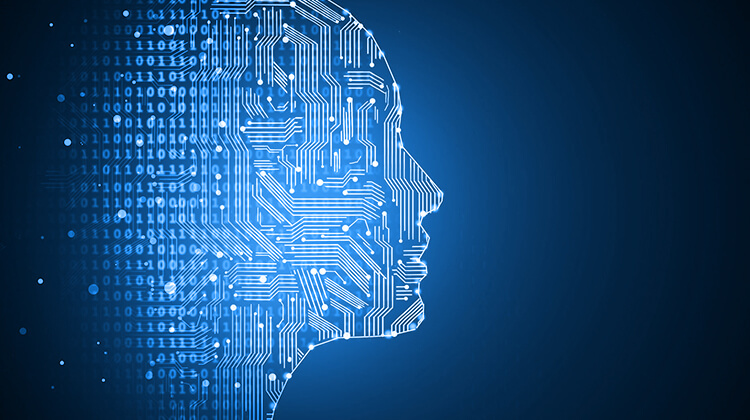
Une image contenant texte, Police, logo, Graphique

Description générée automatiquement

PNDI : De la programmation numérique à la Data Intelligence (Laboratoire : Rapport individuel)



À l'attention de Clerin Frédéric

* Valentin Guillaume (2A)

Table des matières

[Diagramme d’action de la Phase 3 3](#_Toc167450724)

[Premiers résultats 5](#_Toc167450725)

[Tentatives d’améliorations des résultats 6](#_Toc167450726)

[Elimination des valeurs aberrantes 6](#_Toc167450727)

[Augmentation du temps pour le TrainSet et TestSet (1min -> 2min) 7](#_Toc167450728)

[Augmentation du temps pour le TrainSet et TestSet (1min -> +- 3min) 8](#_Toc167450729)

[Génération de pattern à partir d’autres caractéristiques 9](#_Toc167450730)

[Hypothèses 10](#_Toc167450731)

[Génération d’un nouveau vecteur 11](#_Toc167450732)

[Conclusion 12](#_Toc167450733)

# Diagramme d’action de la Phase 3

Une image contenant texte, capture d’écran, Police

Description générée automatiquementUne image contenant texte, capture d’écran, Police, nombre

Description générée automatiquement



# Premiers résultats

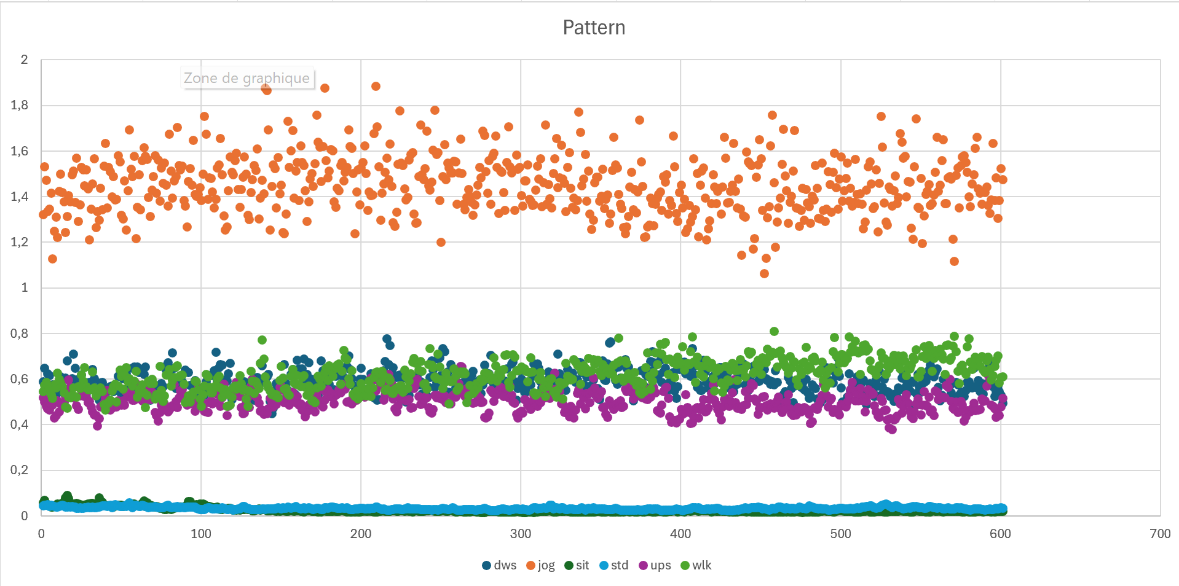
Les premiers résultats obtenus à la suite de la Phase 3 montrent qu’un des six mouvements se distingue avec une identification correcte dans 100 % des cas. Ce mouvement est le numéro 2, appelé « Jog », et il s'avère donc particulièrement facile à reconnaître parmi les autres.

En ce qui concerne le mouvement 3, « Sit », et le mouvement 5, « Ups », ils affichent des taux d’authentification relativement élevés. À l'inverse, les mouvements « dws » (descendre), « std » (se tenir debout) et « wlk » (marcher) présentent des pourcentages d'identification beaucoup plus faibles, rendant leur distinction nettement plus difficile.

En globalisant les résultats, le taux d’identification correct pour l'ensemble des mouvements est de 52,77 %. Ce chiffre est assez bas et signifie qu'il y a presque une chance sur deux de ne pas réussir à identifier correctement un mouvement.

Ces résultats sont encore plus évidents lorsque l'on examine le diagramme en nuage de points généré à partir du modèle des mouvements (PatternSet). Il est clair que le mouvement 2 se distingue nettement des autres, le rendant facilement identifiable et confirmant effectivement les 100 % de reconnaissance obtenus après l'exécution du programme.

En revanche, les autres mouvements se chevauchent considérablement. On observe notamment que les mouvements « dws » , « ups » et « wlk » présentent des similarités marquées, rendant leur différenciation difficile. De même, les mouvements « sit » et « std » se ressemblent fortement, ce qui complique leur identification précise.



# Tentatives d’améliorations des résultats

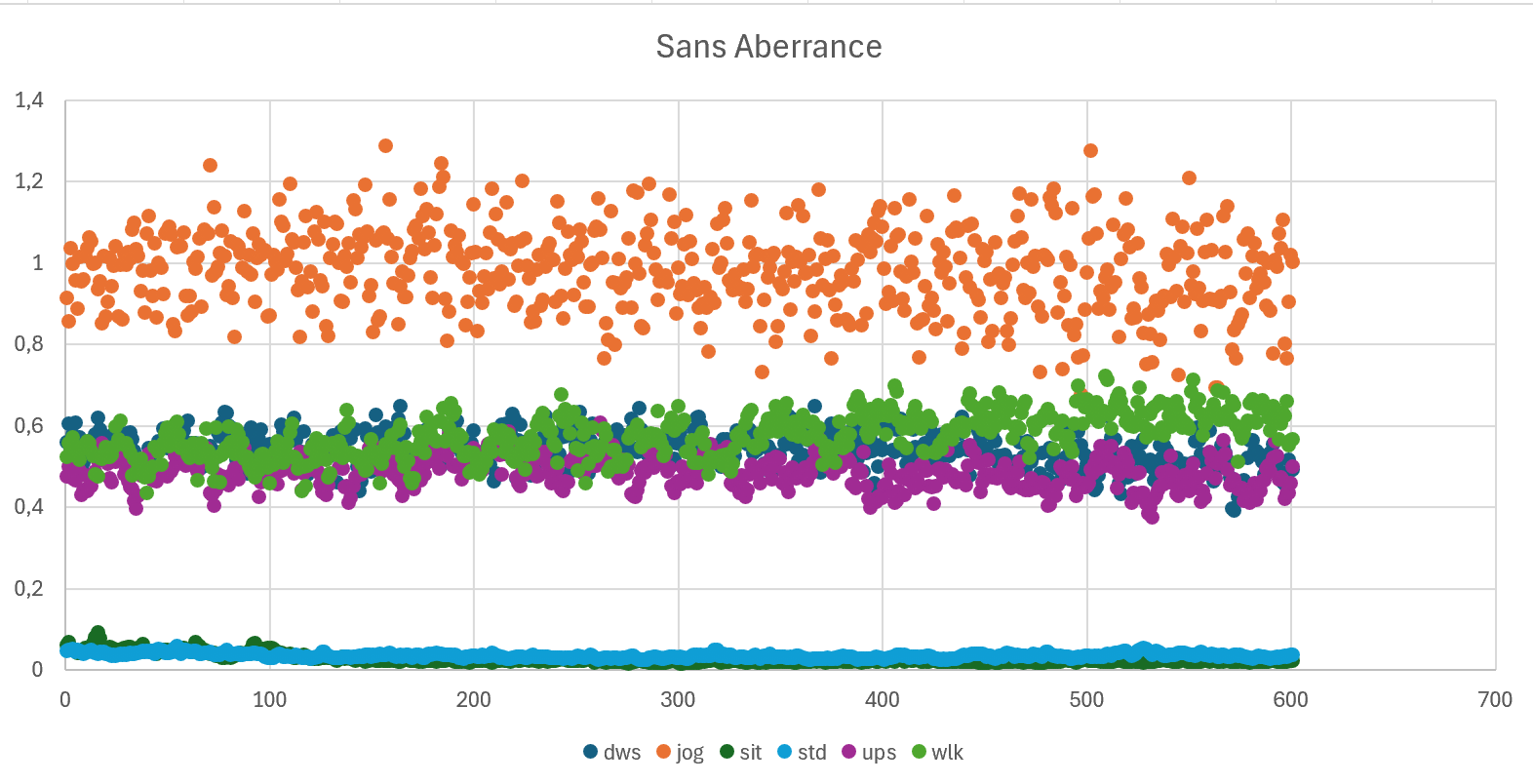
## Elimination des valeurs aberrantes

Pour cette première amélioration, il est essentiel de ne prendre en compte que les données valides et de rejeter les données aberrantes. Selon la Phase 1 complément, une donnée est considérée comme aberrante si elle s'écarte de la moyenne de plus de 3 écarts-types (cf. Phase 1 complément).

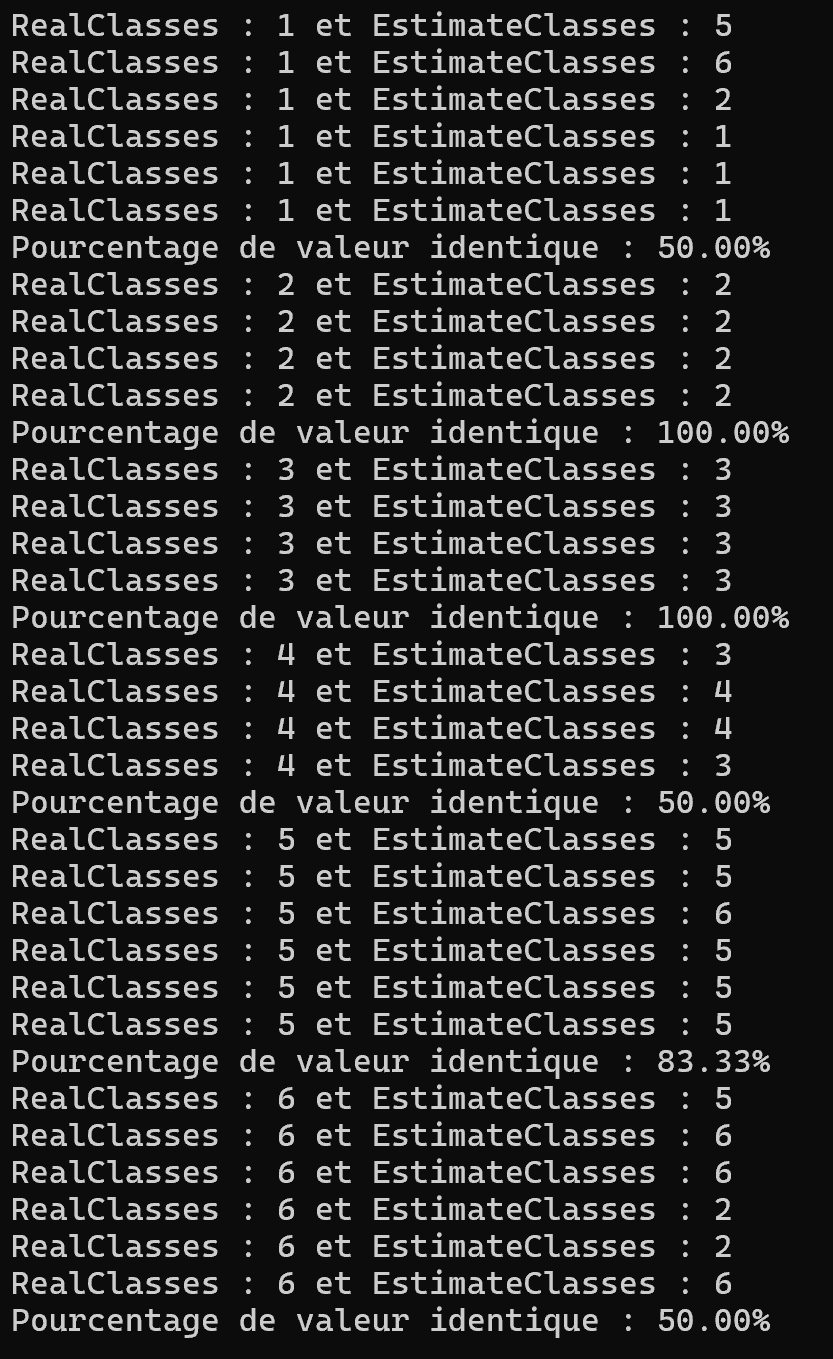
Dans cette étape, j'ai donc utilisé uniquement les valeurs non aberrantes pour créer de nouveaux vecteurs d'accélération. Ensuite, un nouveau modèle a été généré pour évaluer si ces ajustements améliorent les résultats.

Nous constatons une légère amélioration pour le mouvement 1, dont le taux de reconnaissance par rapport à sa RealClasses passe de 16,67 % à 50 %. Cependant, les autres mouvements ne montrent pas de nouveaux résultats, bien que le pourcentage global d'identification des mouvements ait augmenté de 52,77 % à 58,33 %.

Cette amélioration est donc bénéfique et nécessaire pour la reconnaissance des différents mouvements et ne doit pas être négligée.



## Augmentation du temps pour le TrainSet et TestSet (1min -> 2min)

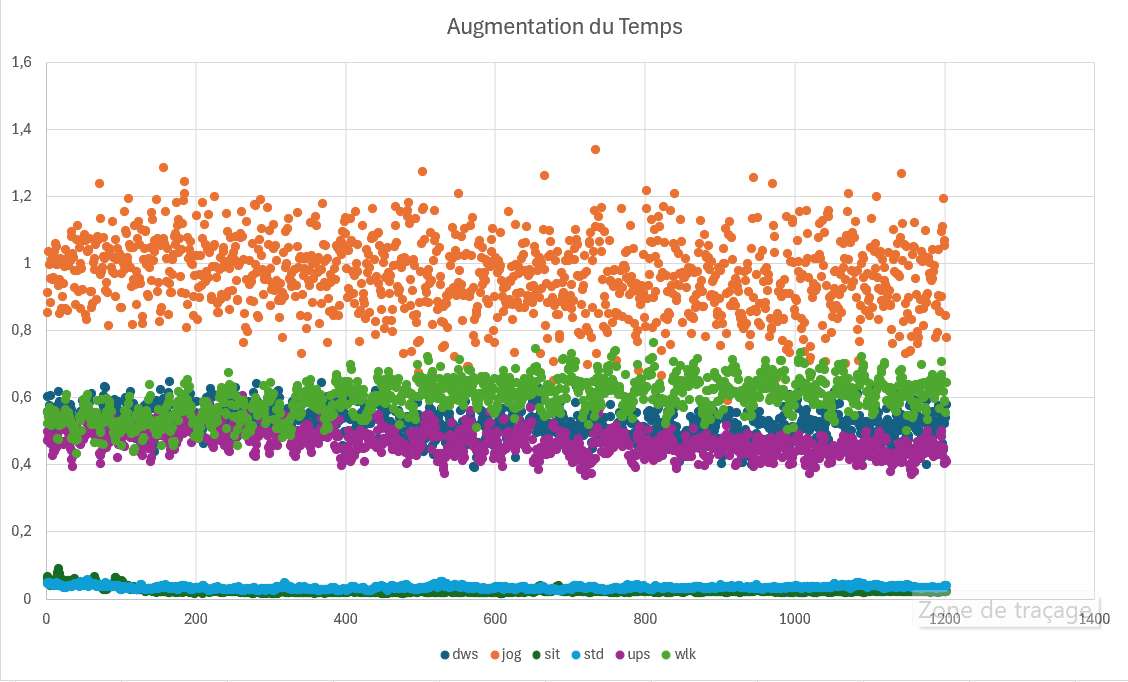
Pour cette deuxième amélioration, j'ai décidé de reprendre l'approche précédente en conservant uniquement les données jugées valides, tout en ajoutant une optimisation supplémentaire : l'augmentation de la durée des ensembles d'entraînement (TrainSet) et de test (TestSet) de 600 à 1200 secondes (de 1 minute à 2 minutes).

Cette modification est particulièrement pertinente, car elle permet de disposer de davantage de données, augmentant ainsi la précision de la distinction des mouvements.

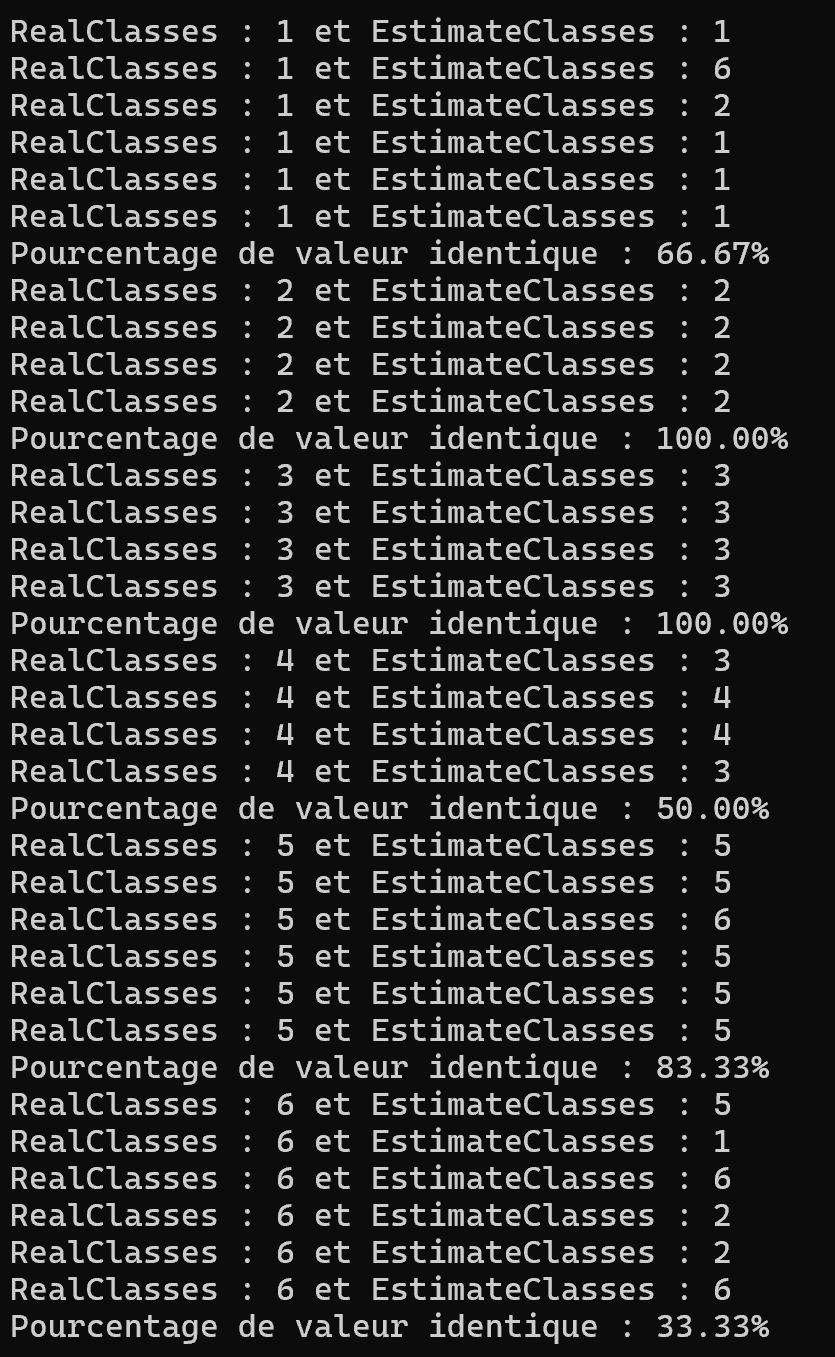
Grâce à cette amélioration, nous observons des résultats significatifs : le taux de reconnaissance du

* Mouvement 3 passe de 75 % à 100 %,
* Mouvement 4 de 25 % à 50 %,
* Mouvement 6 de 16,67 % à 50 %.

Par conséquent, le pourcentage global de distinction des mouvements s'élève de 58,33 % à 72,22 %, ce qui représente une amélioration substantielle et non négligeable.

En conclusion, cette optimisation contribue grandement à la précision de la reconnaissance des différents mouvements, démontrant ainsi l'importance d'une période d'entraînement et de test plus longue.

## Augmentation du temps pour le TrainSet et TestSet (1min -> +- 3min)

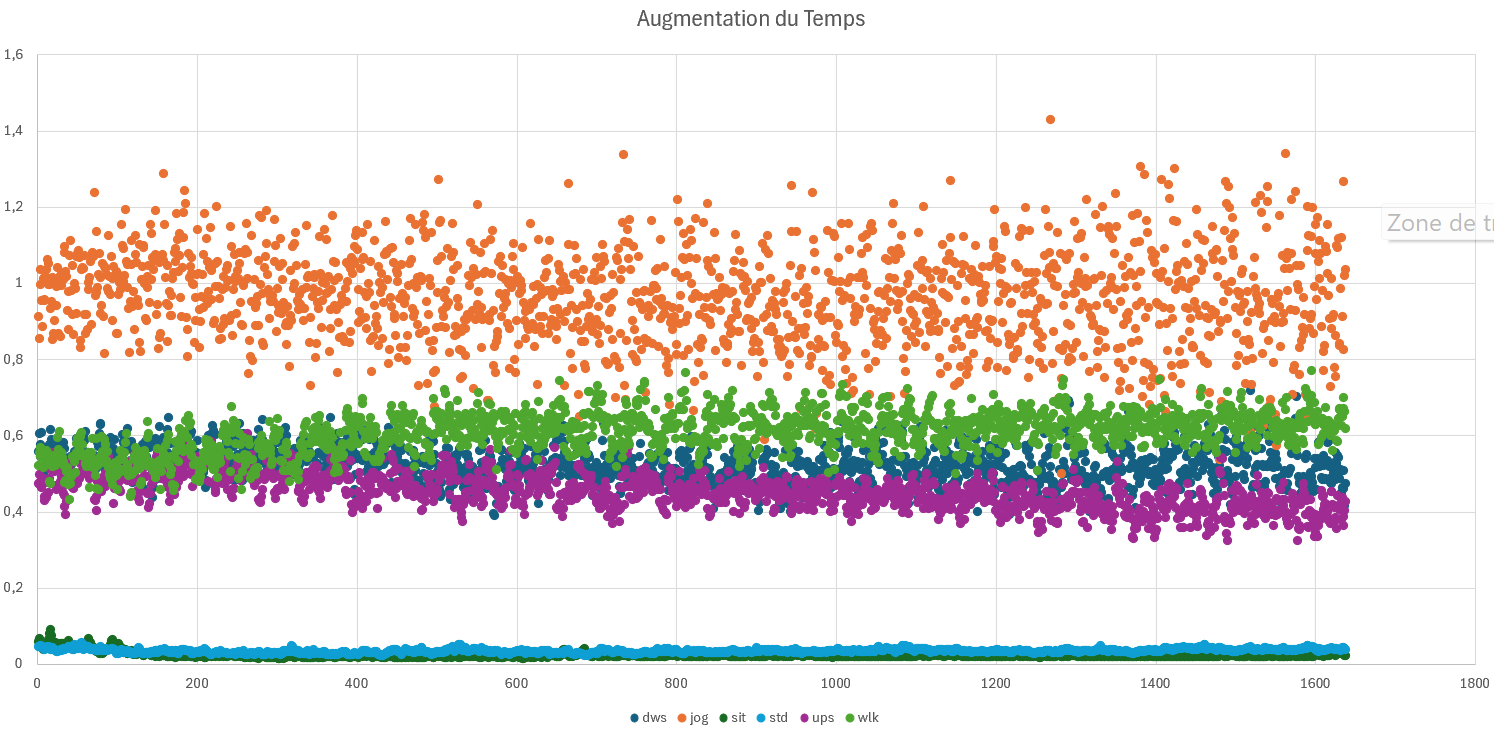
  
  
Dans cette amélioration, je repars des données valides, mais cette fois-ci, j’augmente la durée du TrainSet et du TestSet au maximum permis par les données initiales.

Nous observons que le taux de reconnaissance du mouvement 1 a augmenté de 50 % à 66,67 %, bien que cela ait entraîné une baisse de la reconnaissance du mouvement 6, passant de 50 % à 33,33 %.

Le pourcentage global de distinction des mouvements reste stable à 72,22 %.

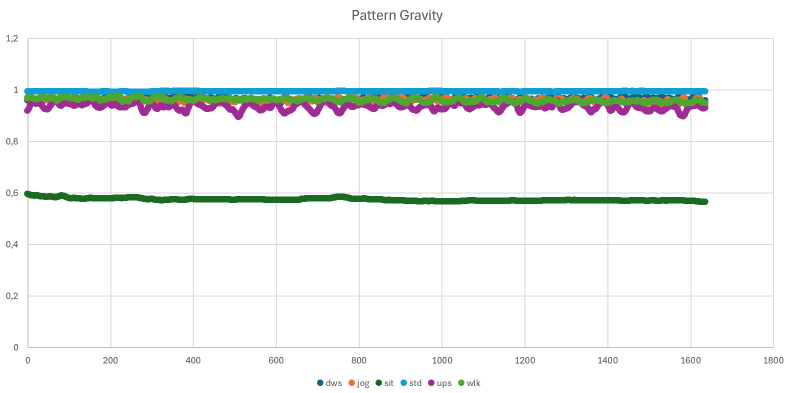
Il est clair que plus la collecte de données pour générer un modèle est étendue, plus la précision de la distinction des mouvements s’améliore.

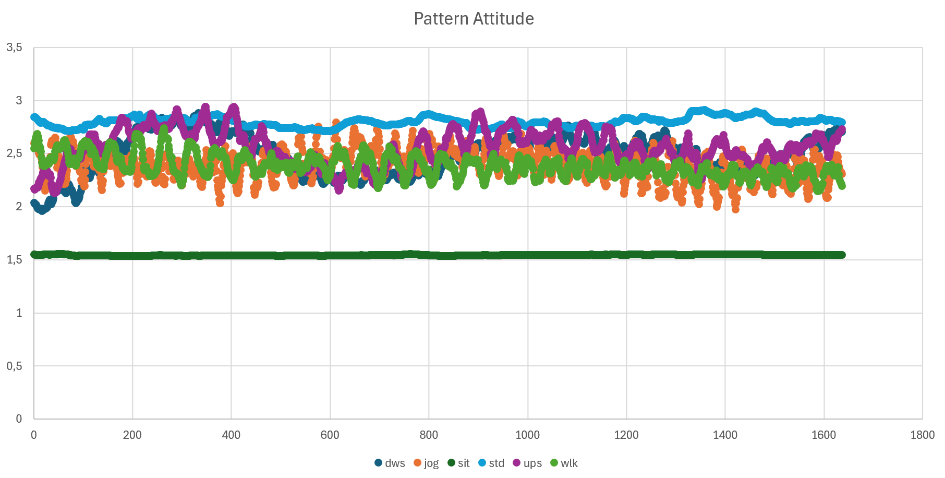
En examinant les graphiques en nuage de points des différents modèles, nous constatons que l'augmentation du temps permet une meilleure séparation et distinction des mouvements « wlk », « dws » et « ups ». Cependant, il demeure difficile de différencier les mouvements « sit » et « std ».

En conclusion, l'augmentation de la durée de collecte des données initiales est cruciale pour générer un modèle précis et pertinent, capable de distinguer efficacement les différents mouvements.

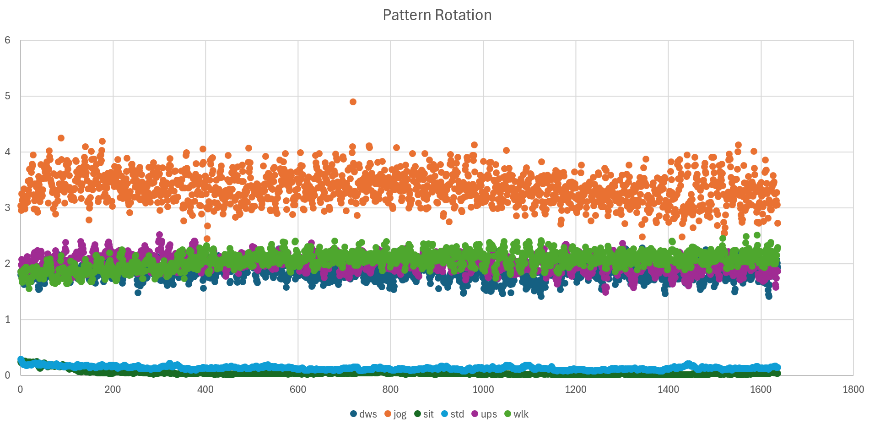
# Génération de pattern à partir d’autres caractéristiques

Dans cette section, j'ai exploré les effets de la génération de modèles à partir d'autres composants des mouvements tels que l'attitude, la gravité et la rotation. Pour cela, j'ai créé des vecteurs spécifiques à ces composants afin de générer des modèles et d'observer les résultats obtenus.

 Ce graphique illustre un nuage de points du modèle créé à partir des données de gravité. On constate que le mouvement « sit » est le plus représentatif et distinctif. En revanche, il est difficile de distinguer les autres mouvements.



Le graphique suivant représente le modèle lié à l'attitude. Ici aussi, le mouvement « sit » se démarque nettement. Toutefois, on peut observer des courbes intéressantes pour les autres mouvements, bien qu'elles ne soient pas aussi distinctes.



Enfin, ce dernier modèle met en évidence les caractéristiques de rotation. Il ressemble beaucoup au modèle basé sur l'accélération, mais il est important de noter que, dans ce cas, la distinction entre les mouvements « dws », « ups » et « wlk » au fil du temps n'est pas présente.

Ces observations montrent que chaque composant des mouvements offre différents niveaux de distinction et de clarté pour la reconnaissance des mouvements, avec le mouvement « sit » étant le plus facilement identifiable dans la plupart des cas.

# Hypothèses

À la suite des premières observations concernant les améliorations du modèle basé sur l'accélération, une hypothèse émerge : il est possible que plus la durée de collecte des données pour générer un modèle est longue, plus la distinction entre les mouvements « dws », « jog », « ups » et « wlk » devienne significative.

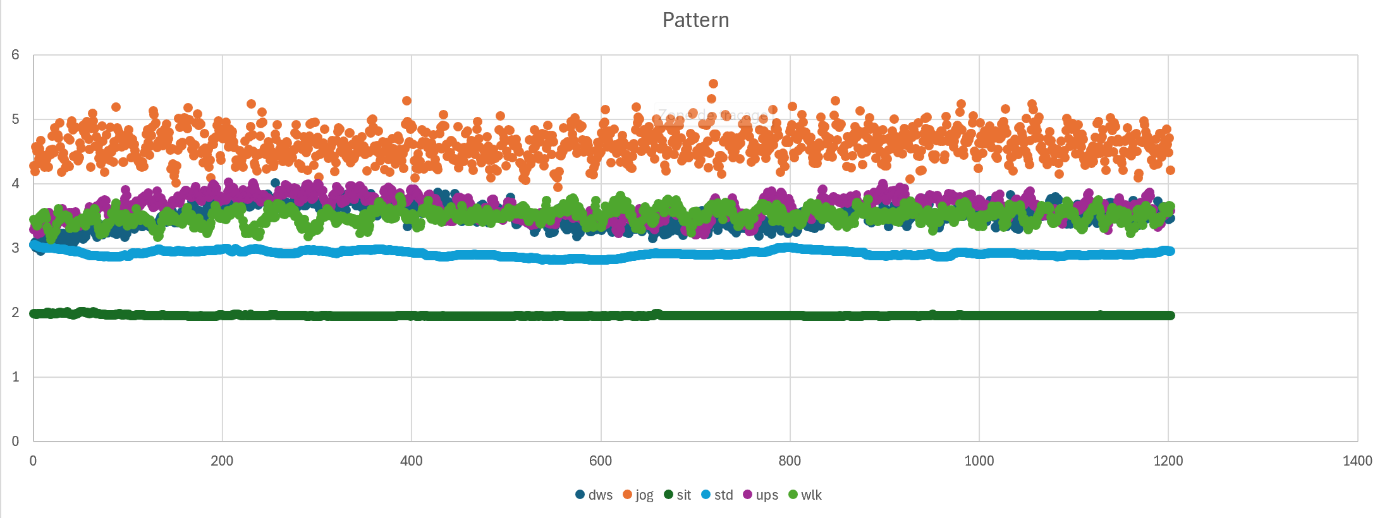
Cependant, pour différencier le mouvement « sit » du mouvement « std », une autre hypothèse pourrait être avancée : l'inclusion des composants de gravité et d'attitude pourrait améliorer cette distinction.

En effet, puisque les modèles basés sur la gravité et l'attitude montrent une distinction claire entre ces deux mouvements, il est légitime de se demander si leur inclusion dans la génération de vecteurs ou de modèles pourrait renforcer cette distinction. De même pour l’inclusion de la rotation pour garder le mouvement « jog » bien distinctif des autres mouvements.

De plus, il est également nécessaire de filtrer toutes les valeurs considérées comme aberrantes dans le cas de la gravité, de l'attitude et de la rotation. Cependant, faute d'informations suffisantes pour vérifier cette dernière, cette étape ne pourra malheureusement pas être effectuée.

En résumé, l'hypothèse suggère que l'augmentation de la durée de collecte des données peut améliorer la distinction entre certains mouvements, tandis que l'inclusion des composants de gravité et d'attitude pourrait renforcer la distinction entre d'autres. Ces réflexions soulignent l'importance d'explorer différentes approches pour améliorer la précision de la reconnaissance des mouvements.

# Génération d’un nouveau vecteur

Dans cette section, j'ai entrepris de générer un vecteur caractéristique regroupant tous ces composants afin d'observer leurs effets sur une période de 1200 secondes. Pour ce faire, j'ai calculé la racine des carrés de tous les composants (Attitude, Gravité, Rotation, Accélération).

Ce processus met en évidence une nette distinction entre certains mouvements. Nous remarquons que le mouvement « jog » reste toujours facilement identifiable. De plus, cette fois-ci, nous constatons que les mouvements « sit » et « std » sont également très distinguables.

Cependant, la séparation observée précédemment entre les mouvements « dws », « ups » et « wlk » n'est plus aussi marquée. Cette observation suggère que l'inclusion de plusieurs caractéristiques en plus de l'accélération permet une meilleure distinction entre certains mouvements avec ses désavantages pour d’autres.

En résumé, cette approche met en lumière l'importance de prendre en compte plusieurs composants dans la génération de vecteurs caractéristiques pour améliorer la précision de la distinction entre les différents mouvements.

# Conclusion

L'extension de la période de collecte de données peut être une ressource précieuse pour distinguer certains mouvements dans le temps. En outre, l'intégration de plusieurs paramètres supplémentaires dans la génération de vecteurs et de modèles peut également améliorer la reconnaissance des mouvements.

La meilleure précision que j'ai pu obtenir en ne tenant compte que de l'accélération est de 72,22 % pour tous les mouvements, en utilisant une durée de collecte maximale et en éliminant les valeurs aberrantes.

En incluant plusieurs caractéristiques telles que la gravité, l'attitude et la rotation, j'ai pu obtenir un modèle qui distingue correctement 3 des 6 mouvements. Cependant, en raison du manque de connaissances et de temps, je n'ai pas pu obtenir de résultats meilleurs que ceux mentionnés.

Je peux seulement émettre l'hypothèse qu'introduire une nouvelle formule pour générer des vecteurs en formalisant les données pourrait être bénéfique, de même qu'une autre formule de calcul de distance que celle de la distance euclidienne pourrait être pertinente.

En conclusion, bien que ces résultats représentent une avancée significative dans la reconnaissance des mouvements, il reste encore des opportunités d'amélioration à explorer pour optimiser davantage la précision et l'efficacité de cette analyse.