Unmanned Aerial Vehicle (UAV) Intrusion Detection Datasets

Jeyaratnam Joshua

Lévy Daniel

Datascience & IA, ESILV 2023

Enjeux du projet

Depuis 2014, on constate une multiplication de survols de zones sensibles par drone.

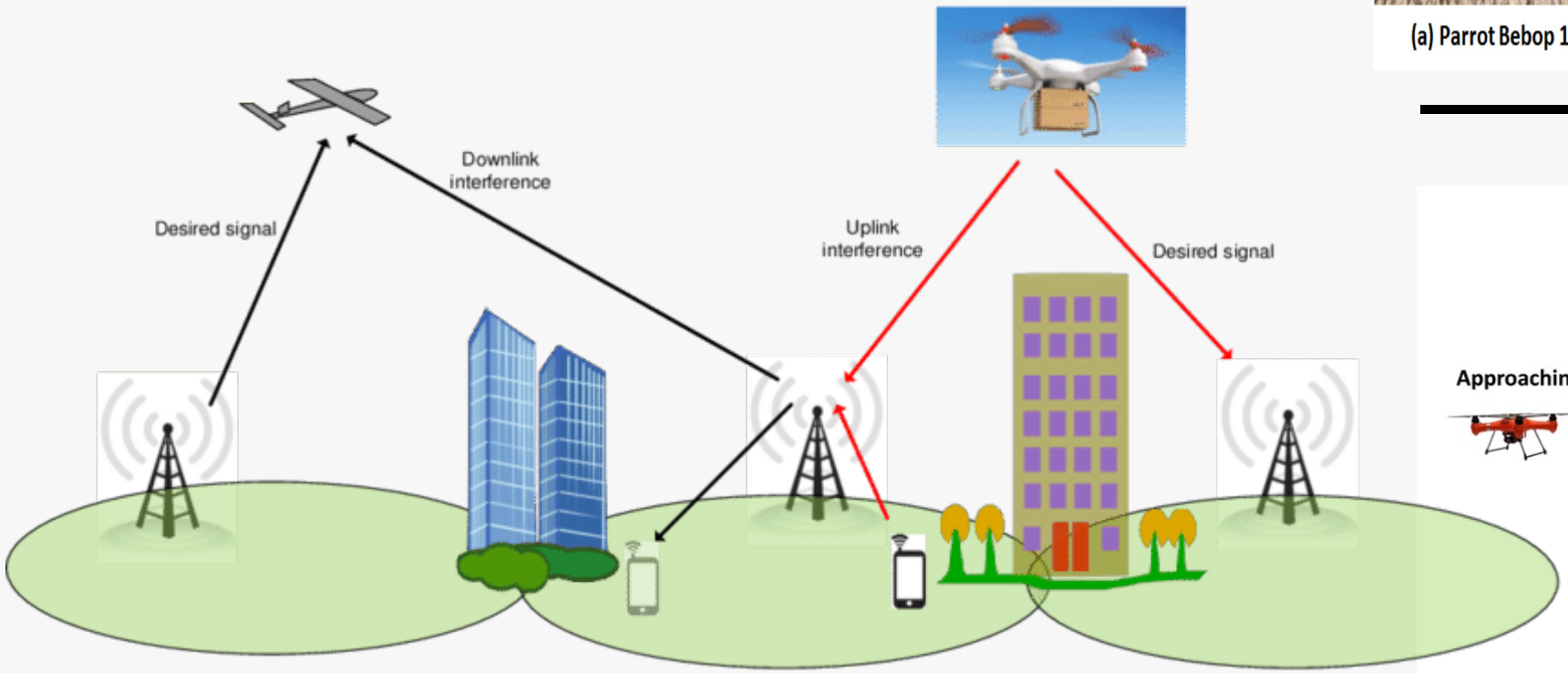
En France en 2018, 135 infractions impliquant des drones ont été constatées dont 35 survols de zones interdites, auxquels s’ajoutent les nombreux survols de la prison de Fresnes.

Des risques potentiels pour la sécurité publique ou la vie privée.

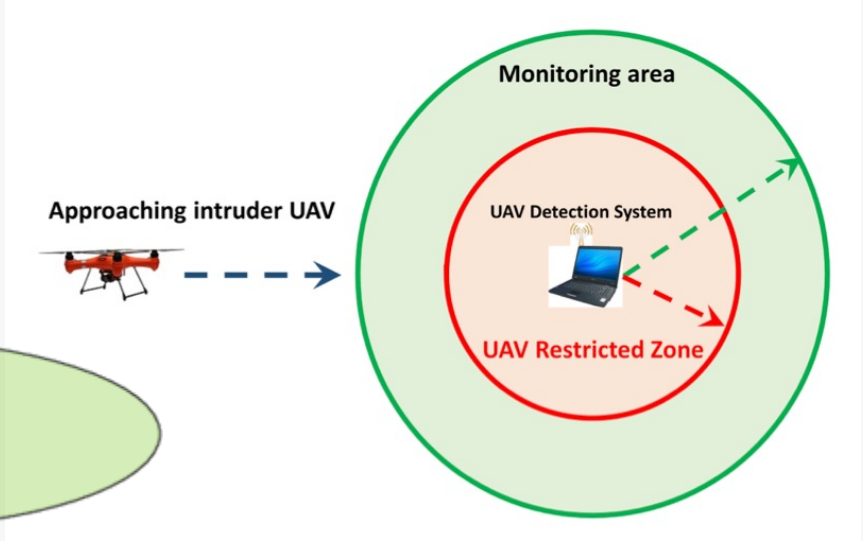
Les méthodes de détection physique existantes (telles que le radar, la vision et le son) peuvent devenir inefficaces.

Il devient urgent de développer de nouvelles techniques de détection et d'identification des drones envahisseurs.

Détection d’UAV



(a) Parrot Bebop 1 (b) DBPower UDI (c) DJI Spark



Features

56 features

17800 lignes

uplink size	downlink size	both_links size	uplink_interval	downlink interval	both_links interval	Label
0.01363306	0.009906635	0.012473	0.01319514	0.2069389653	0.0136330612	1
0.17687626	0.328221843	0.0509895	0.072376825	2.7467961576	7.7993522431	0
			...	...		

Features

- uplink size
- downlink size
- both_links size
- uplink_interval
- downlink interval
- both_linksinterval

Feature ID:Name	Description
V_1: mean	$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x(i)$
V_2: median	The higher half value of a data sample.
V_3: MAD[1]	$MAD = median(\lvert x(i) - median(x)\rvert)$
V_4: STD[1]	$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x(i) - mean(x))^2}$
V_5: Skewness	$\gamma = \frac{1}{N}\sum_{i=1}^{N}(x(i) - mean(x)/\sigma)^3$
V_6: Kurtosis	$\beta = \frac{1}{N}\sum_{i=1}^{N}(x(i) - mean(x)/\sigma)^4$
V_7: MAX	$H = (Max(x(i))\lvert_{i=1\ldots N})$
V_8: MIN	$L = (Min(x(i))\lvert_{i=1\ldots N})$
V_9: Mean Square (ms)	$MS = \frac{1}{N}\sum_{i=1}^{N}(x(i))^2$

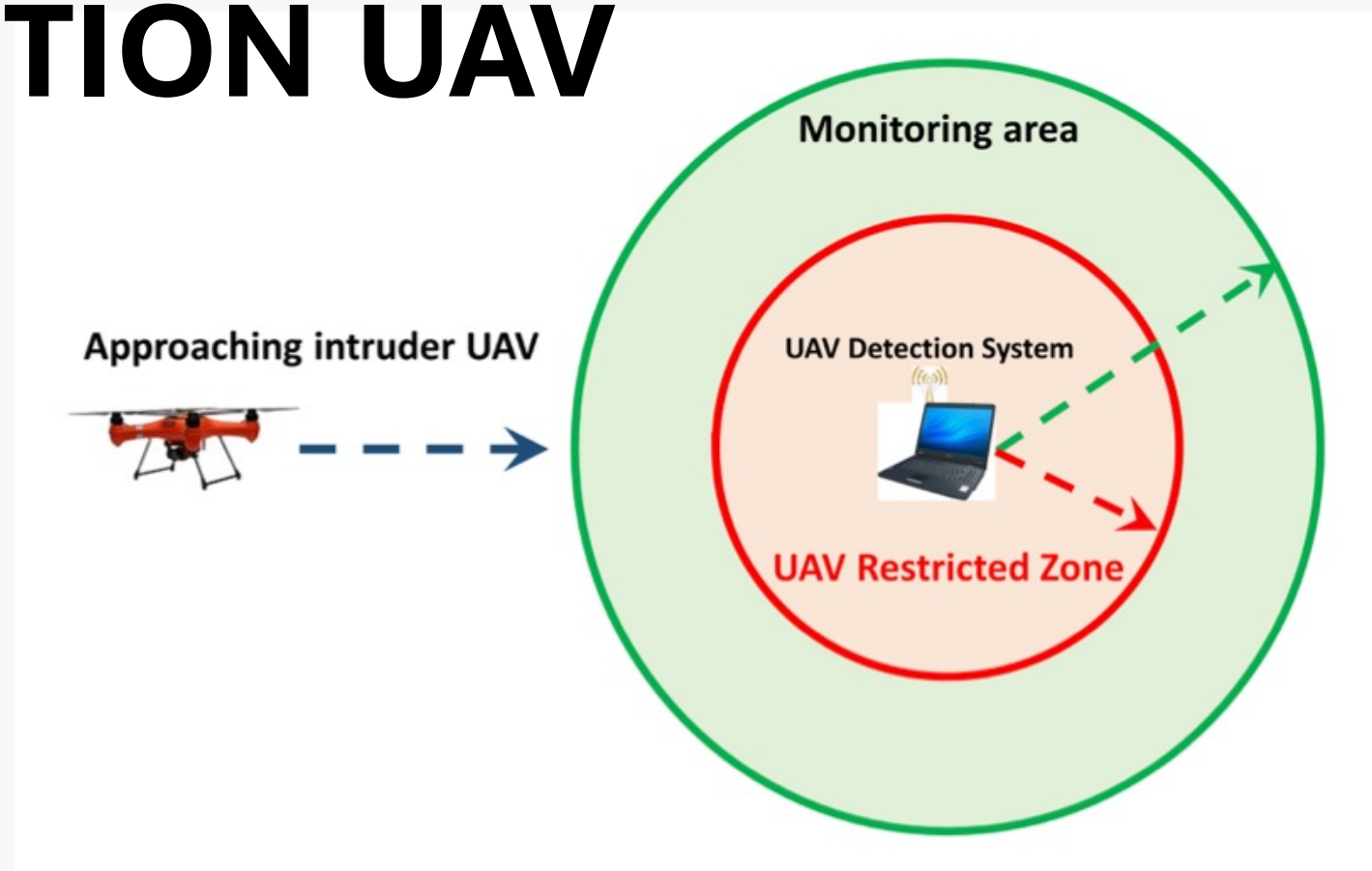
1. uplink_size_mean
2. uplink_size_median
3. uplink_size_MAD
4. uplink_size_STD
5. uplink_size_Skewness
6. uplink_size_Kurtosis
7. uplink_size_MAX
8. uplink_size_MIN
9. uplink_size_MeanSquare
10. downlink_size_mean
11. downlink_size_median
12. downlink_size_MAD
13. downlink_size_STD
14. downlink_size_Skewness
15. downlink_size_Kurtosis
16. downlink_size_MAX
17. downlink_size_MIN
18. downlink_size_MeanSquare
19. both_links_size_mean
20. both_links_size_median
21. both_links_size_MAD
22. both_links_size_STD
23. both_links_size_Skewness
24. both_links_size_Kurtosis
25. both_links_size_MAX
26. both_links_size_MIN
27. both_links_size_MeanSquare
28. uplink_interval_mean
29. uplink_interval_median
30. uplink_interval_MAD
31. uplink_interval_STD
32. uplink_interval_Skewness
33. uplink_interval_Kurtosis
34. uplink_interval_MAX
35. uplink_interval_MIN
36. uplink_interval_MeanSquare
37. downlink_interval_mean
38. downlink_interval_median
39. downlink_interval_MAD
40. downlink_interval_STD
41. downlink_interval_Skewness
42. downlink_interval_Kurtosis
43. downlink_interval_MAX
44. downlink_interval_MIN
45. downlink_interval_MeanSquare
46. both_links_interval_mean
47. both_links_interval_median
48. both_links_interval_MAD
49. both_links_interval_STD
50. both_links_interval_Skewness
51. both_links_interval_Kurtosis
52. both_links_interval_MAX
53. both_links_interval_MIN
54. both_links_interval_MeanSquare
55. label

Problématiques

Reconnaitre si le flux est celui d’un UAV ou non.

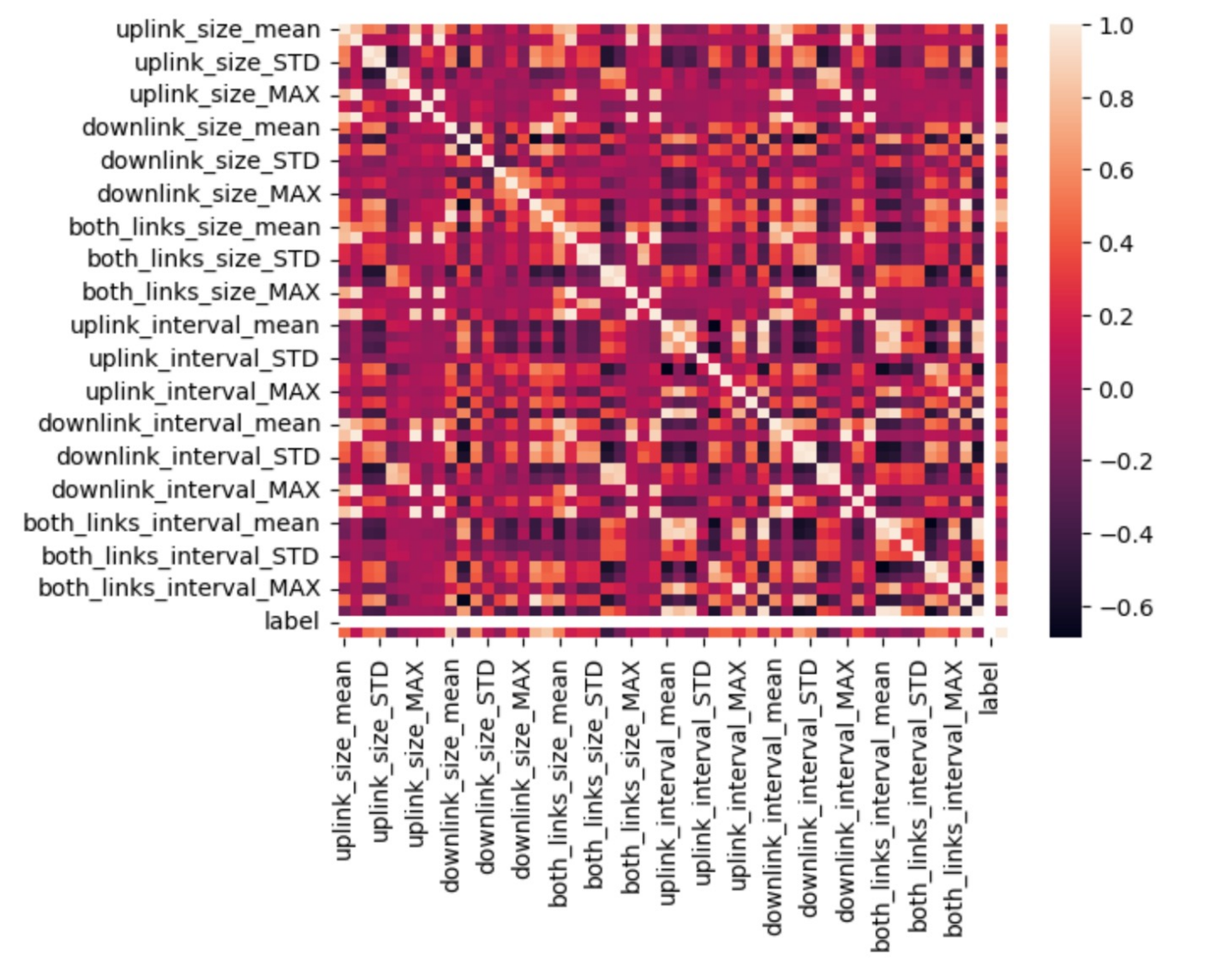
Reconnaître le modèle dont provient le flux WIFI.

Problématique n°1: DETECTION UAV

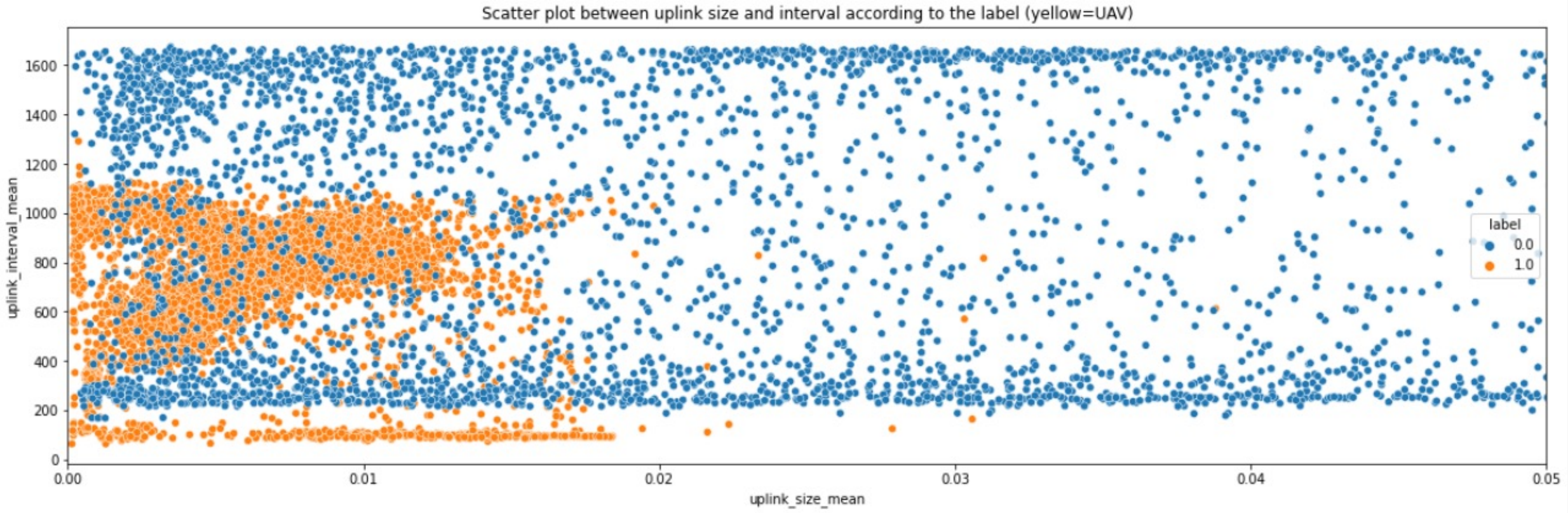


Matrice de corrélation

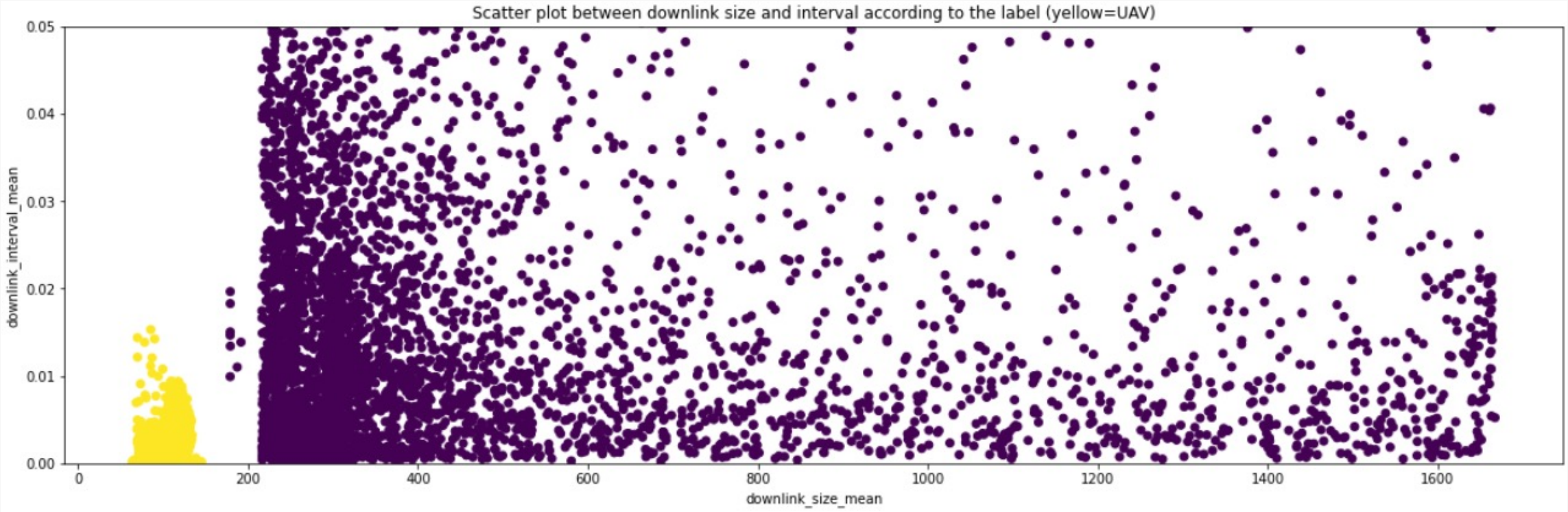
label	1.000000
both_links_interval_MIN	0.976437
downlink_interval_MIN	0.849652
downlink_size_MAX	0.742546
downlink_size_MIN	0.730564
downlink_size_MeanSquare	0.665829
downlink_size_mean	0.644901
downlink_size_median	0.598818
uplink_interval_MIN	0.566412
downlink_size_MAD	0.503821
both_links_interval_MAX	0.439436
uplink_interval_Kurtosis	0.379975
both_links_interval_Skewness	0.344019
uplink_size_Skewness	0.343666
downlink_interval_Kurtosis	0.324745
uplink_interval_STD	0.305793
uplink_size_Kurtosis	0.304020
both_links_interval_Kurtosis	0.298865



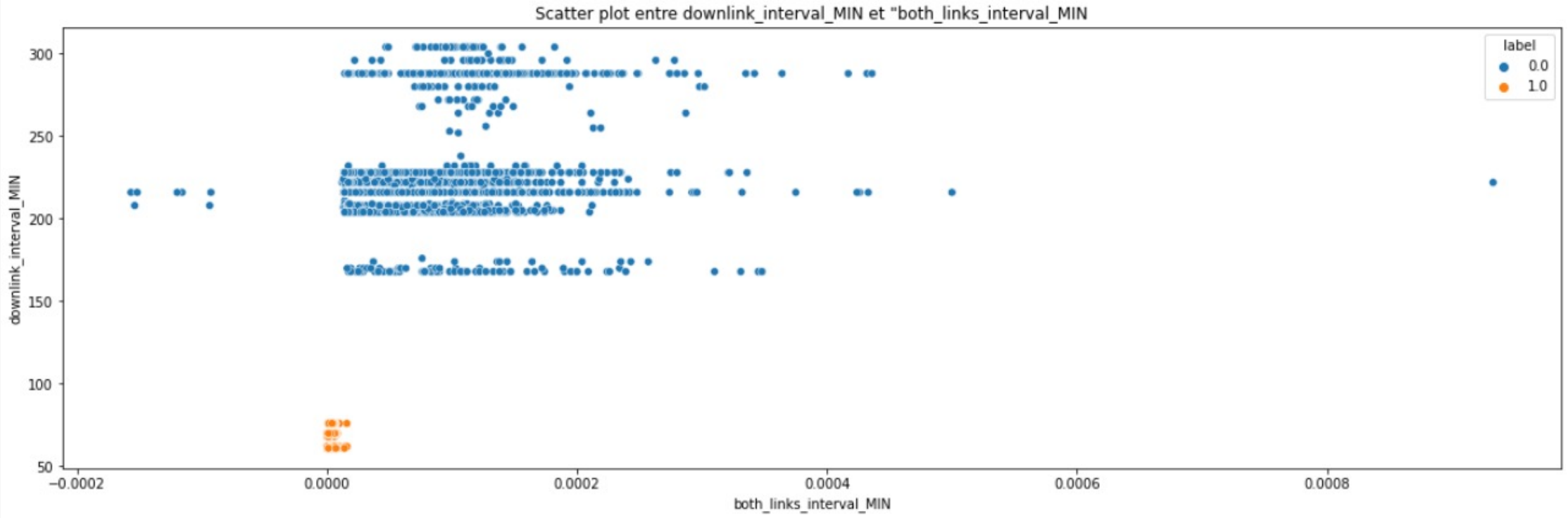
Visualisation



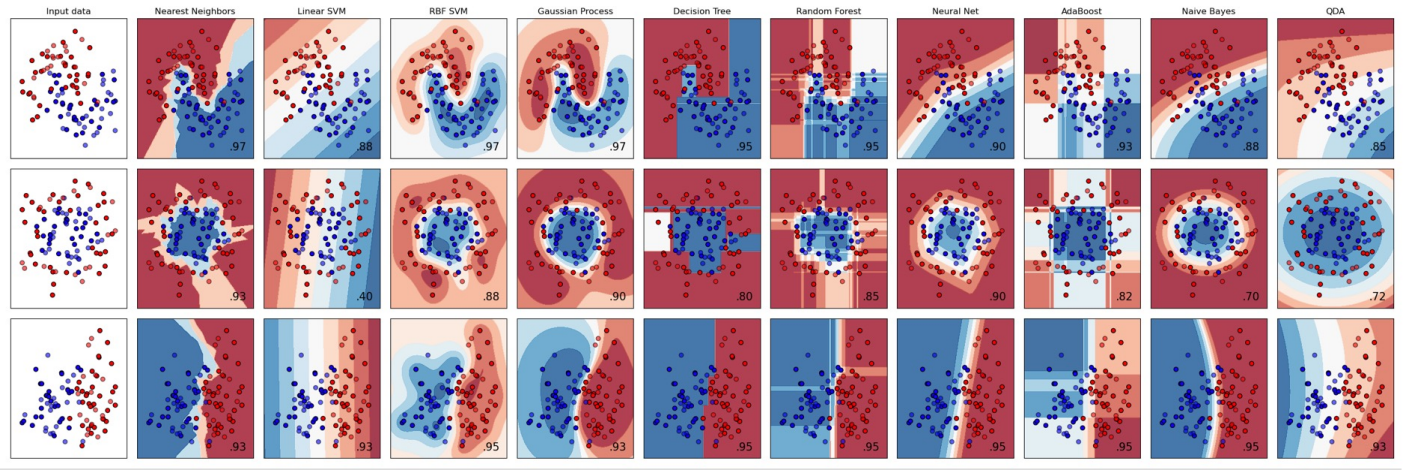
Visualisation



Visualisation



Techniques de classification



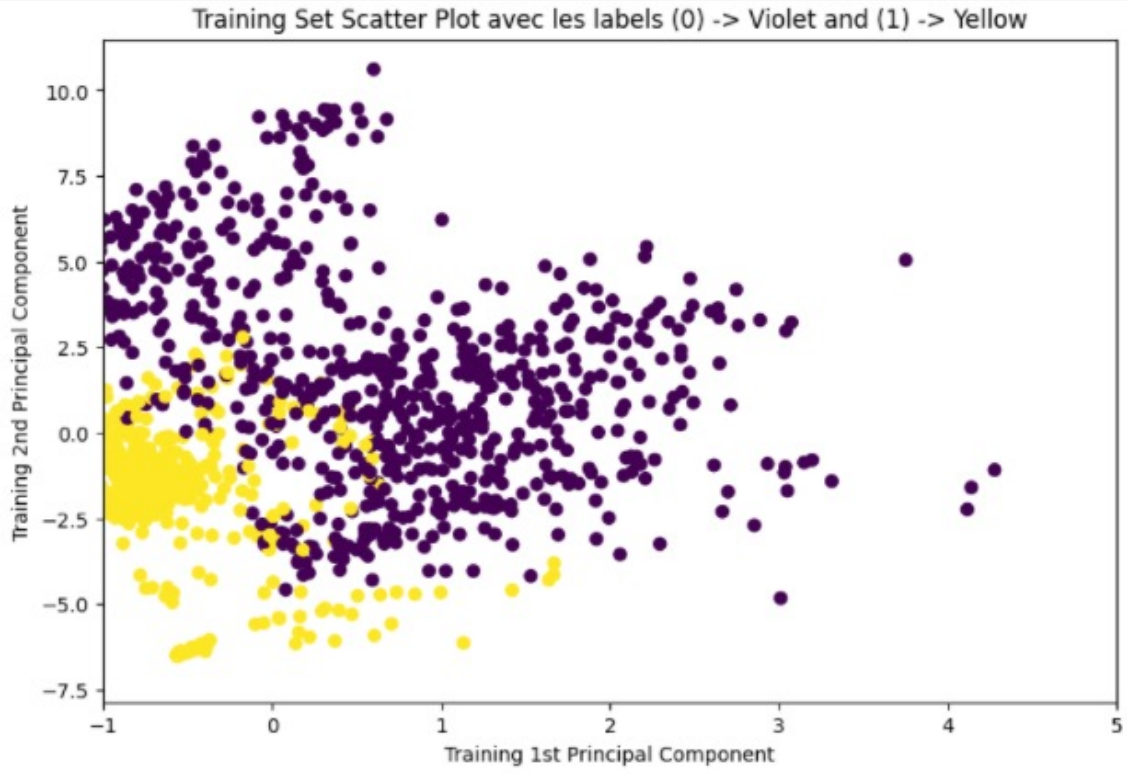
1) Travailler sur les données

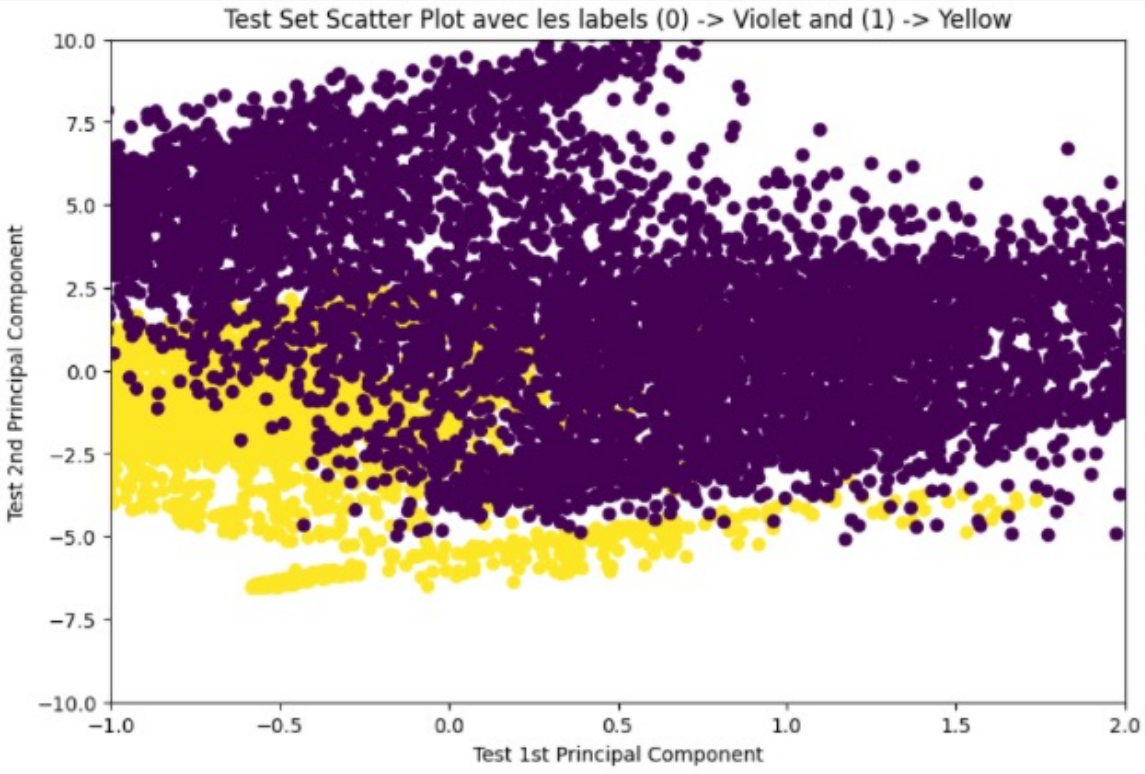
- Scale the data
- PCA

2) Classification

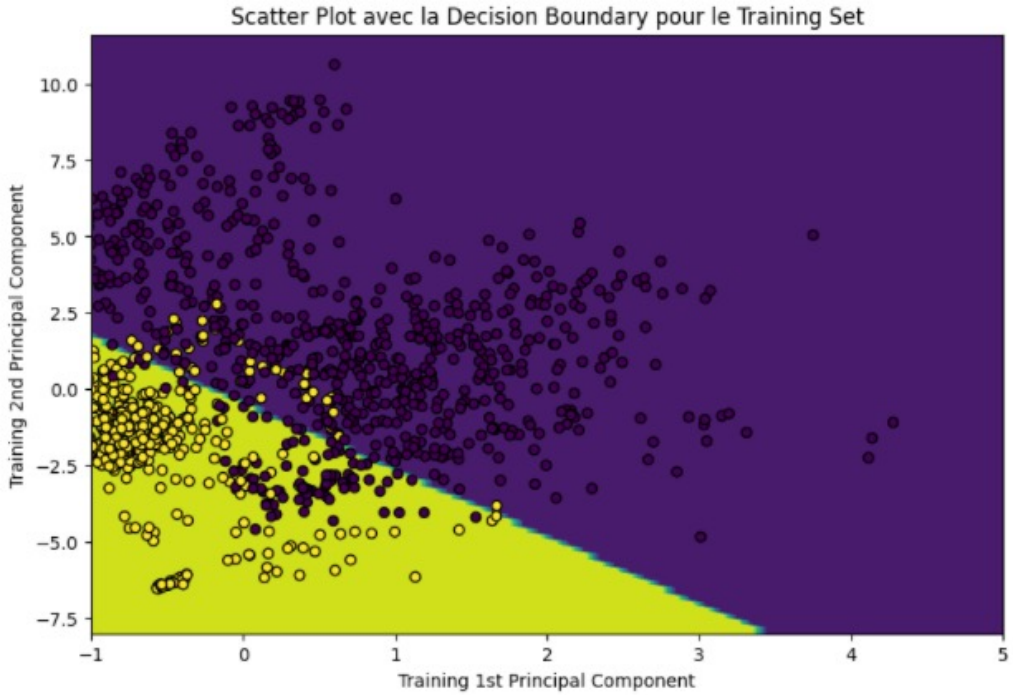
- KNN
- SVC
- Logistic Regression
- Decision tree Classifier

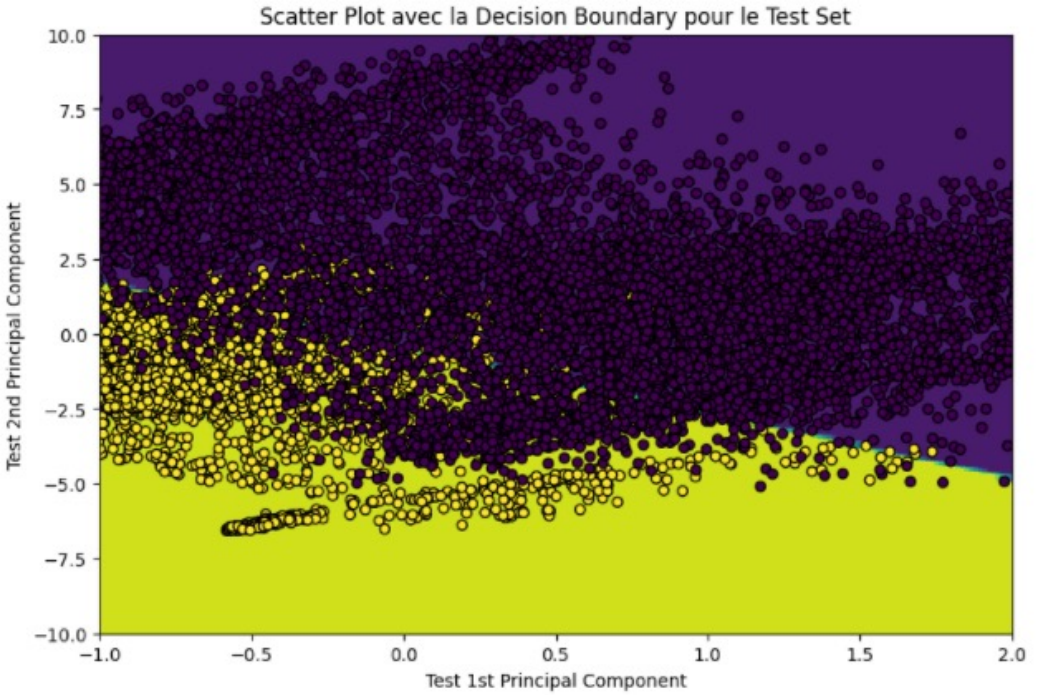
Visualisation après PCA





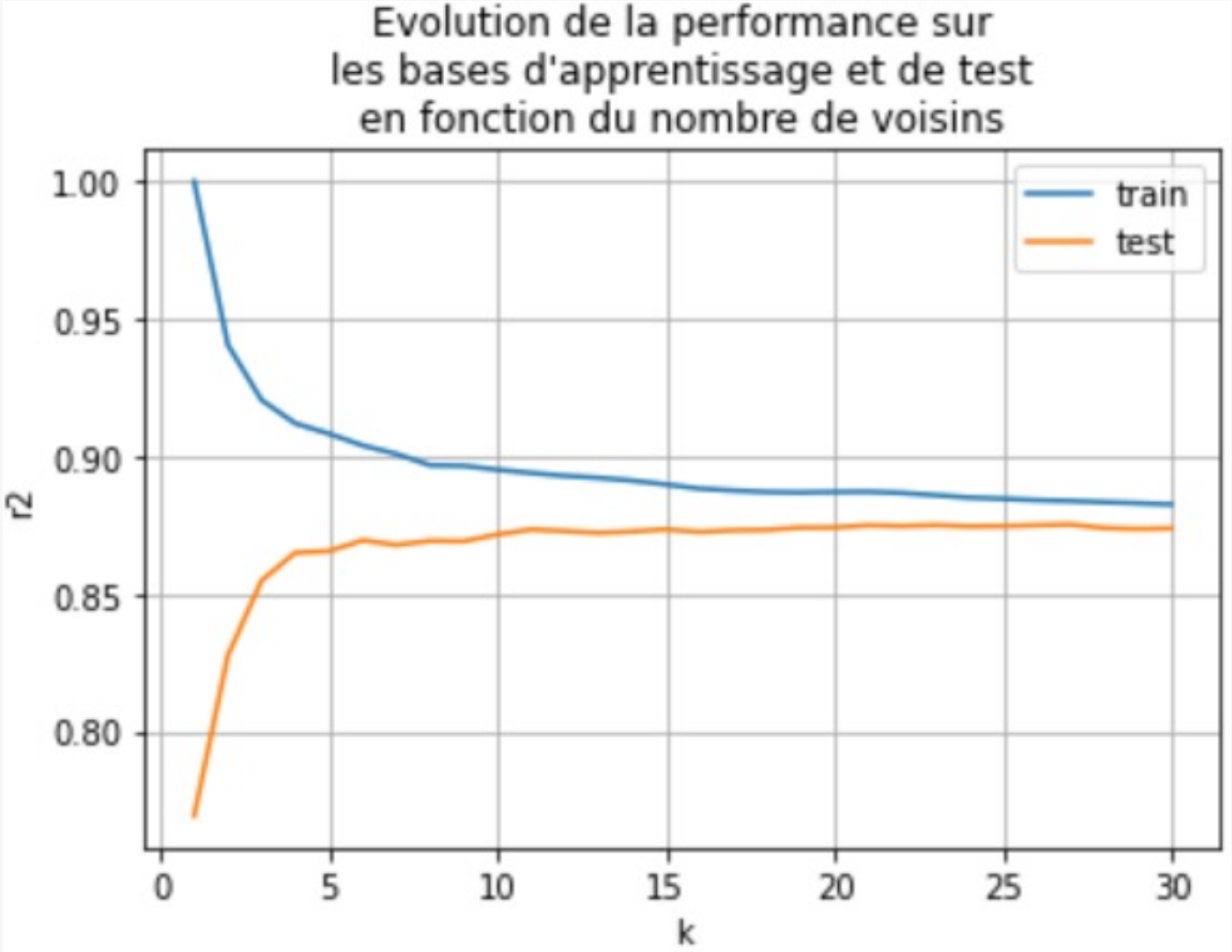
Logistic regression et GridSearch






```
Accuracy du Model:  0.9234723015419 76
```

Performance des paramètres



Résultats :

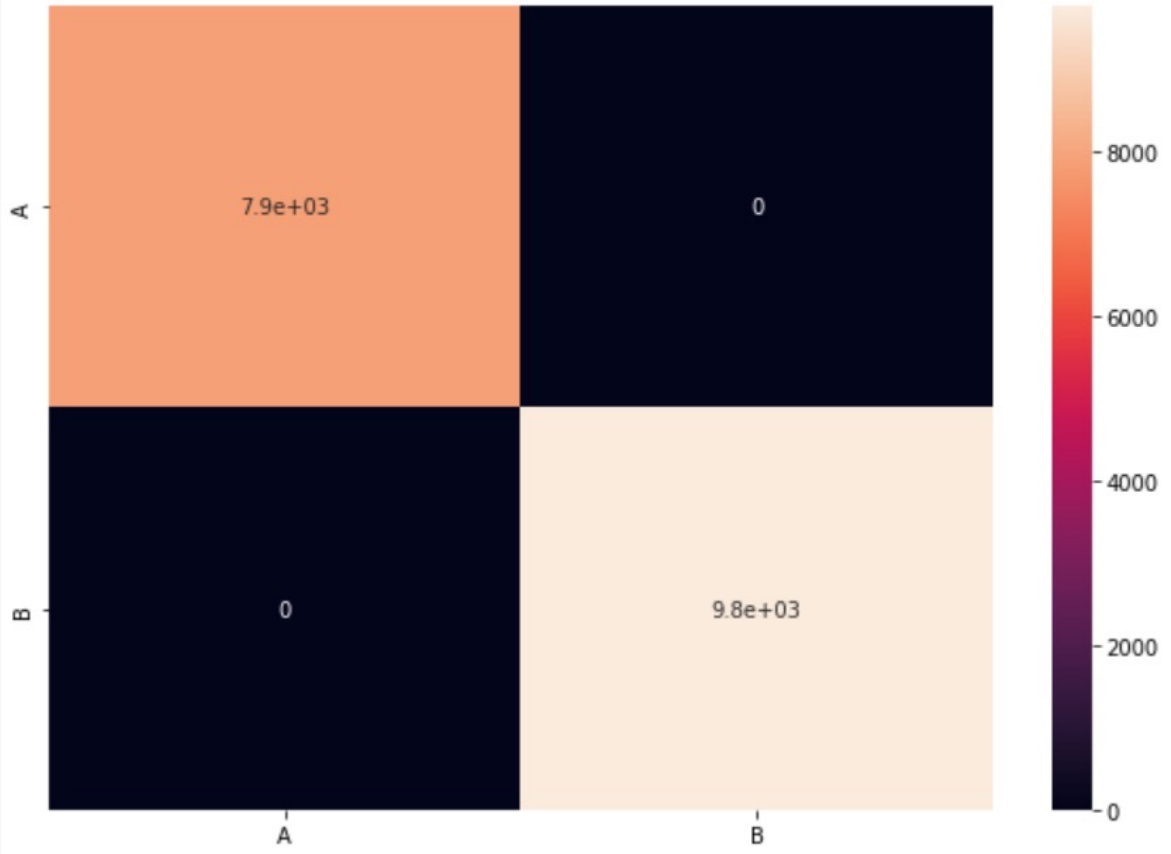
SVC

KNN

Features utilisées:
both_links_interval_MIN
downlink_interval_MIN

	precision	recall	f1-score	support
0.0	1.00	1.00	1.00	7869
1.0	1.00	1.00	1.00	9760
accuracy			1.00	17629

Matrice de confusion



Conclusion PB1

Conclusion:

- Pas besoin de centré réduire ni de faire une PCA.
- 100 % d'accuracy mais si le constructeur décide de changer la taille des packages alors le modèle perdra de sa précision.
- Le modèle SVC est le plus performant dans ce cas.

Pour aller plus loin:

- Effectuer l'expérience nous même avec un drone et tester l’IA.

Problématique n°2
QUEL MODELE D'UAV ?

(a) Parrot Bebop 1

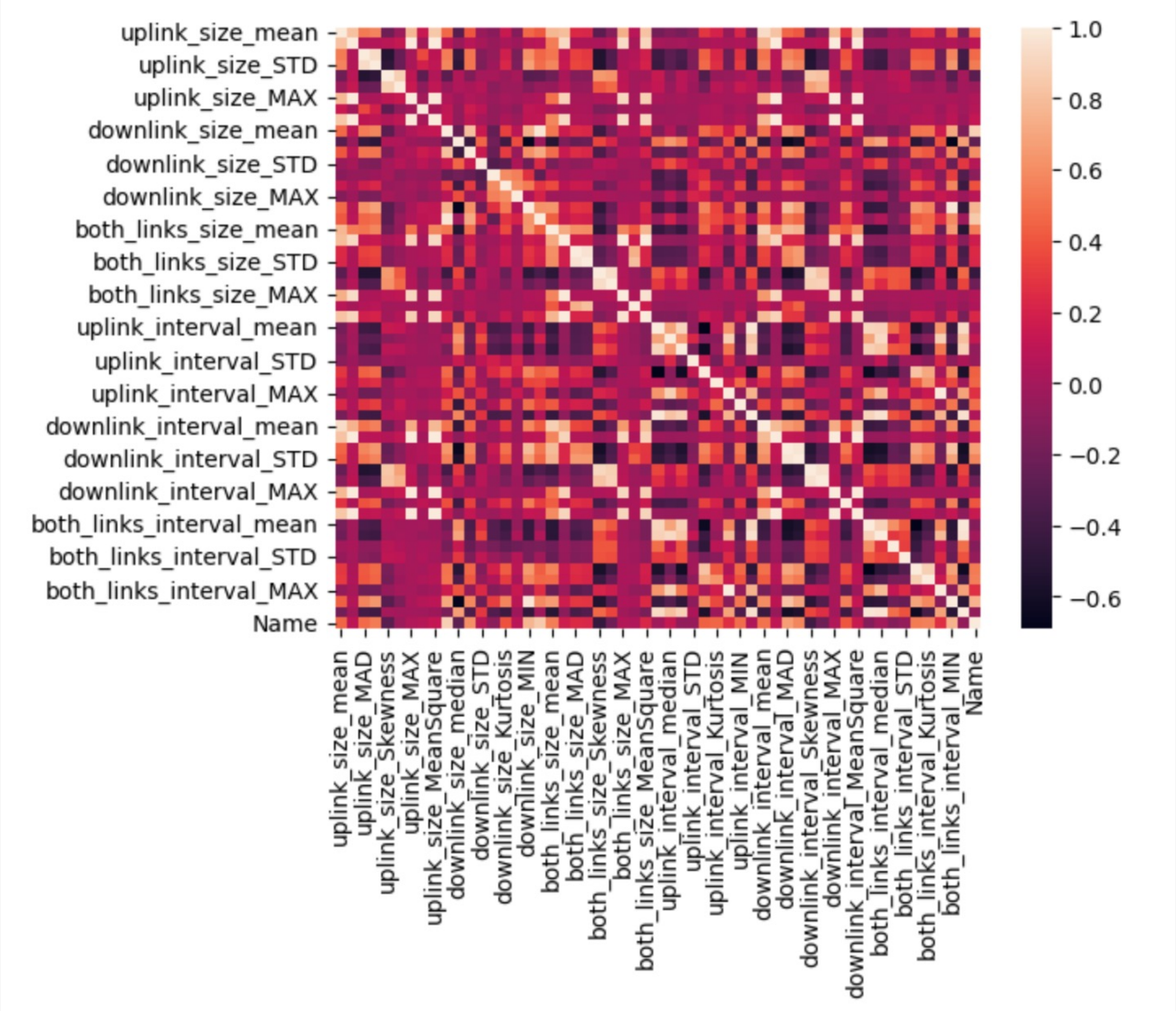
(b) DBPower UDI

(c) DJI Spark

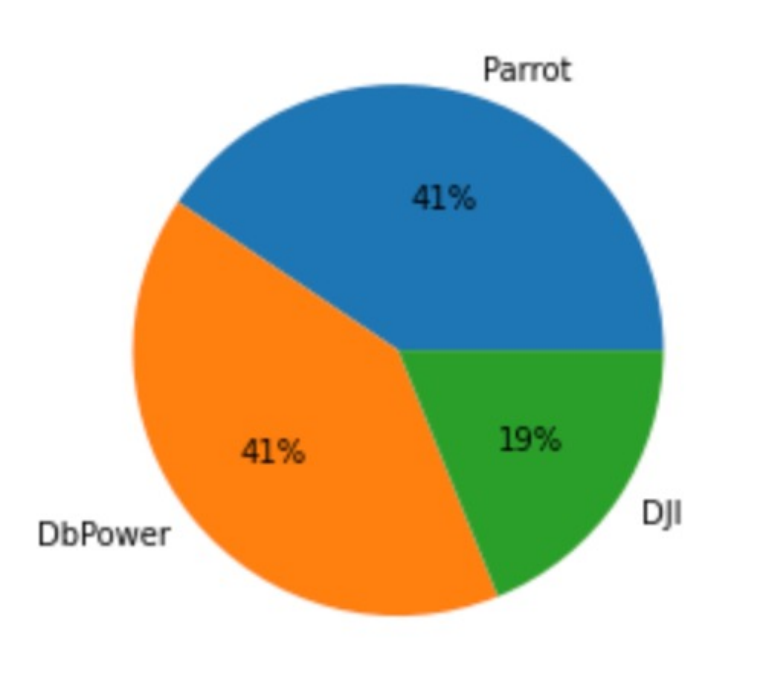
Traffic mode	Parrot Bebop	DBPower UDI	DJI Spark
Bidirectional-flow mode	Dataset 1	Dataset 2	Dataset 3

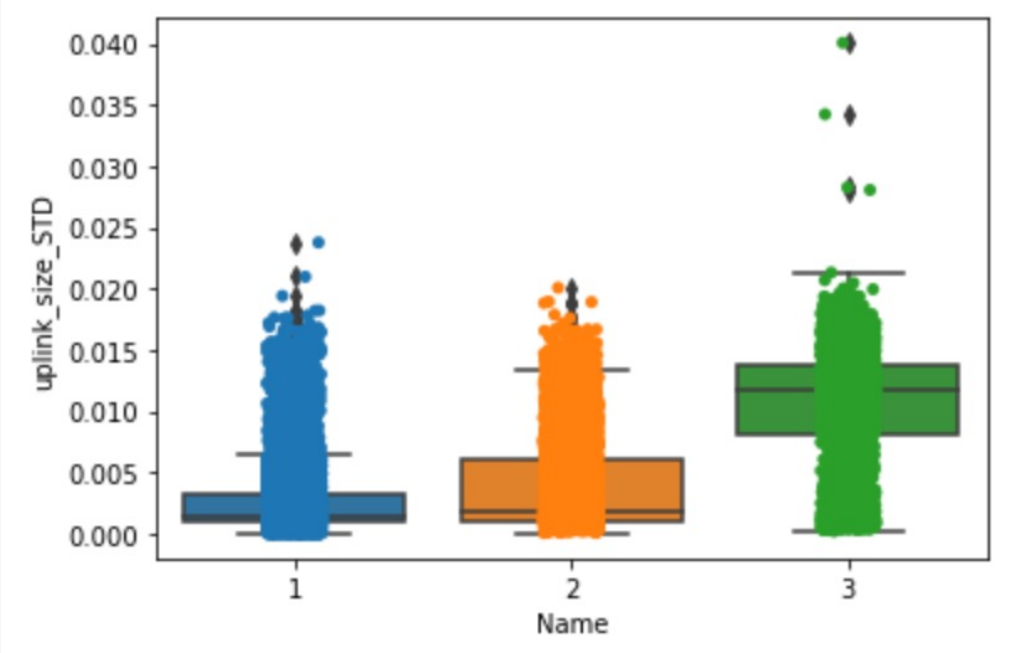
Matrice de corrélation

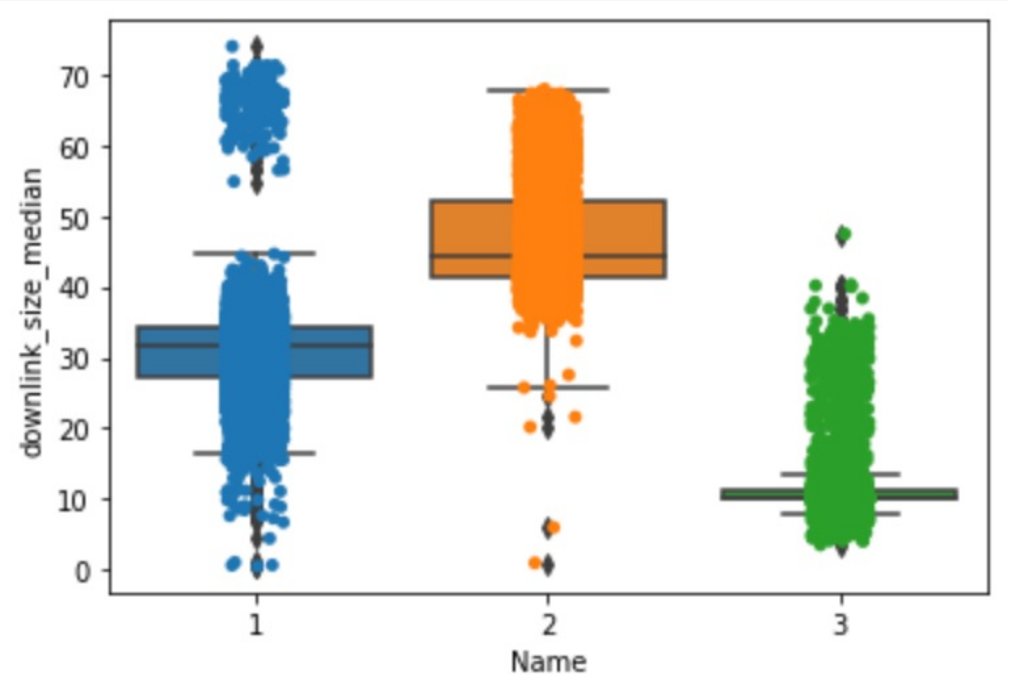
Name	1.000000
downlink_size_mean	0.881025
downlink_size_MeanSquare	0.853912
downlink_size_MIN	0.764552
both_links_interval_MIN	0.754825
downlink_interval_MAD	0.664044
downlink_size_MAD	0.608923
downlink_interval_STD	0.571360
both_links_interval_Kurtosis	0.554614
both_links_size_mean	0.546708
uplink_size_STD	0.537764
both_links_interval_Skewness	0.533526
downlink_interval_mean	0.509253
uplink_interval_MIN	0.483993
uplink_size_MAD	0.472108
uplink_size_mean	0.452570
uplink_interval_Kurtosis	0.452144
uplink_interval_Skewness	0.443406
both_links_size_Skewness	0.440316
downlink_interval_MIN	0.404422



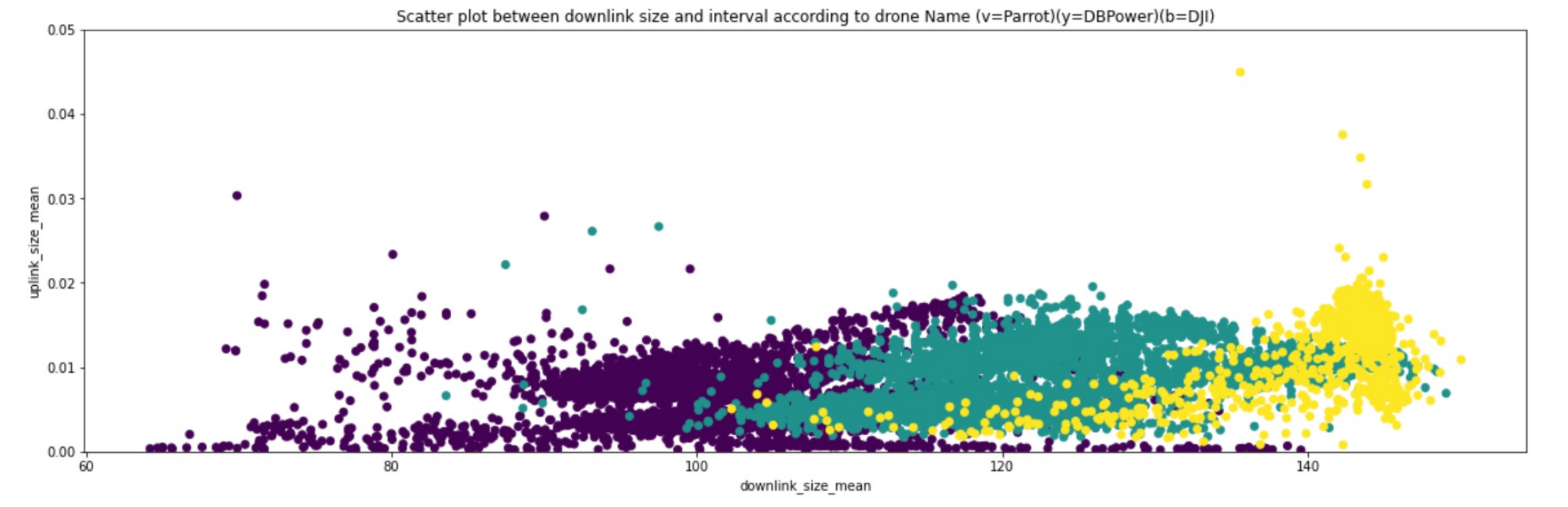
Visualisation







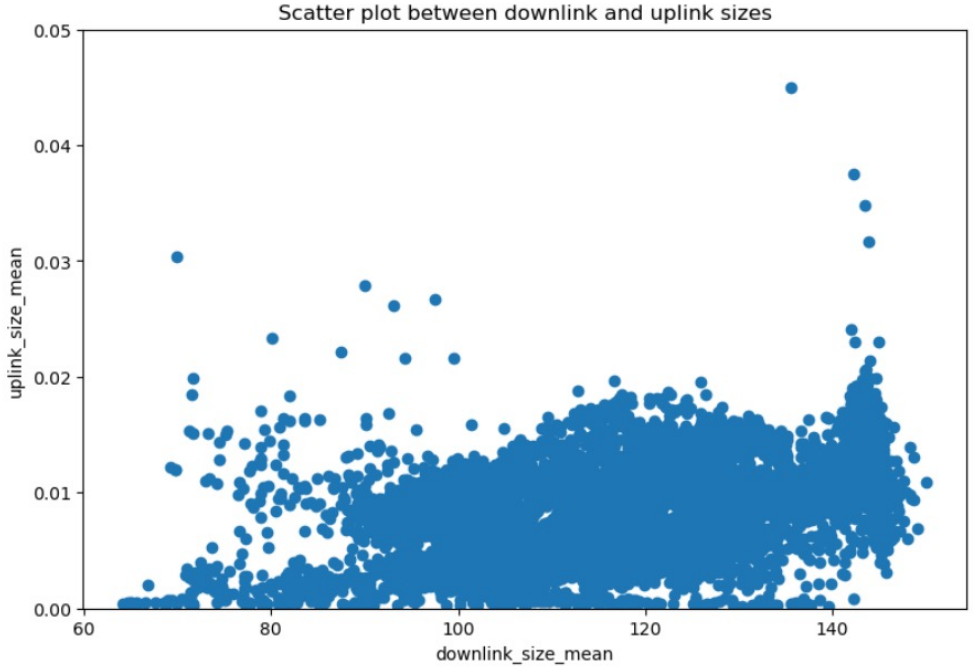
Visualisation

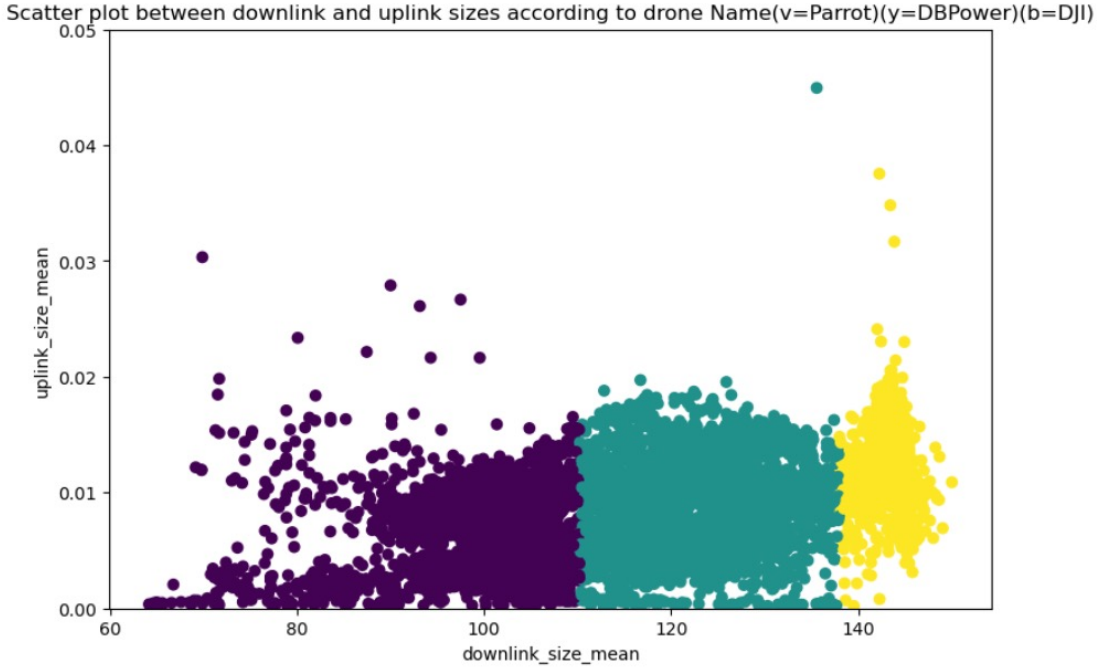


Model

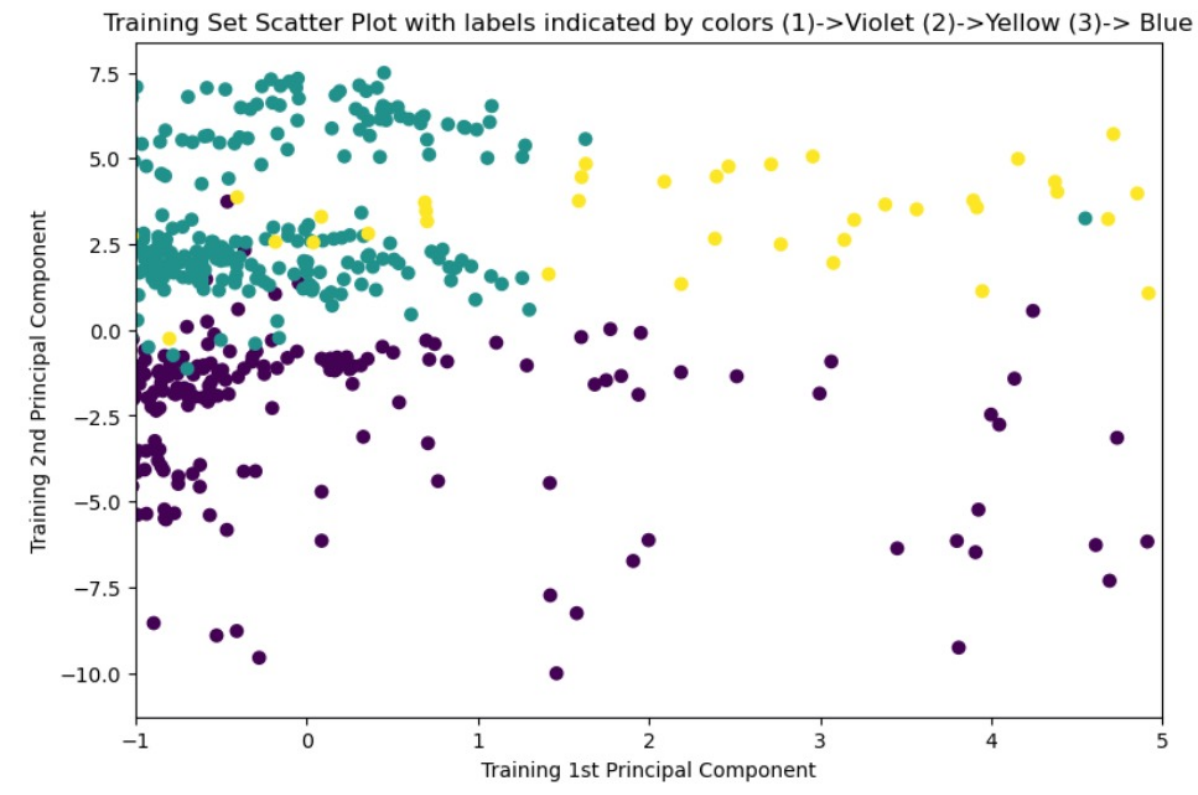
	precision	recall	f1-score	support
1	0.91	0.89	0.90	5377
2	0.85	0.86	0.86	5377
3	0.90	0.92	0.91	2500
accuracy			0.88	13254
macro avg	0.89	0.89	0.89	13254
weighted avg	0.88	0.88	0.88	13254

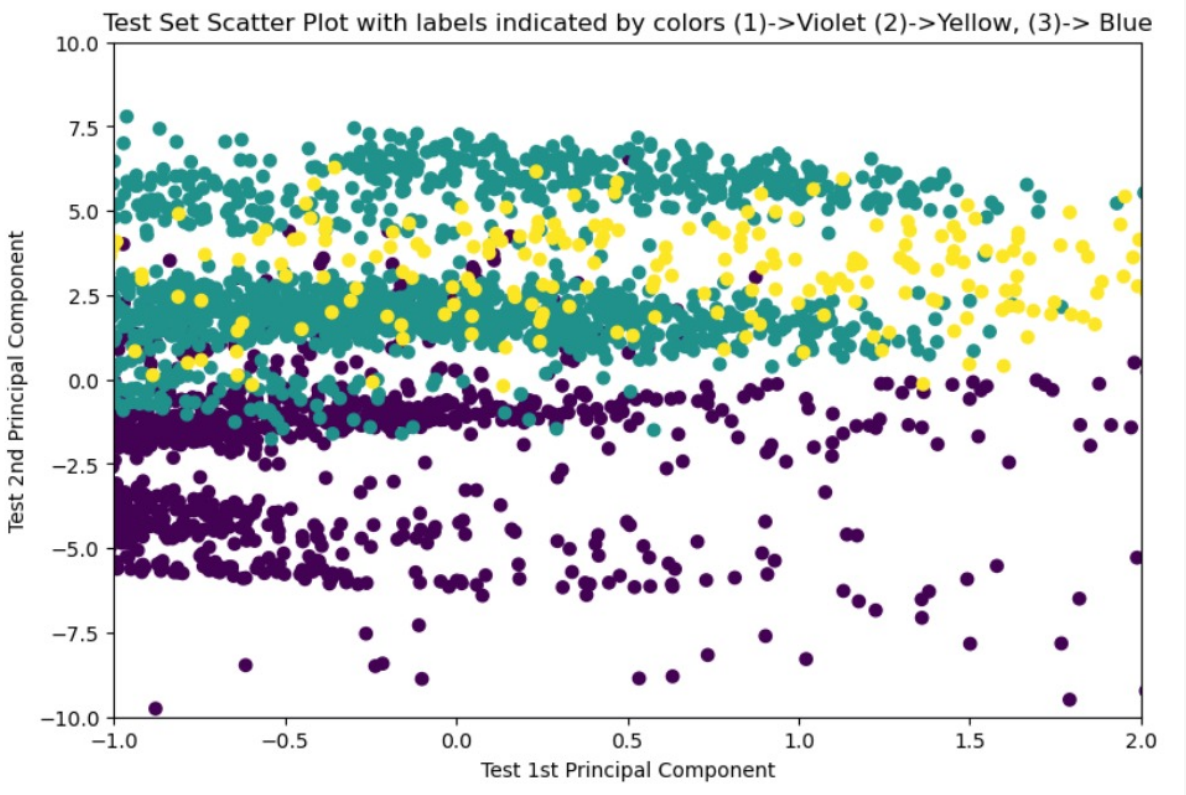
Features utilisées:
downlink size mean
Uplink siize mean





PCA

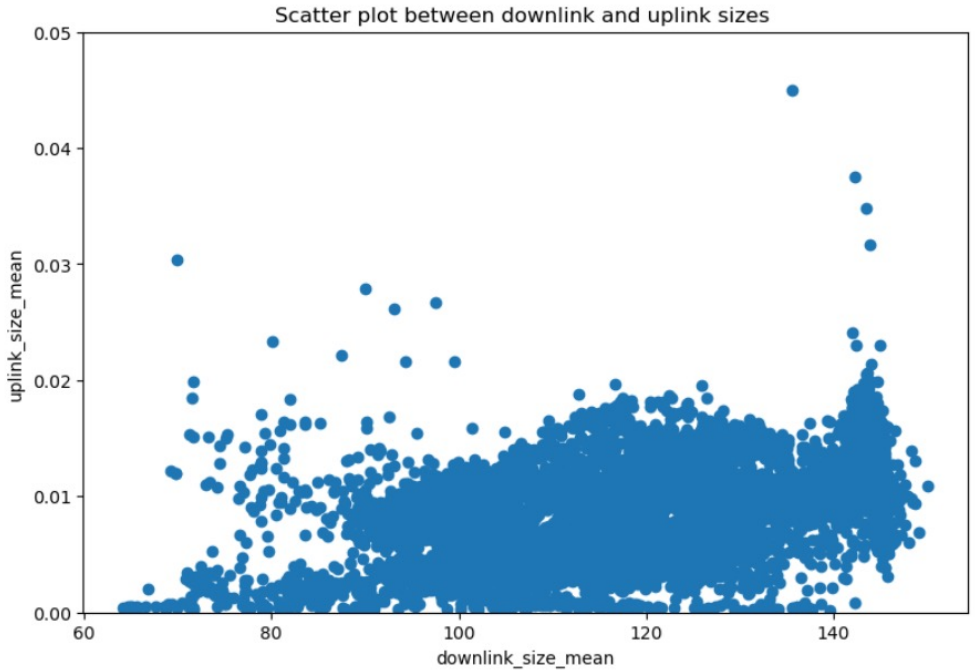

Training Set Scatter Plot with labels indicated by colors (1)->Violet (2)->Yellow (3)-> Blue
Training 2nd Principal Component
Training 1st Principal Component
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−1
0
1
2
3
4
5

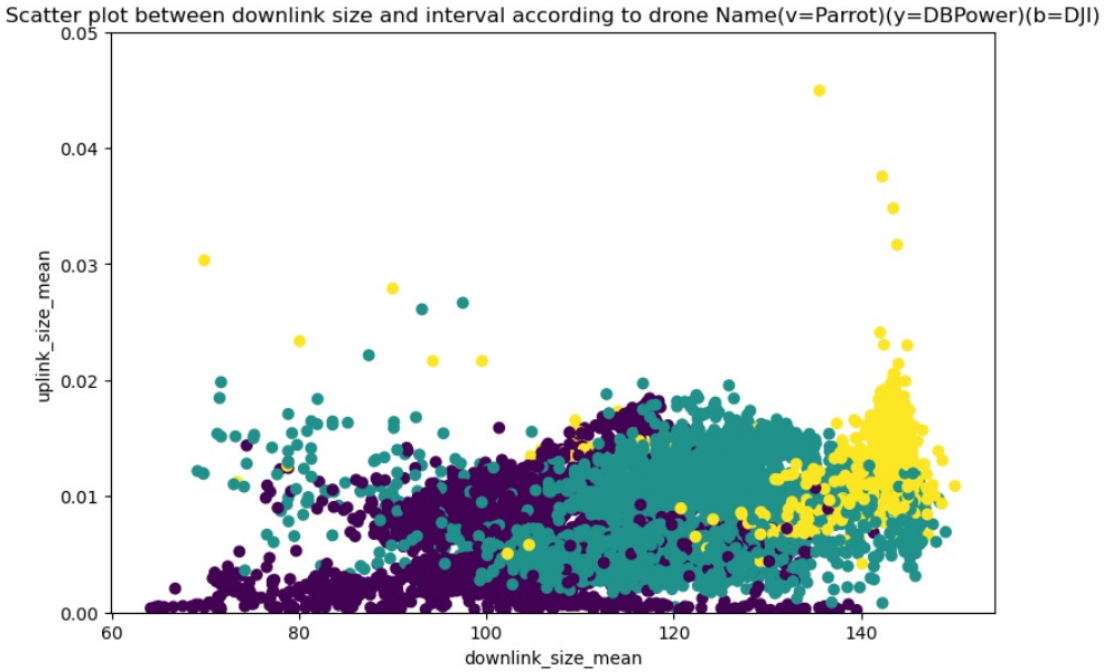

Test Set Scatter Plot with labels indicated by colors (1)->Violet (2)->Yellow, (3)-> Blue
Test 2nd Principal Component
Test 1st Principal Component
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

Model

	precision	recall	f1-score	support
1	0.93	0.95	0.94	5377
2	0.91	0.94	0.92	5377
3	0.99	0.89	0.94	2500
accuracy			0.93	13254
macro avg	0.94	0.92	0.93	13254
weighted avg	0.93	0.93	0.93	13254

Features utilisées:
1st 2nd Principal Component

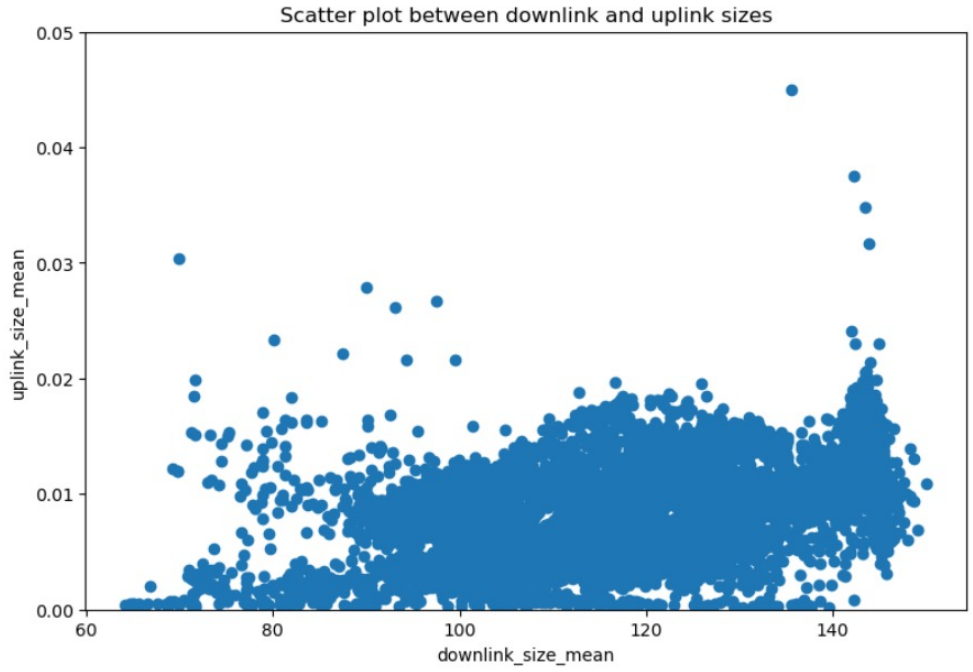


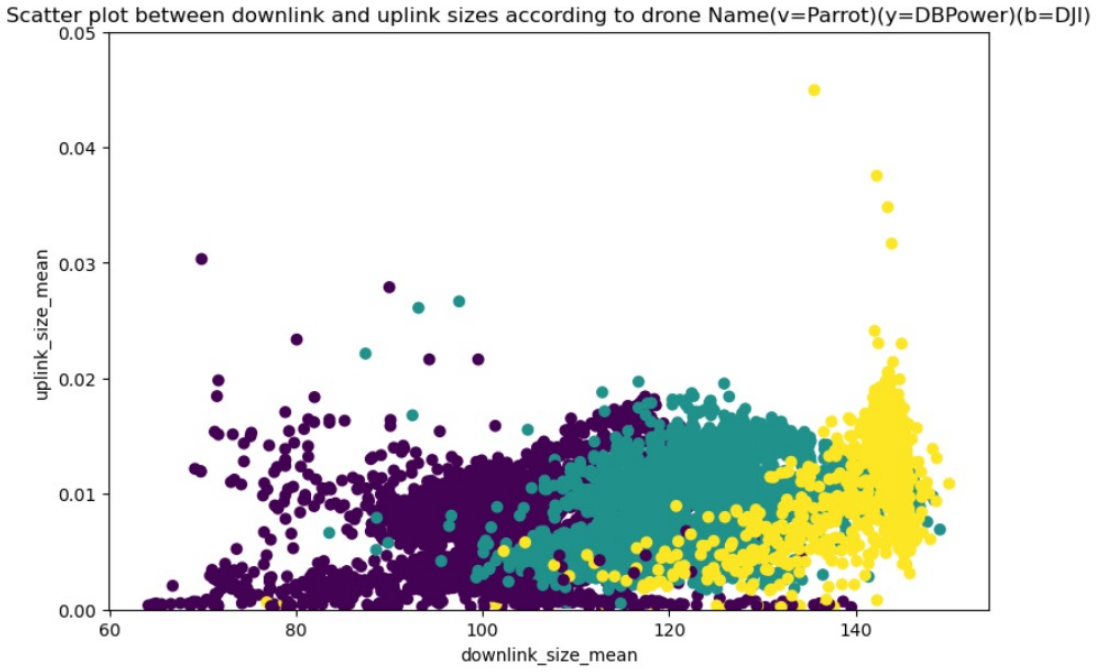


Model

	precision	recall	f1-score	support
1	1.00	1.00	1.00	5377
2	1.00	1.00	1.00	5377
3	0.99	0.99	0.99	2500
accuracy			1.00	13254
macro avg	1.00	1.00	1.00	13254
weighted avg	1.00	1.00	1.00	13254

Features utilisées:
TOUTES





Conclusion PB2

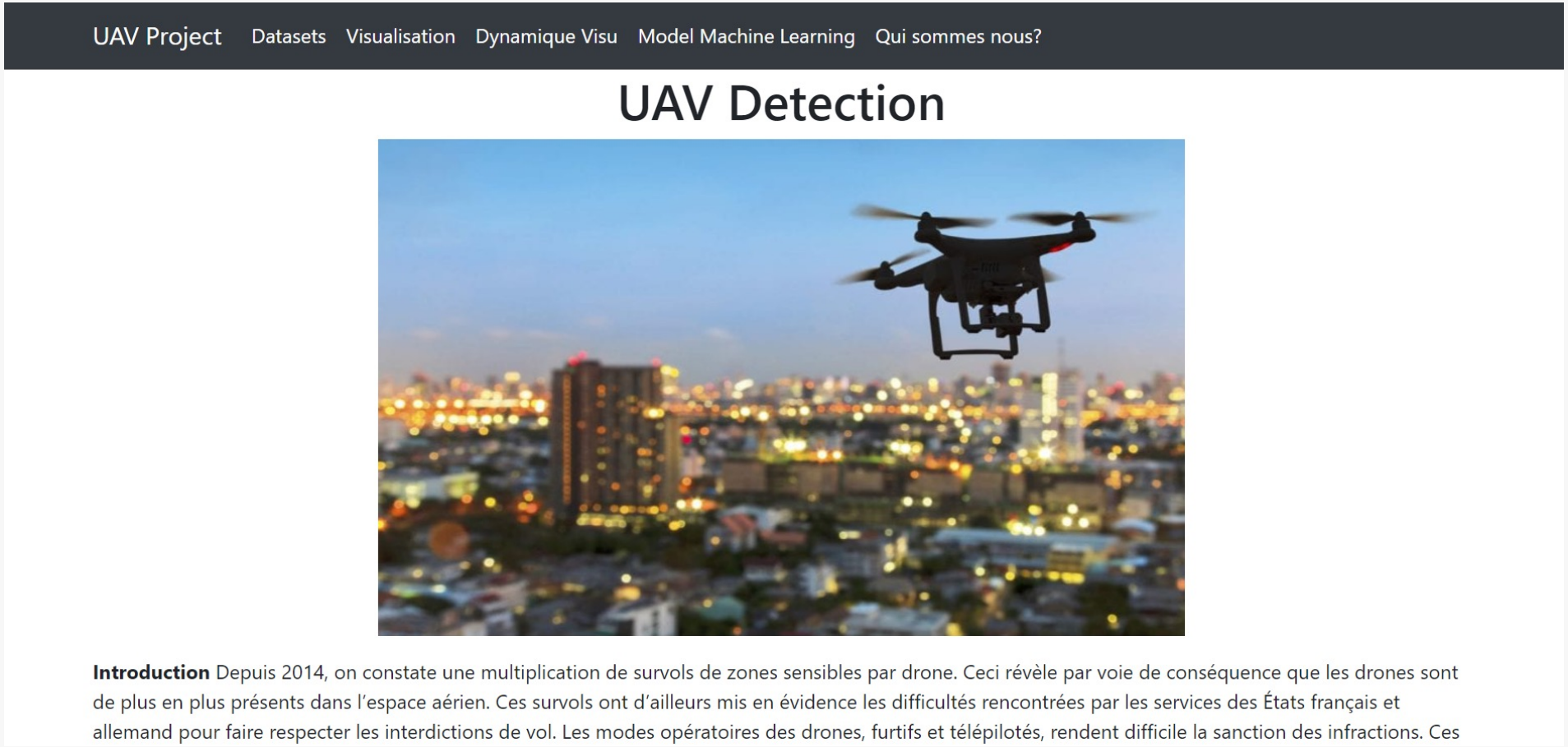
Conclusion:

- Le model SVC est le model le plus performant dans cette configuration.
- Ni centré réduire, ni étudier les composantes principales n'a été pertinent ici.

Pour aller plus loin:

- Agrandir le dataset et essayer de comprendre les features en captant les flux de plusieurs drones nous même
- Application de l'IA dans plusieurs cas concrets.

Démo API Django



Calendrier

COACHING

- **25/11 18H30**

Répartition des taches

Dataset (champs labels ect...)

Quelles techniques

- **29/11 17H30**

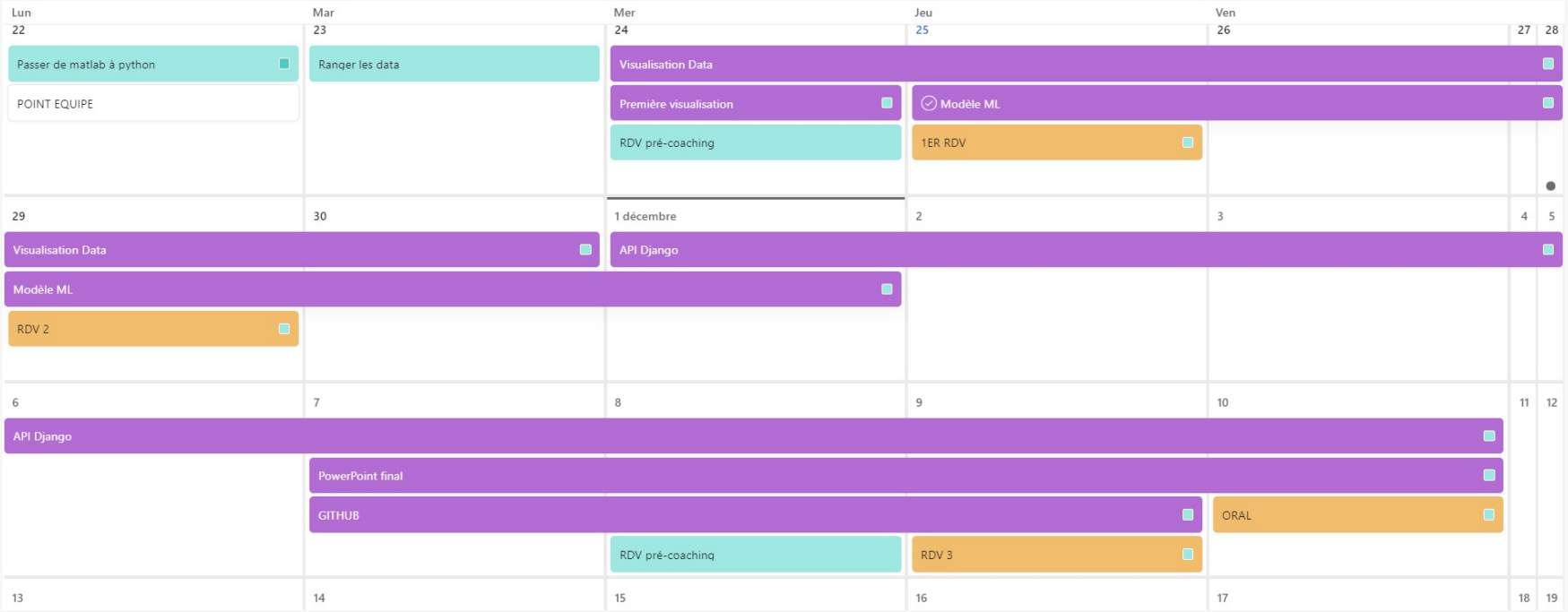
Premiers résultats

Avancées, Bouts de code

Structure de la présentation

- **9/12 18H30**

Présentation a blanc



Travail à faire sur chaque problématique :

- Data Vizualisation
- Modeling

Autres Taches :

- Github
- PPT
- API, Django

Problèmes rencontrés

DOWNLOAD DES DATASETS

COMPRENDRE LES FEATURES

TEMPS DE TRAVAIL LIMITÉ

Conclusion

Conclusion:

- Détecter si un drone rentre dans une zone équipée d’une antenne wifi.
- Détecter le model de drone détecté.
- Travail très formateur en mode sprint.

Pour aller plus loin...

- Discuter avec Liang Zhao de ce projet et de la création du dataset.
- L'objectif à terme est de tester les model en condition réel.

Liens utiles

- Datasets :
- http://mason.gmu.edu/~lzhao9/materials/data/UAV/
- https://archive.ics.uci.edu/ml/datasets/Unmanned+Aerial+Vehicle+%28UAV%29+Intrusion+Detection

Papiers traitant de ce type de projets:

- https://par.nsf.gov/servlets/purl/10135595
- https://arxiv.org/pdf/1905.06396.pdf

Contact de Liang Zaho : lzhao9@gmu.edu