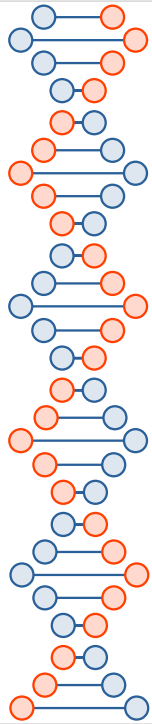


Traitement de données Projet de Holter tensionnel intelligent

Robin SHAMSNEJAD - E2I5



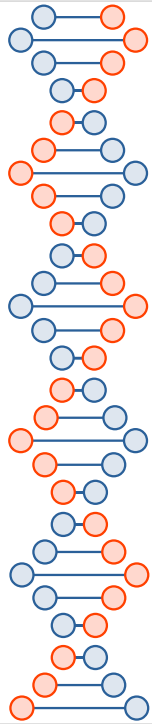
Objectif

→ Créer un système de mesure de la pression artérielle « intelligent »

2

Plusieurs mesures ont été faites sur des patients pendant des journées entières, et donc peuvent contenir des artefacts parmi les mesures exploitables notamment à cause de mouvements parasites.

A partir de l'évaluation de leur validité par un médecin cardiologue, l'objectif est de créer un système permettant de prédire efficacement et en temps réel si une onde de pouls mesurée est exploitable ou non.



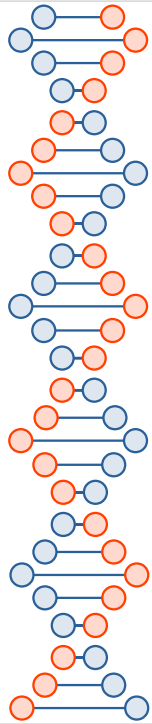
Données en entrée

- data1, data2 : Patient A
- data3 : Patient B

Entraînement	Validation	Test
data1	data2	data3

Les jeux data1 et data2 contiennent des mesures pour un premier patient, et data3 pour un second.

On va donc pouvoir entraîner et vérifier nos modèles avec data1, puis comparer les performances en prédiction intra-sujet avec data2 et en prédiction inter-sujet avec data3.

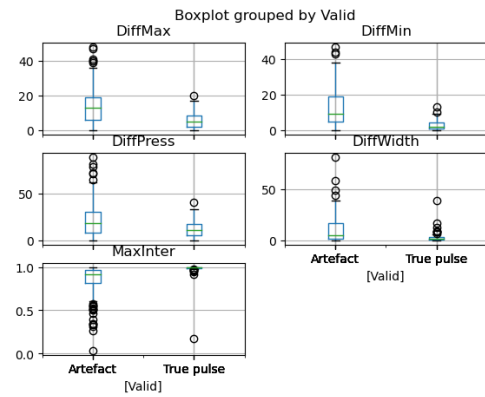
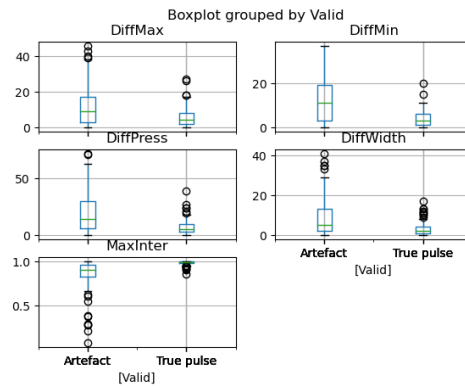
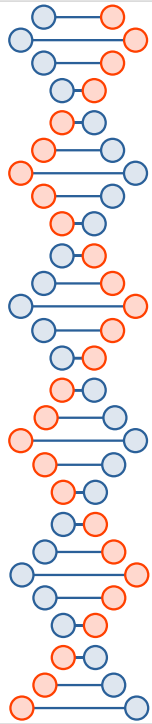


Description des données

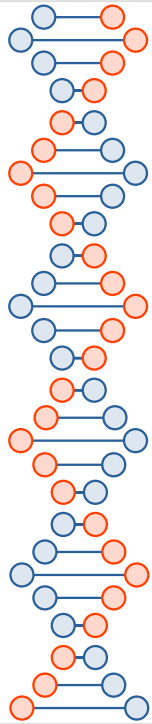
	Nombre de mesures	Nombre de mesures valides	Nombre d'artefacts	Taux d'artefacts
data1	189	108	81	42 %
data2	240	91	149	62 %
data3	218	76	142	65 %

Les mesures ayant donné le jeu data1 semblent avoir été faites dans de meilleures conditions que les autres, car les artefacts y sont minoritaires alors qu'ils sont majoritaires dans les deux autres jeux.

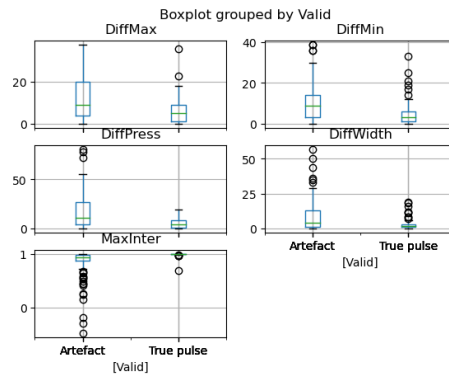
Description des données



(Notes à la diapositive suivante)

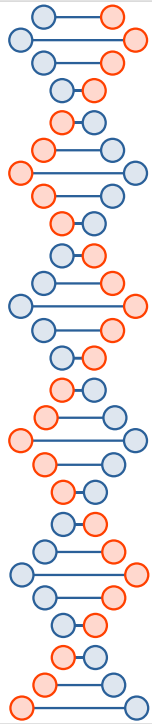


Description des données

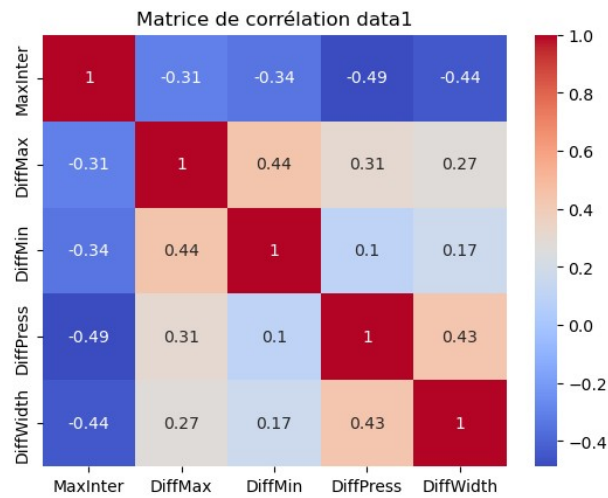


data3

Les données de chaque groupe se chevauchent pour toutes les variables, il est donc impossible d'en utiliser une comme élément discriminant.



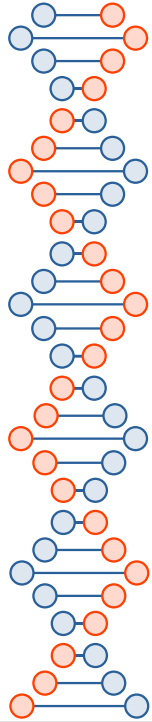
Matrice de corrélation



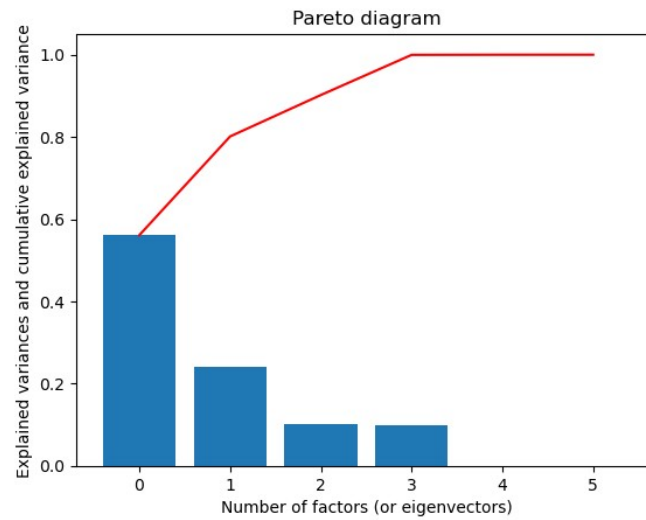
7

La matrice de corrélation nous indique que les variables sont également très peu intercorrélées.

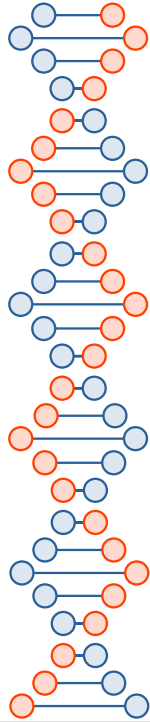
On va donc tenter une ACP afin de trouver des composantes principales dans le jeu de données pour en réduire les dimensions.



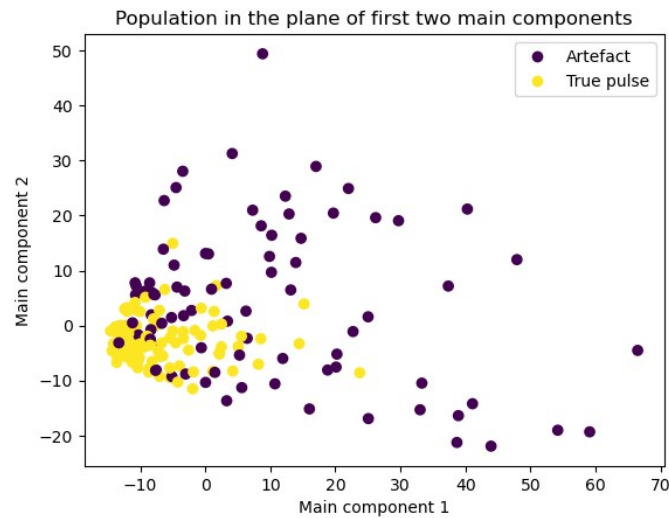
Analyse en Composantes Principales



La variance expliquée cumulée montre que 2 composantes suffisent pour conserver environ 80 % de la variance. On va donc pouvoir faire une ACP à 2 composantes pour la suite



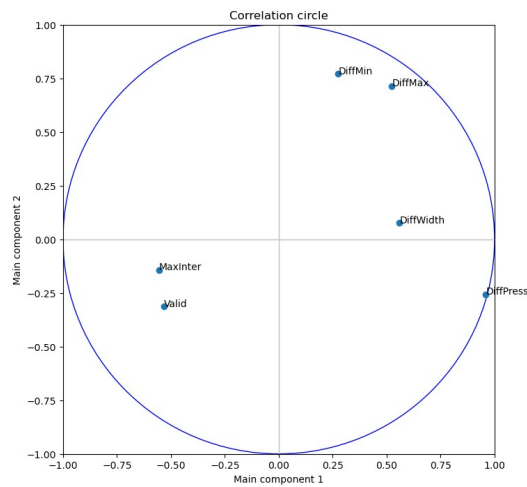
Analyse en Composantes Principales



9

En projetant sur chaque axe, on remarque que chaque groupe a une distribution tendant vers une loi normale, ce qui nous permettra d'utiliser des classifieurs par la suite.

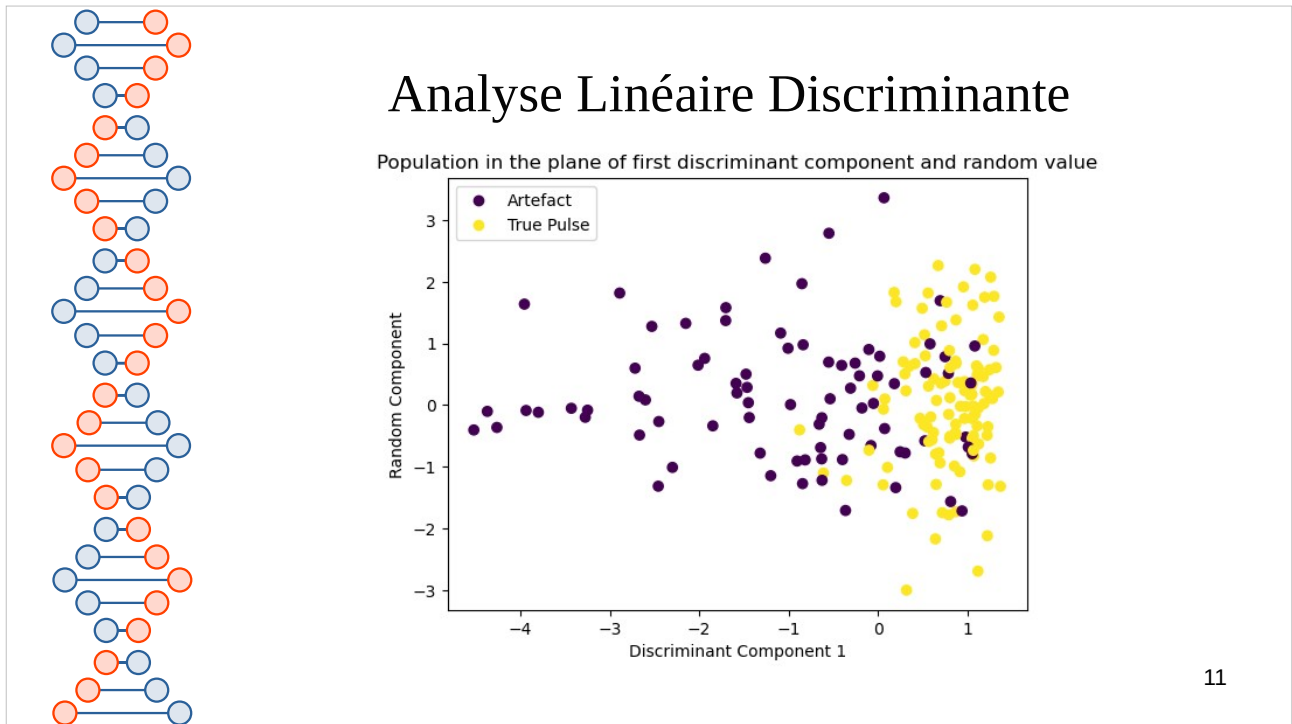
Analyse en Composantes Principales



10

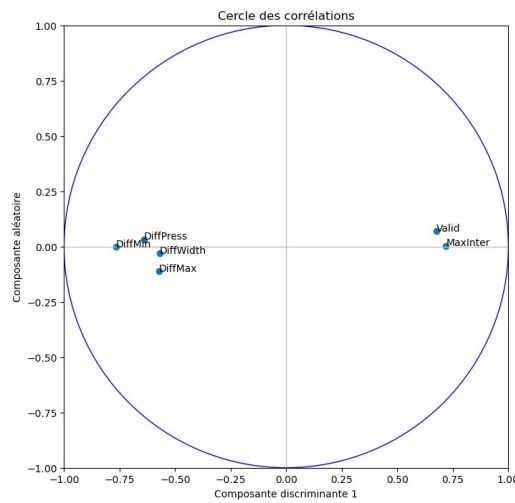
Le cercle des corrélations nous indique que DiffMin et DiffMax sont très corrélées avec la CP2, et DiffPress avec la CP1, mais guère plus.

Difficile à interpréter pour faire des prédictions pour la personne chargée du développement ne pratiquant pas la médecine.



On retrouve également une distribution normale des groupe sur l'axe de la composante discriminante

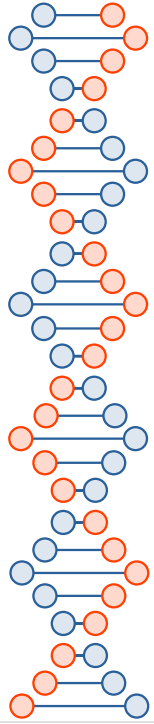
Analyse Linéaire Discriminante



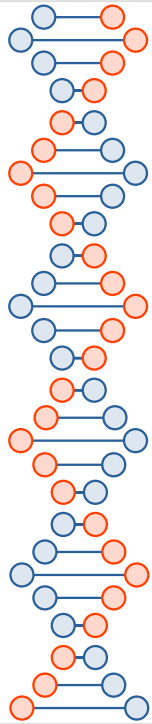
12

Le cercle de corrélation par ALD ne nous donne pas vraiment d'informations exploitables non plus.

Face à toutes ces difficultés, il va falloir utiliser un outil plus puissant pour permettre une prédiction efficace, à savoir un classifieur.



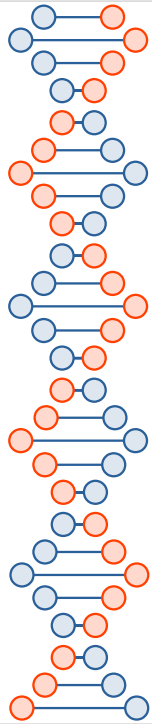
Classifieurs



Classifieurs testés

- Linear Discriminant Analysis (LDA)
 - Quadratic Discriminant Analysis (QDA)
 - Gaussian Naive Bayes (GNB)
 - K Neighbors Classifier (KNC)
-
- Objectif : Trouver le bon compromis entre minimisation des faux négatifs et minimisation des faux positifs

Minimiser les faux négatifs permet d'éviter de prolonger inutilement une mesure qui serait bonne, mais si le modèle présente trop de faux positifs, l'efficacité du système est affectée. Il faudra donc placer le curseur judicieusement lors du choix du modèle de prédiction.

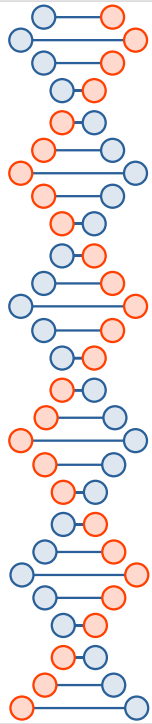


Méthodologie

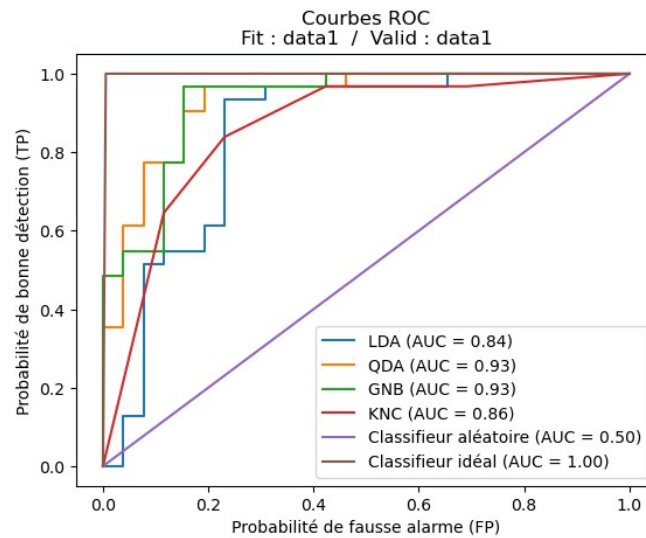
- Division de data1 en deux sous-jeux de données
- Entraînement des 4 classifieurs avec une partie
- Vérification du bon fonctionnement avec l'autre partie
- Entraînement des 4 classifieurs avec data1 entier
- Comparaison des performances de prédiction sur data2 (intra-sujet) et data3 (inter-sujet)

15

data1 et data2, bien que mesurés sur un même patient, semblent avoir été générés dans des conditions différentes : il serait donc intéressant de comparer les performances de prédiction sur les deux. Cependant, comme on utilise data1 pour l'entraînement, il n'est pas très intéressant de le tester sur le même jeu de données. C'est pour quoi dans ce cas là j'ai décidé de diviser les données pour l'entraînement et la validation (avec `train_test_split()` sous Python)



Comparaison des courbes ROC : data1

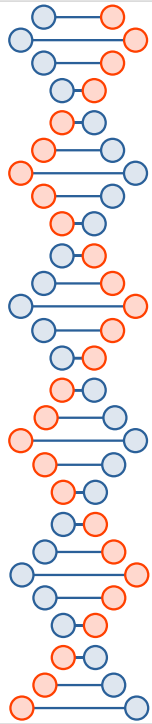


16

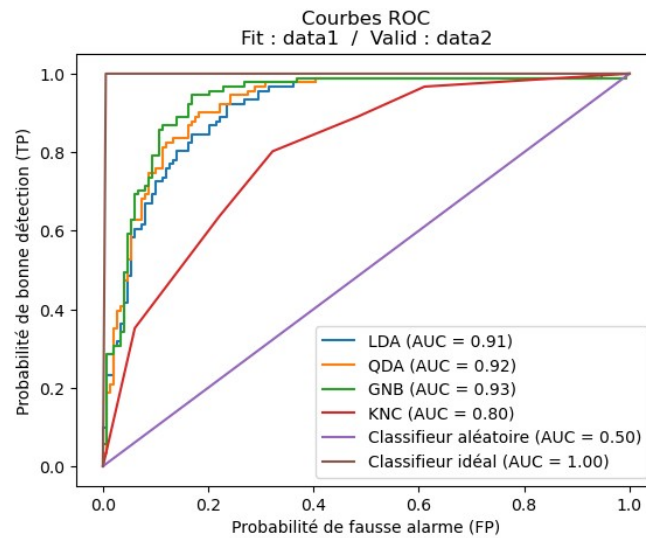
L'entraînement et la validation des classifieurs par le jeu data1 divisé en deux semble montrer que LDA et KNC sont moins performants que les autres sur cette prédiction.

Les deux autres présentent de relativement bons scores avec un AUC de 0.93, mais seront différenciés lors des tests suivants.

À noter cependant : le jeu de données étant relativement petit à l'origine, il l'est encore plus après division dans ce cas là. La précision n'est donc probablement pas très représentative des performances réelles.



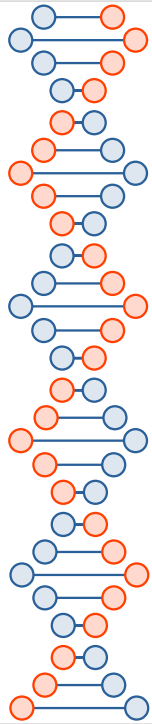
Comparaison des courbes ROC : data2



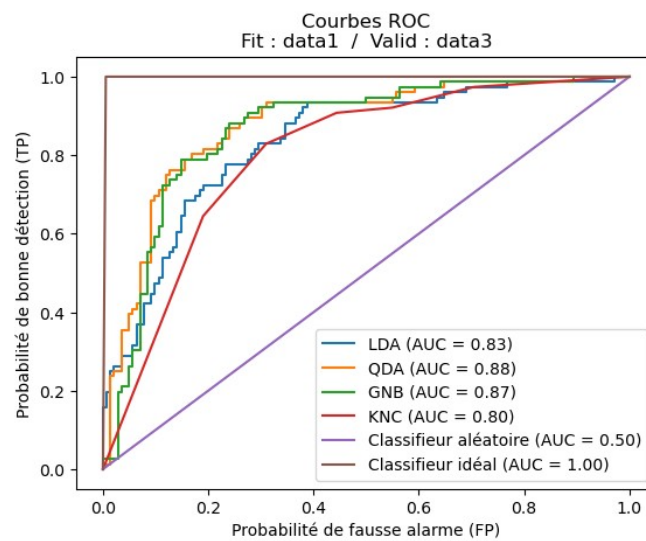
17

Cette fois ci, les modèles ont été entraînés avec data1 complet. Le jeu de test est maintenant data2, nous sommes donc dans le cas intra-sujet.

Cette fois-ci le KNC montre nettement qu'il n'est pas le bon choix par rapport aux autres options, et le meilleur choix semble être GNB qui présente les mêmes performances que dans le cas précédent avec un AUC de 0.93. Cependant QDA et LDA ont également des performances proches, donc ils pourraient être de bons candidats alternatifs si jamais GNB n'est plus une option.



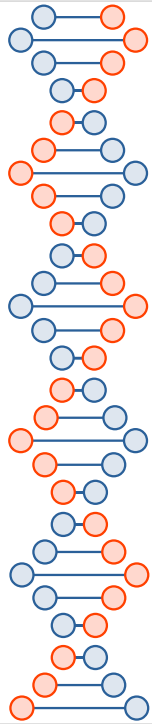
Comparaison des courbes ROC : data3



18

Enfin, on se trouve maintenant dans le cas inter-sujet avec data3.

Dans ce cas, QDA présente un meilleur score avec un AUC de 0.88, cependant GNB reste toujours une bonne option avec un AUC très proche de 0.87.



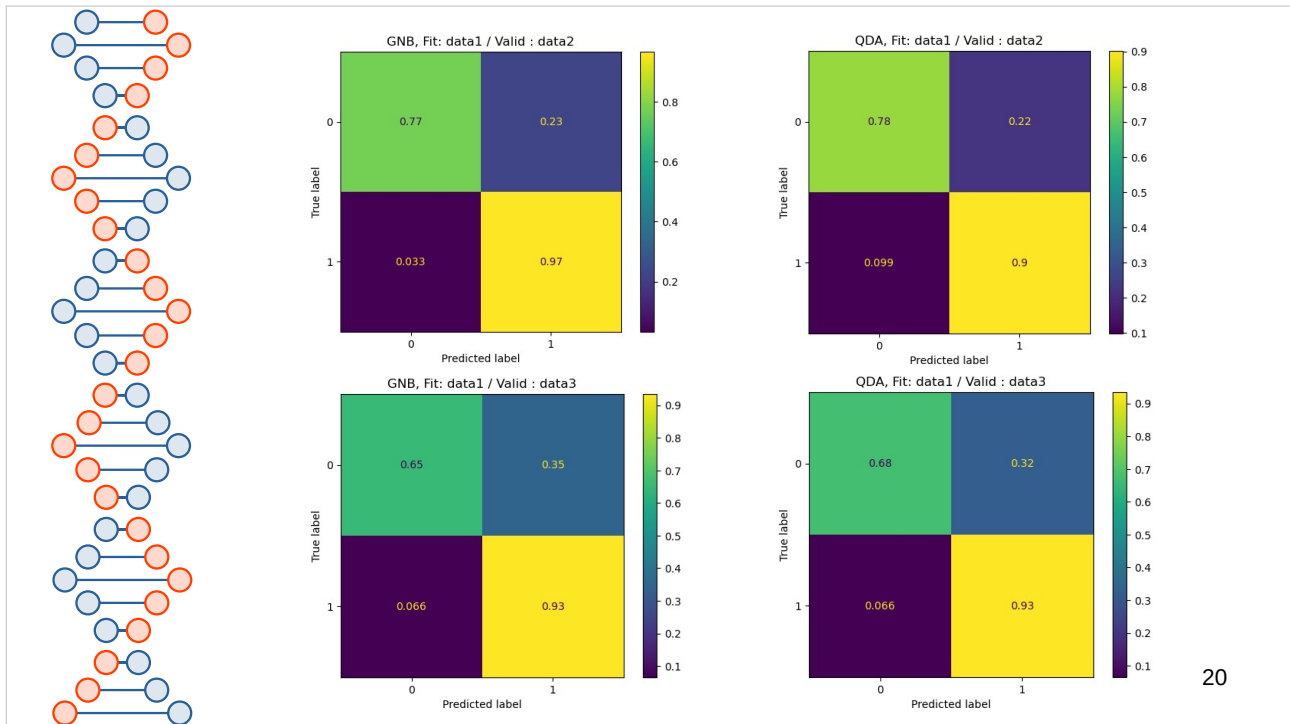
Quel classifieur choisir ?

- KNC : rejeté, car il présente de mauvaises performances dans tous les cas
- LDA : rejeté, car il présente de mauvaises performances dans le cas inter-sujet
- QDA et GNB : candidats à départager

19

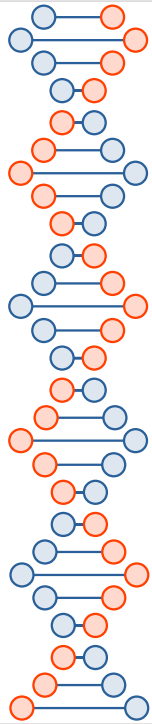
GNB présente de meilleures performances dans le cas intra-sujet, alors que c'est le cas de QDA dans le cas inter-sujet.

Cependant, dans les deux cas les deux classifieurs sont des options viables, il va donc falloir les départager.



Pour tenter de trouver un élément pour départager GNB et QDA, on observe leurs matrices de confusion dans les cas intra- et inter-sujet.

- Dans tous les cas, le taux de faux négatifs est très bas ($< 10\%$), ce qui est une bonne chose pour éviter des mesures qui dureraient trop longtemps et seraient potentiellement douloureuses pour le patient.
- GNB présente de bien meilleures performances en intra-sujet
- Les performances sont similaires dans le cas inter-sujet, à l'exception du fait que le taux de faux positif est plus élevé chez GNB



Conclusion

- Classifieurs rejetés : KNC & LDA
- Classifieur retenu : GNB
 - QDA peut être un bon choix aussi

21

Extrait du cahier des charges :

« Lors de l'installation du holter sur le patient [...], le système fait l'acquisition d'ondes de pouls de référence [...]. »

- Dans ce cas, je choisirais GNB, car il a les meilleures performances en intra-sujet.
- Dans le cas où le cahier des charges serait modifié, et où les ondes de pouls de référence utilisées pourraient être celles d'une autre personne, alors je choisirais plutôt QDA en raison de son taux de faux positifs plus faible (32 % contre 35 % pour GNB), ce qui donnerait moins de travail de tri au médecin.
- Cependant, le taux de faux positifs reste assez élevé avec un tiers des mesures, il pourrait donc être intéressant d'explorer d'autres classifieurs pour tenter d'améliorer ce point.