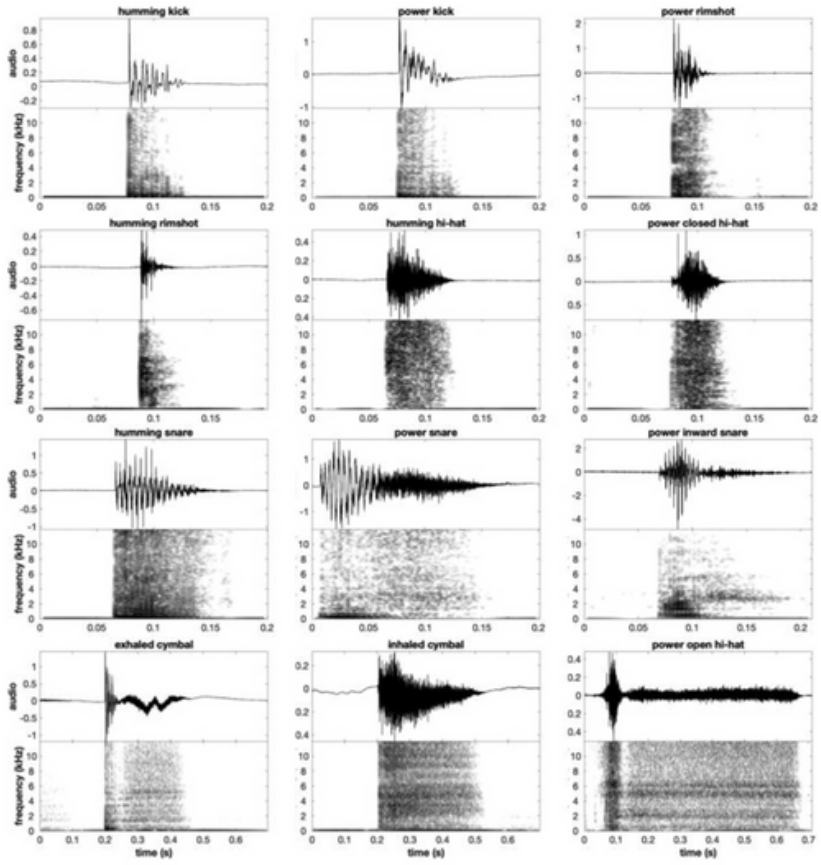
UNIVERSITÉ TOULOUSE III PAUL SABATIER — Université de Toulouse

OCTOBRE 2023

Analyse des Techniques de Classification pour la Détection de Sons de Beatbox



RÉALISÉ PAR : NOUH CHELGHAM ET NOURA FAIZ

Introduction

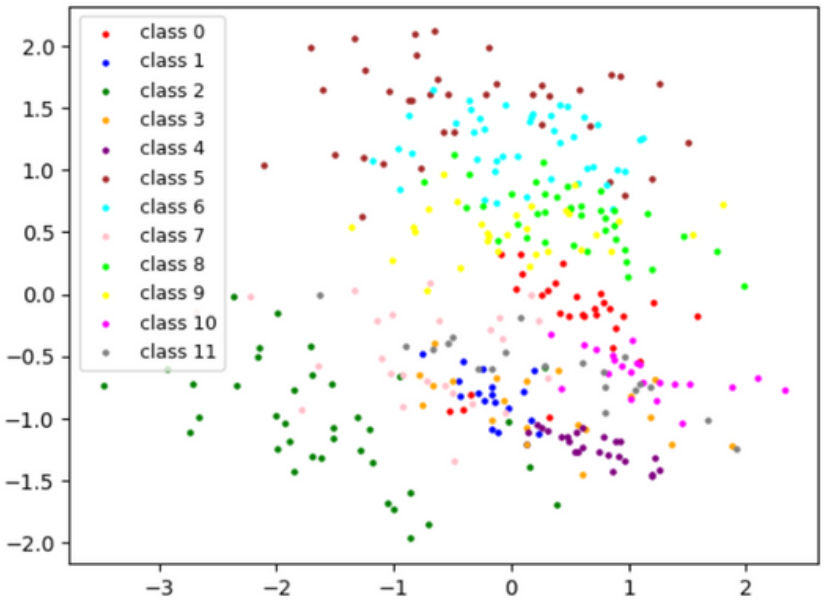
L'art du beatbox consiste à produire des sons de percussions et d'instruments avec la bouche, a connu une popularité grandissante au fil des années. Dans ce travail pratique, nous nous plongeons au cœur de ce phénomène musical en explorant la possibilité de détecter et de reconnaître automatiquement différents sons de beatbox. Nous disposons pour cela d'une base d'enregistrements sonores au format .wav, que nous cherchons à analyser pour identifier 12 sons de beatbox spécifiques.

Classification sans prétraitement

I. Méthode supervisée (KNN : k nearest neighbors)

Le KNN est un algorithme de classification supervisé. Pour chaque observation $x_i = (x_i^1, x_i^1, x_i^2, \ldots x_i^n)$, nous avons une étiquette associée **yi** , qui représente la classe de x_i
Pour classer un nouveau vecteur (x) à l'aide de **KNN** :

- Calculez la distance (souvent euclidienne, mais d'autres métriques peuvent être utilisées) entre x et chaque échantillon xi dans l'ensemble d'entraînement.
- Identifiez les k échantillons xi les plus proches de x .
- La classe attribuée à x sera celle qui est majoritairement représentée parmi ces **k** voisins les plus proches.

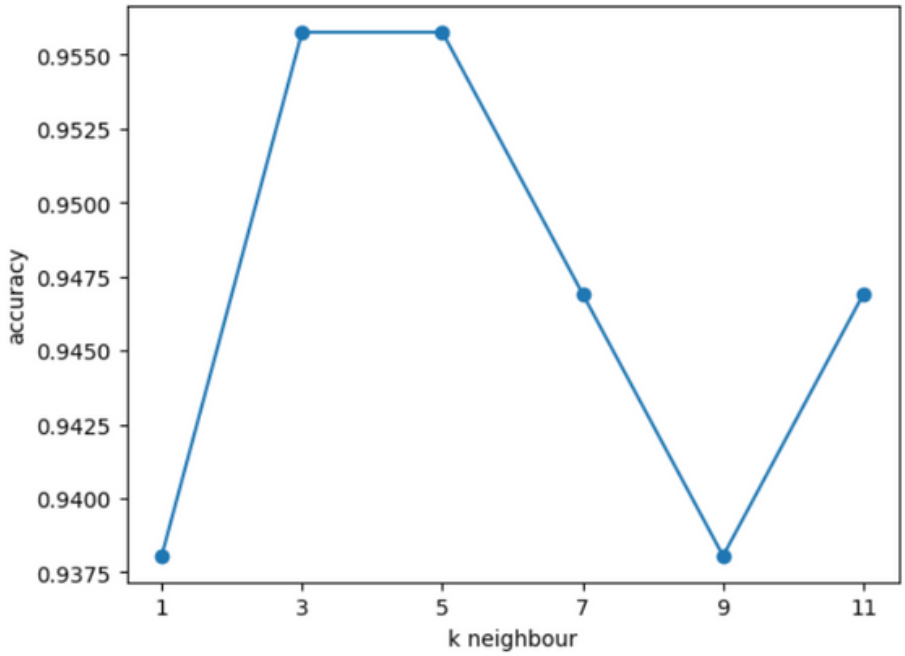


A. Étude sur les paramètres inhérents à la méthode supervisée (k-NN)

Le choix du paramètre k dans l'algorithme k-NN est crucial pour sa performance. La valeur optimale de k varie selon les données et nécessite une détermination expérimentale. Cette section analyse l'effet de différentes valeurs de k sur notre modèle.

1. Impact de la valeur de k sur la précision :

Le graphique ci-dessous illustre comment la précision du modèle k-NN fluctue avec différentes valeurs de k:

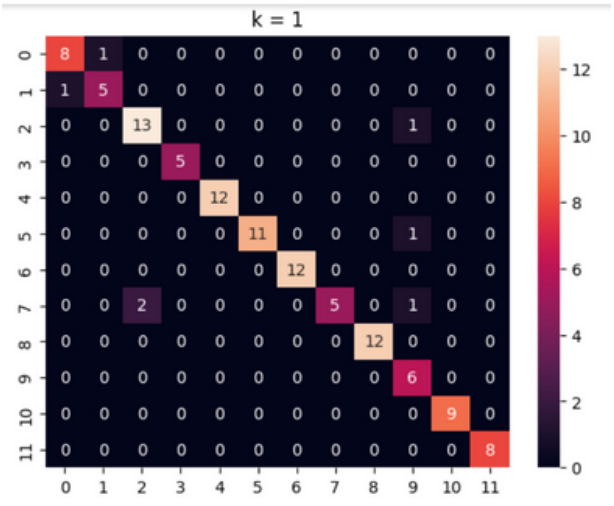


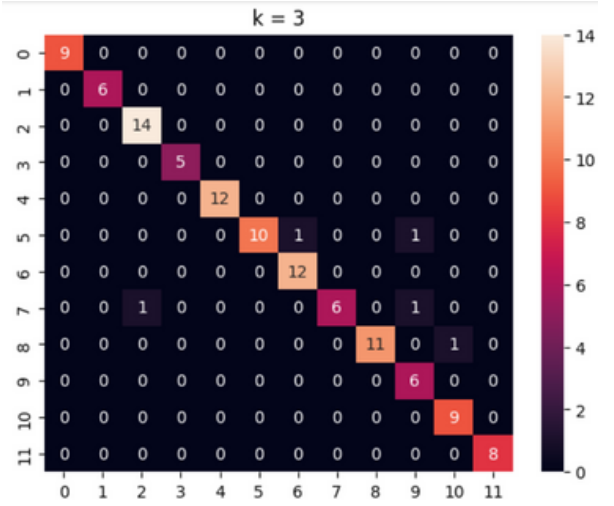
Selon le graphique, la précision atteint son apogée à k=5, puis diminue à k=7, avant de remonter légèrement à k=9 et k=11. Cela suggère que k=5 est la valeur la plus optimale pour cette dataset en termes de précision.

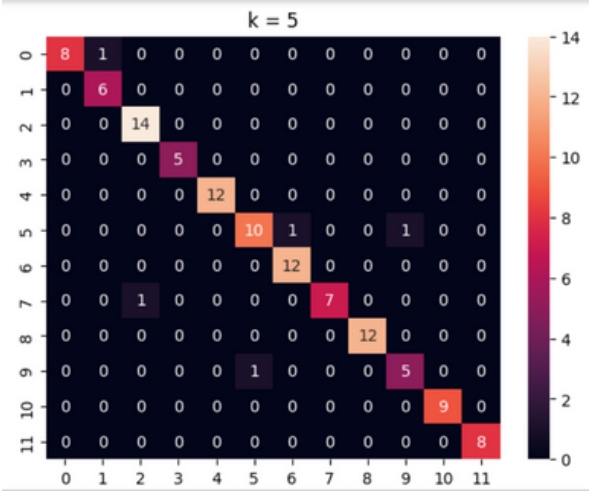
2. Matrices de confusion pour différentes valeurs de k :

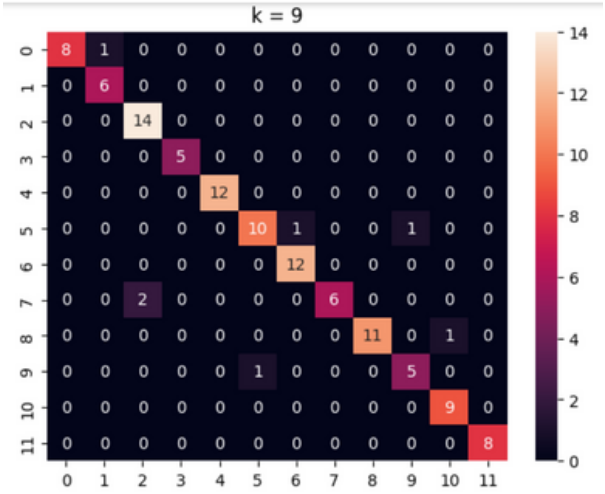
Les matrices de confusion permettent de comprendre comment le modèle classe chaque son de beatbox pour différentes valeurs de k. Voici les matrices de confusion pour quelques valeurs sélectionnées de k:

A. Étude sur les paramètres inhérents à la méthode supervisée (k-NN)









- Pour chaque matrice, la diagonale montre les bonnes classifications.
- Hors de la diagonale, les valeurs montrent les erreurs de classification.

_ Pour k = 5_:
la plupart des classes soient correctement classées, il existe quelques erreurs de classification, en particulier pour les classes 5 et 7.

A. Étude sur les paramètres inhérents à la méthode supervisée (k-NN)

3.Précision (Accuracy):

C'est une mesure de performance pour les modèles de classification. Elle représente le pourcentage de prédictions correctes par rapport au total des prédictions faites. Elle est calculée comme suit:

Accuracy = **(**Nombre total de prédictions**)/(**Nombre de prédictions correctes**)**

```
k = 1   |       accuracy = 0.9380530973451328
k = 3   |       accuracy = 0.9557522123893806
k = 5   |       accuracy = 0.9557522123893806
k = 7   |       accuracy = 0.9469026548672567
k = 9   |       accuracy = 0.9380530973451328
k = 11  |       accuracy = 0.9469026548672567
```

Interprétation : À partir des résultats que nous avons obtenus pour différentes valeurs de k, nous observons que notre modèle atteint une accuracy la plus élevée de 95.58% pour k = 3 et k = 5. Cela signifie que pour ces valeurs de k, notre modèle prédit correctement 95.58% des cas dans notre ensemble de test. Cette haute accuracy suggère que notre modèle est performant et fiable pour ces paramètres. De plus, la faible variation de l'accuracy avec différentes valeurs de k indique la robustesse de notre modèle face à ces variations.

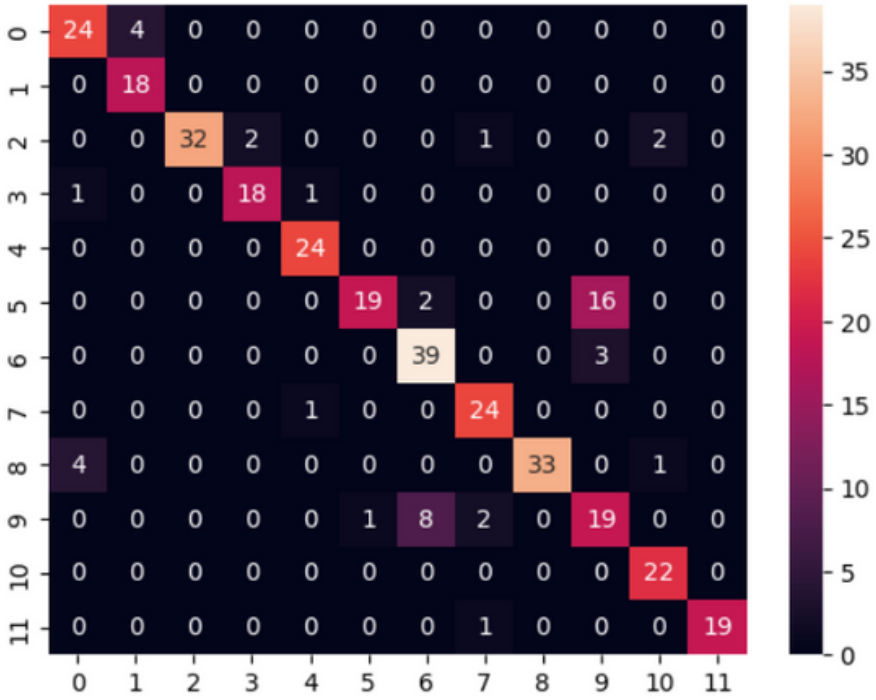
II. Méthode non supervisée (K-means)

C'est un algorithme non supervisé de clustering qui permet de séparer les données en groupes homogènes ayant des caractéristiques communes. L'objectif est de diviser les points en **k** groupes, appelés clusters, homogènes et compacts. Pour chaque point,

- on calcule la distance euclidienne entre ce point et chacun des centroïdes puis on l'associe au centroïde le plus proche c'est-à-dire celui avec la plus petite distance.
- Le résultat final est obtenu lorsque les nouveaux centroïdes ne bougent plus des précédents

B. Étude sur les paramètres inhérents à la méthode non supervisée (K-means)

1. Matrices de confusion :



Score de précision : 0.8533724340175953

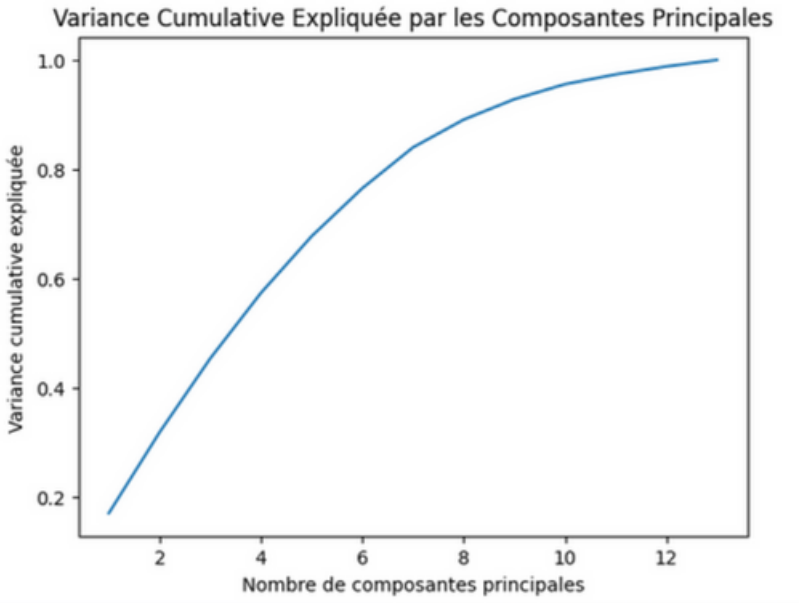
Interprétation : La diagonale principale de la matrice indique où le modèle a correctement identifié les clusters. Par exemple, 24 points ont été correctement classés dans le cluster 0. Cependant, des erreurs existent, comme les 4 points qui appartenaient en réalité à un autre cluster mais ont été classés dans le cluster 0.
**Le score de précision global** de **85,33%** indique que, dans 85,33% des cas, le modèle a assigné correctement les points à leurs clusters respectifs.

Classification avec prétraitement

A. Méthode supervisée avec réduction de dimension par ACP

L'analyse en composantes principales (ACP) est une méthode statistique multivariée qui permet de réduire la complexité d'un ensemble de données en projetant les variables sur un espace de dimensions inférieures. Elle permet de déterminer les relations entre les variables et de les visualiser sous forme de graphiques.

L'ACP est souvent utilisée pour explorer des corrélations entre les variables, identifier des groupes homogènes ou des observations atypiques, et pour réduire le nombre de variables dans un modèle.



Interprétation: Le graphique montre la variance cumulée expliquée par les composantes principales. Il nous guide sur le nombre de composants à conserver pour représenter au mieux la variance totale des données. Le point idéal à choisir est celui où la courbe commence à s'aplatir, indiquant que les composants supplémentaires n'ajoutent que peu de variance expliquée. Dans ce cas, "**12**" semble être le choix optimal, mais "**10**" pourrait également être considéré.

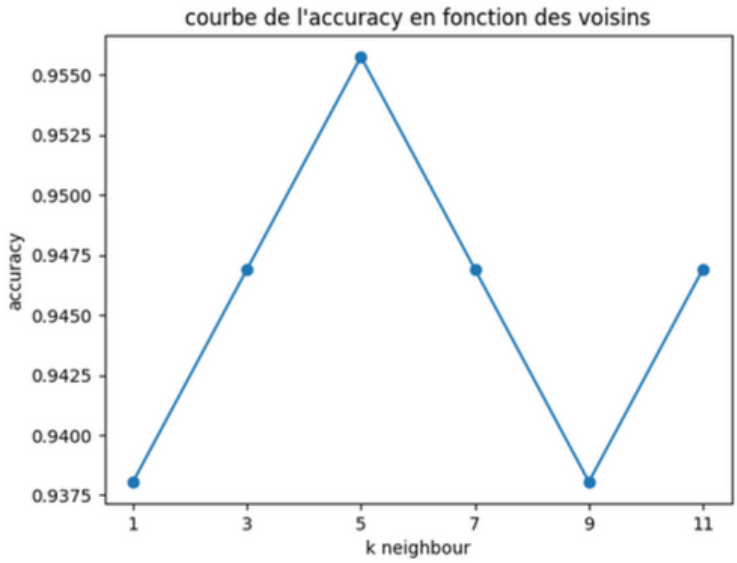
- Après avoir déterminé le nombre optimal de composantes principales à partir de notre analyse, nous fixons `num_components` à 12.

```
num_components = 12
pca = PCA(n_components=num_components)
X_train_pca = pca.fit_transform(X_train)
X_test_pca = pca.transform(X_test)
```

I. Méthode non supervisée (KNN : k nearest neighbors)

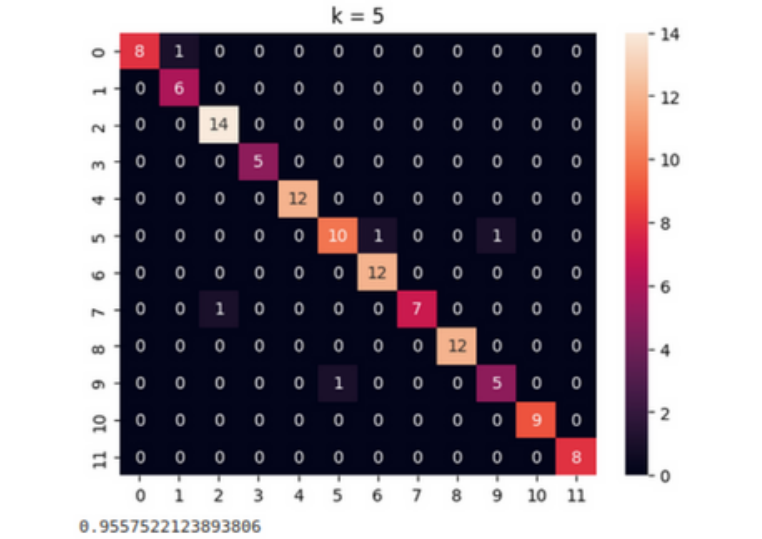
1. Impact de la valeur de k sur la précision :

Le graphique ci-dessous illustre comment la précision du modèle k-NN fluctue avec différentes valeurs de k:



Selon le graphique, la précision est la plus élevée à k=5. Elle diminue ensuite à k=7, puis remonte progressivement à k=9 et continue de croître à k=11. Cela indique que, pour cet ensemble de données, k=5 offre la meilleure précision

2. Matrices de confusion :



Interprétation: Le nombre de voisins optimal est passé de “3” sans ACP à “5” avec ACP. La matrice de confusion montre une forte performance : la plupart des prédictions sont correctes (diagonale principale), avec seulement 5 erreurs. Ceci donne un taux d'erreur de "0.04", alors que la précision est de "0.96" (le modèle est très précis).

3.Précision (Accuracy):

```
k = 1   |       accuracy = 0.9380530973451328
k = 3   |       accuracy = 0.9469026548672567
k = 5   |       accuracy = 0.9557522123893806
k = 7   |       accuracy = 0.9469026548672567
k = 9   |       accuracy = 0.9380530973451328
k = 11  |       accuracy = 0.9469026548672567
```

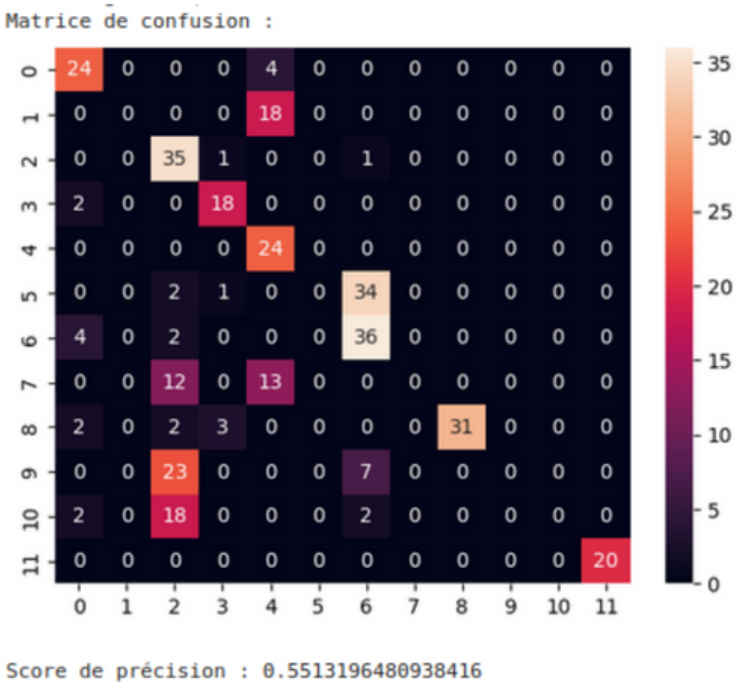
Interprétation: À partir des résultats obtenus pour différentes valeurs de k, nous constatons que notre modèle réalise une accuracy la plus élevée de 95.57% pour k = 5. Cela indique que pour cette valeur de k, notre modèle prédit correctement 95.57% des cas dans notre ensemble de test. Cette haute accuracy suggère que notre modèle est performant et fiable pour ce paramètre.

II. Méthode non supervisée (K-means)

D'abord, nous standardisons les données en utilisant la bibliothèque `**sklearn.preprocessing**`. Ensuite, notre objectif est de déterminer le nombre de clusters idéal pour notre tâche de clustering. Pour ce faire, nous faisons appel à la méthode du coude, qui illustre la variation de l'inertie en fonction du nombre de clusters à travers un graphique.

Sur le graphique, bien qu'il n'y ait pas de cassure évidente, nous identifions le point où la courbe commence à s'infléchir, soit "7". Ce chiffre représente le nombre de clusters optimal, où l'inertie est ni trop grande ni trop faible, indiquant une distance relativement uniforme entre les points à l'intérieur des clusters.

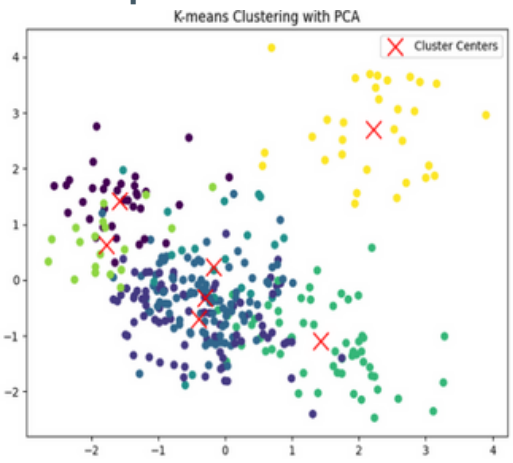
1. Matrices de confusion :

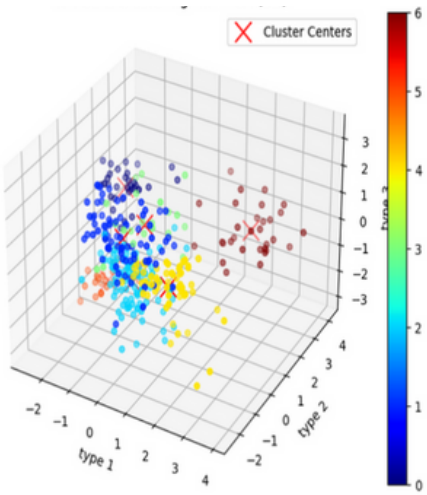


Interprétation: on remarque que les éléments non-nul hors la diagonal sont plus élévé , on a un taux d'erreur de classification de "0.45" mais on est quand même satisfait car les élément qui sont dans la diagonal sont plus élevé et ce qui veut dire que notre modèle arrive à prédire les "vrai positifs" et les "vrais négatifs" mieux que les "faux positifs " ou les "faux négatifs" .
on peut remarquer aussi :

- les 1 sont totalement confondus par des 4
- les 5 sont trop confondus avec des 6 et sont mal classés
- les 7 sont divisé en 2 clusters de 2 et 4
- les 8 sont un peu reparties mais la majorité est bien classée
- les 9 sont confondus avec dès 2 (23 entités) et 7 entités avec 6
- les 10 sont trop confondus avec 2

2.Représentation des données 2D / 3D :





Graphique 1 (2D) :

- Représentation en deux dimensions des clusters formés avec les centres indiqués par des croix rouges.

Graphique 2 (3D) :

- Vue tridimensionnelle offrant une perspective plus détaillée des clusters.
- La barre de couleur à droite indique la profondeur des points.

- Ces visualisations mettent en évidence la capacité de K-means combiné à la PCA à segmenter clairement les données.

Conclusion :

Nous avons exploré des méthodes supervisées et non supervisées, combinées à l'ACP. Les approches supervisées ont montré une précision élevée, tandis que le clustering non supervisé a révélé des structures de données intrinsèques. L'ACP a efficacement réduit la dimensionnalité. L'ensemble des techniques utilisées offre une perspective complète de notre jeu de données.