

Projet Final

Python

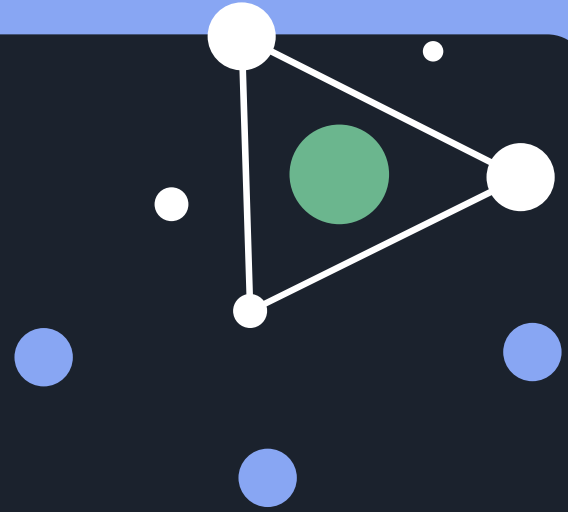




Table of contents

01

Intro

We will talk about the topic
of the study

03

Our Visualization

We will show our
visualization

02

Our dataset

We will describe the dataset

04

Our Modeling

You can explain our choice



The ins and out of the **problem** ?



**Data
preprocessing**

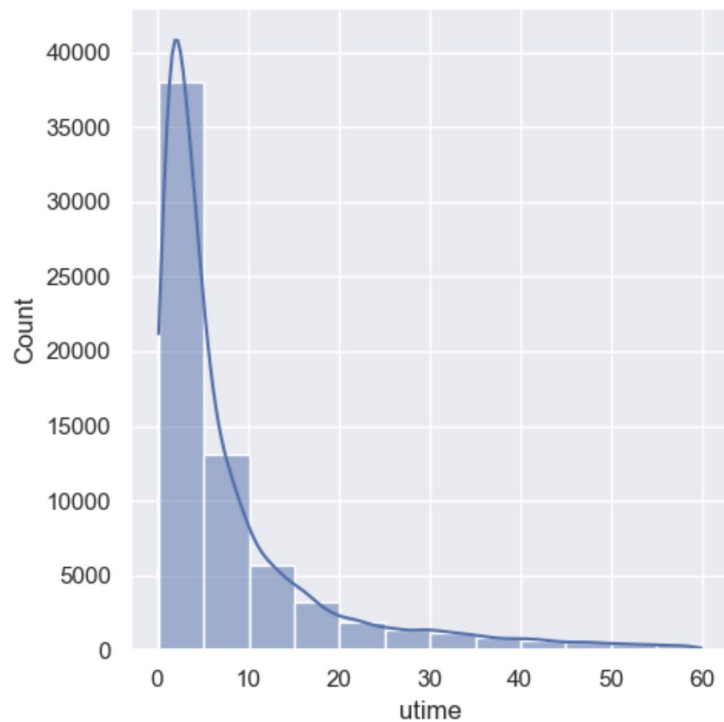


**Data
visualization**



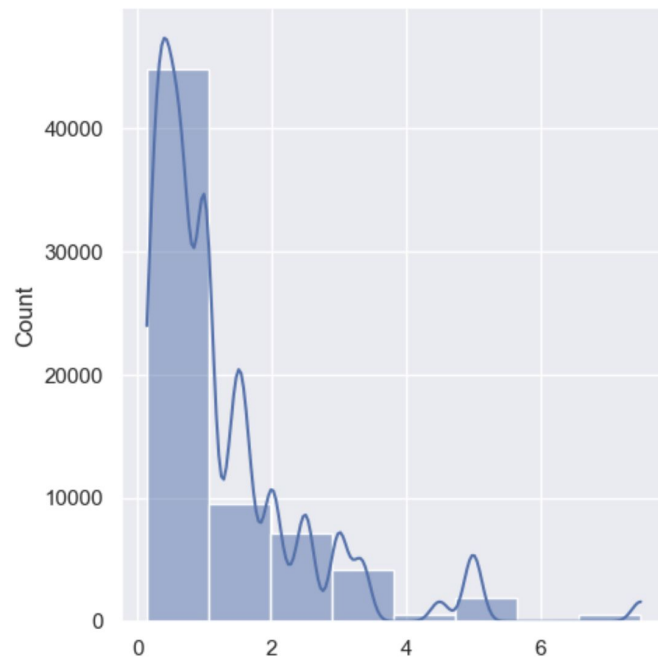
Modeling

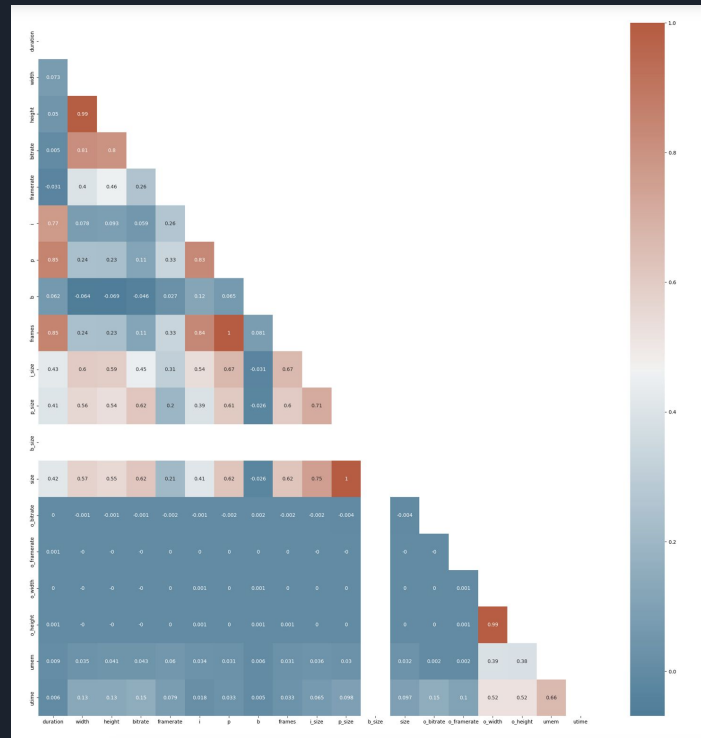




```
sns.displot(Rapport_hauteur_io, kde=True, bins=8)
```

```
<seaborn.axisgrid.FacetGrid at 0x295f94510>
```





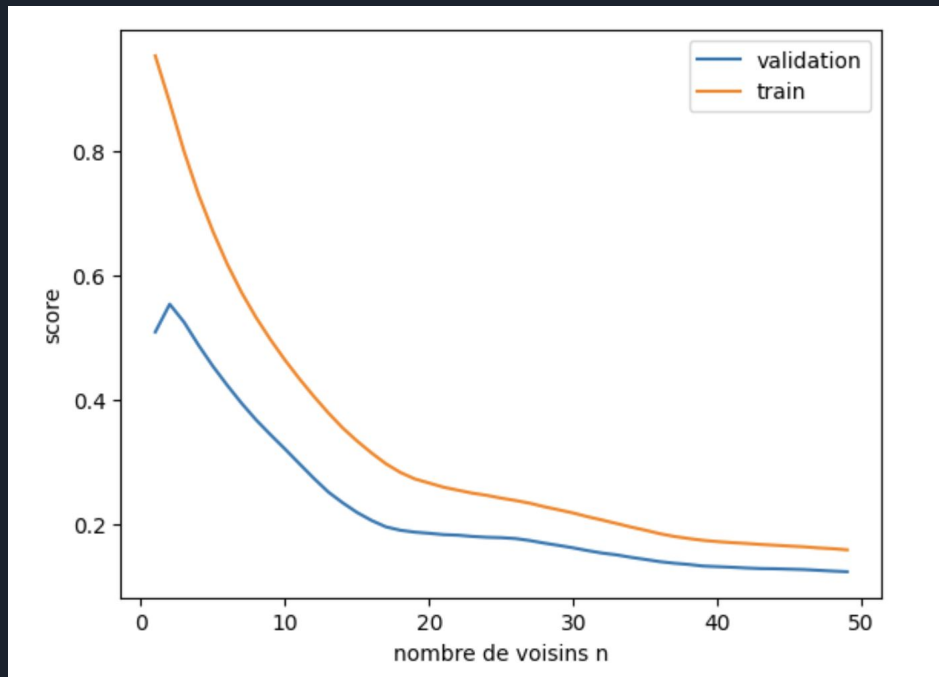


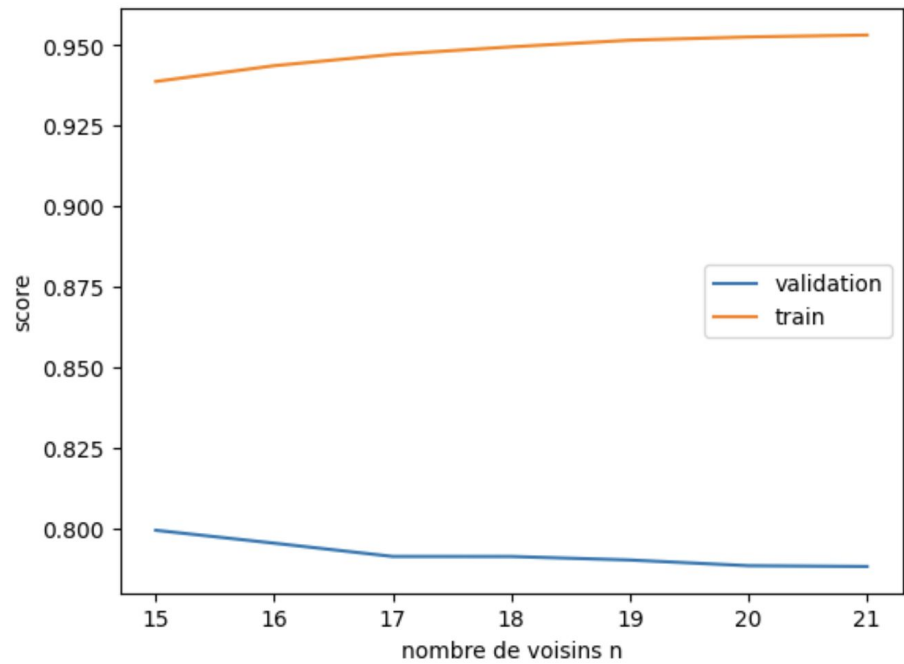
1) KN Regressor


```
from sklearn.neighbors import KNeighborsRegressor  
  
neigh = KNeighborsRegressor(n_neighbors=3)  
neigh.fit(X_train, y_train)  
print('Train score: ',neigh.score(X_train,y_train))  
print('Test score: ',neigh.score(X_test,y_test))
```

Train score: 0.819415162730038

Test score: 0.5784490223166803







```
In [22]: from sklearn.model_selection import GridSearchCV

param_grid = { 'max_depth': np.arange(15,30), 'random_state': np.arange(0,
grid = GridSearchCV(RandomForestRegressor(), param_grid, cv=5)
grid.fit(X_train, y_train)
```

```
Out[22]: GridSearchCV(cv=5, estimator=RandomForestRegressor(),
                    param_grid={'max_depth': array([15, 16, 17, 18, 19, 20, 21
, 22, 23, 24, 25, 26, 27, 28, 29]),
                    'random_state': array([0, 1, 2])})
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.

On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

```
In [24]: grid.best_score_
```

```
Out[24]: 0.7989467655287086
```

```
In [25]: grid.best_params_
```

```
Out[25]: {'max_depth': 15, 'random_state': 1}
```

