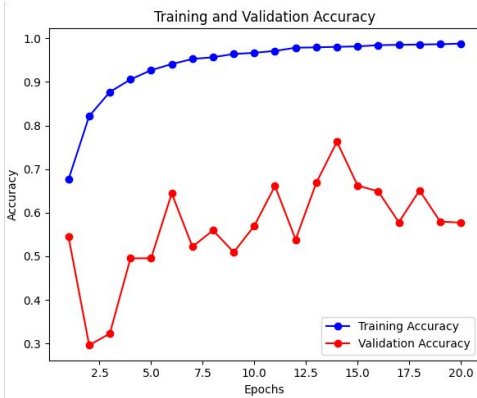
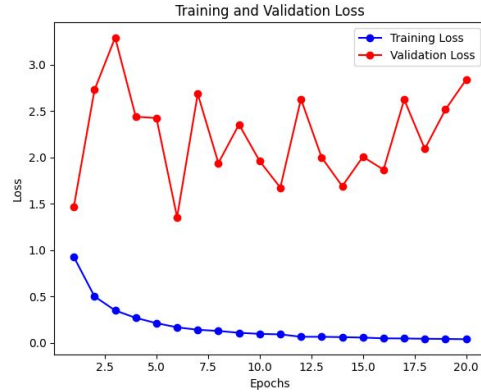
(Train): acc = ~64% | top3 acc = ~72%
 (Val): acc = ~54% | top3 acc = ~58%

(Train): acc = ~54% | top3 acc = ~75%
 (Val): acc = ~50% | top3 acc = ~70%

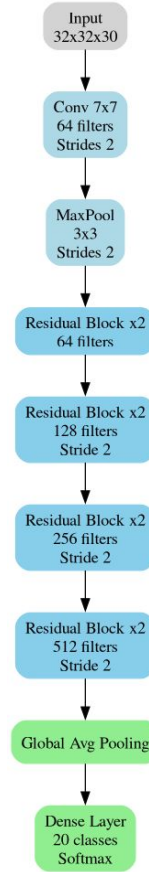
(Train): acc = ~55% | top3 acc = ~83%
 (Val): acc = ~52% | top3 acc = ~81%



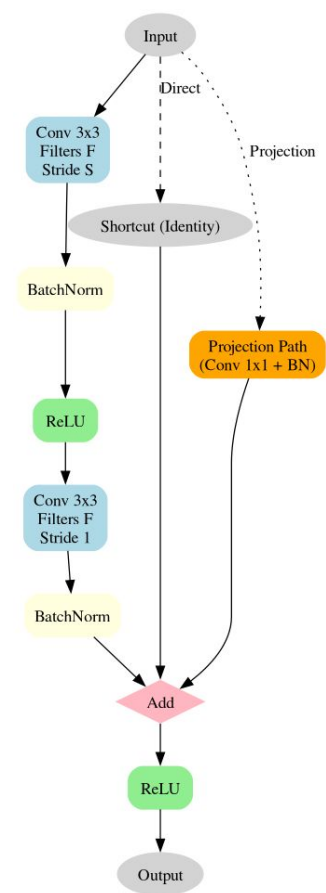
5. Méthode: Resnet



ResNet:



Residual Block:

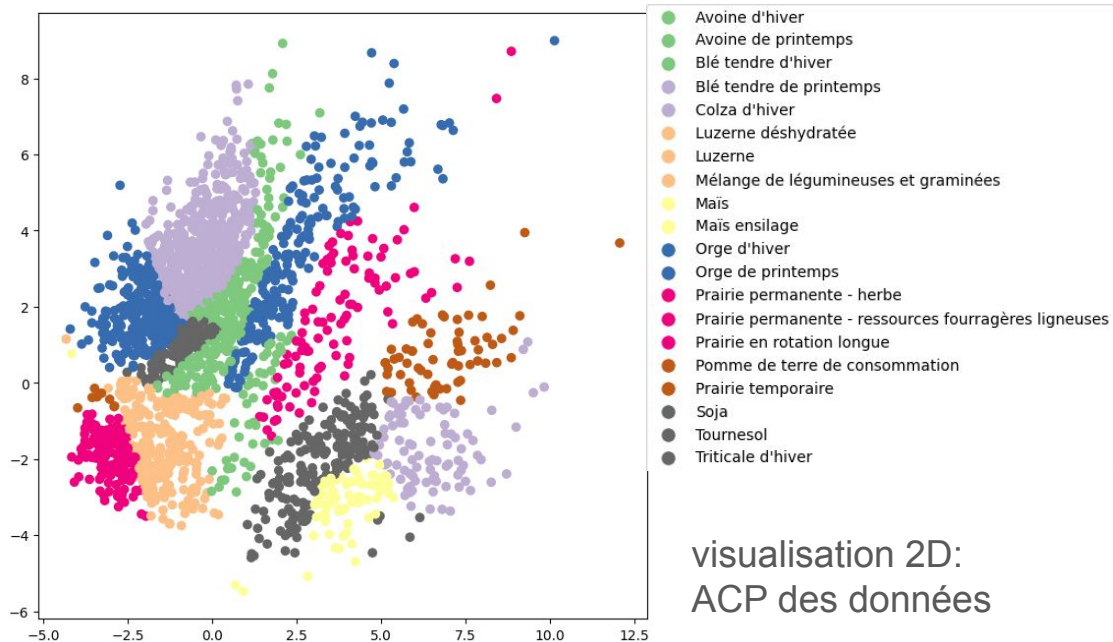


`model.count_params()` = 11 285 396 !

6. Méthode: SVM + ACP

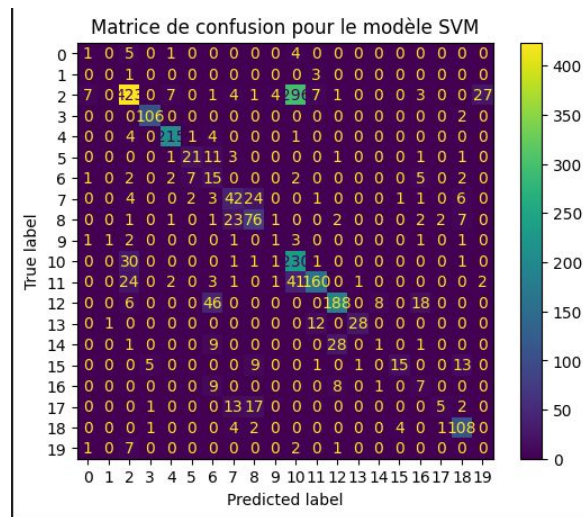
Pourquoi utiliser une SVM ?

- Simple
- Classes non linéaires
- Adapté à tout nombre de caractéristiques



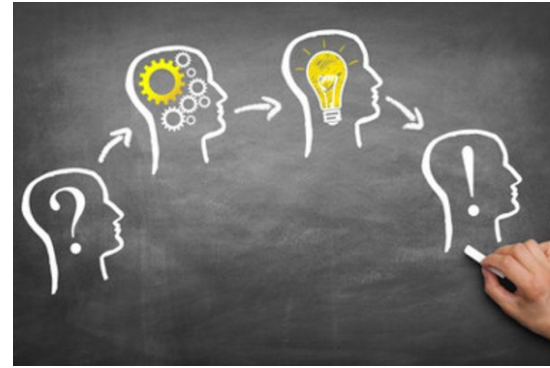
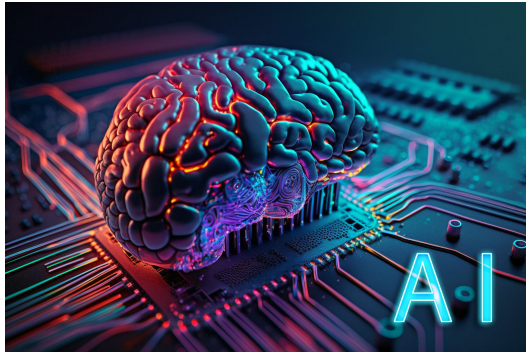
6. Méthode: SVM + ACP

- Mise en oeuvre
- Résultats
 - (base): 60% accuracy
 - (optimaux): 65% accuracy
- Conclusion
 - Classes proches



7. Conclusion

- Difficultés
- Performances
- Limites
- Nouvelles perspectives



Questions ?

