

Hyperparameters and Grid Search

Overview

1

What are hyperparameters?

2

Hyperparameters in decision trees and random forest

3

Grid search as a way to optimise hyperparameters

4

NOTEBOOK: Optimising the decision tree.

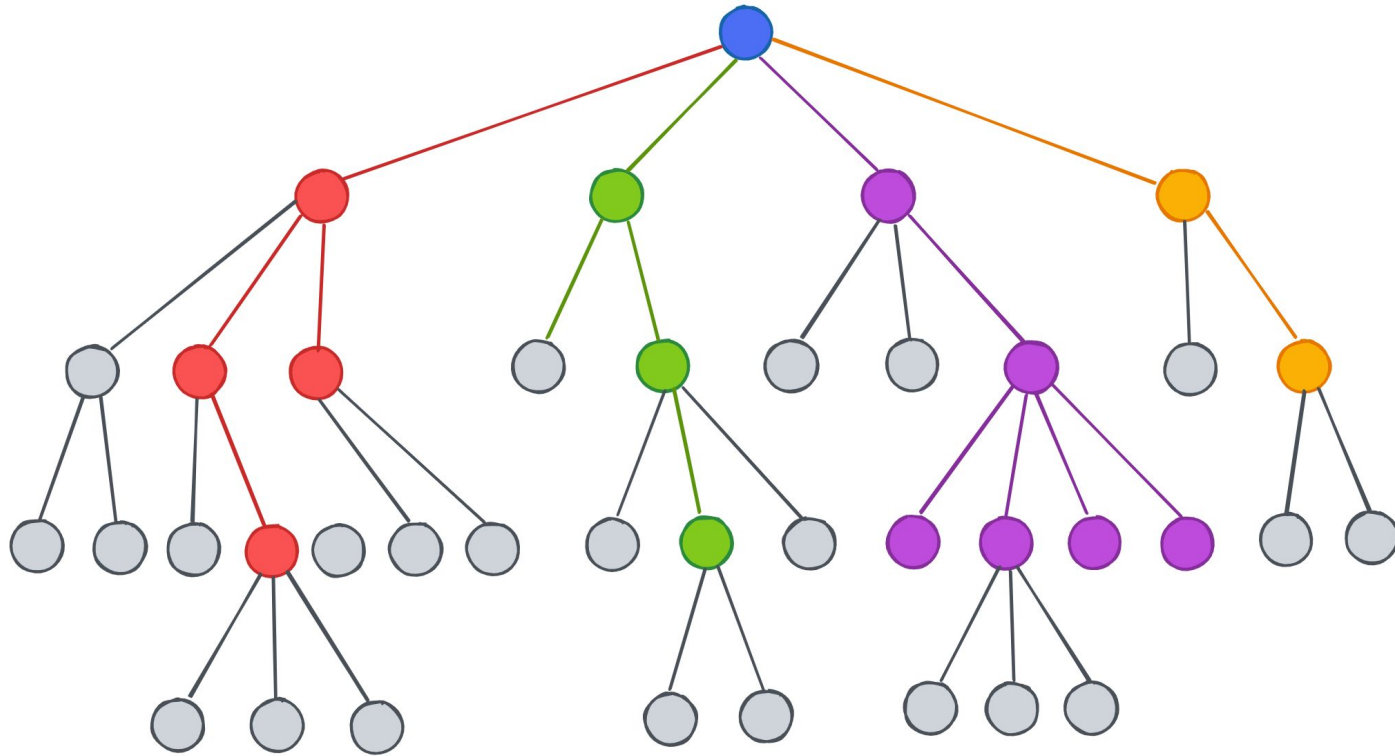
Hyperparameters

Parameters vs Hyperparameters

Parameters: Learnt from the data. Determine how the input is mapped to the output.

Hyperparameters: Parameters controlling the learning process. Not learnt from the data. User choices! E.g. number of trees to include in your random forest.

Hyperparameters in a decision tree?



[Source](#)



CONFIDENTIAL: Oxmedica Ltd. 10347756



What do these look like in Sklearn?

```
class sklearn.tree.DecisionTreeClassifier(*, criterion='gini',
splitter='best', max_depth=None, min_samples_split=2, min_samples_leaf=1,
min_weight_fraction_leaf=0.0, max_features=None, random_state=None,
max_leaf_nodes=None, min_impurity_decrease=0.0, class_weight=None,
ccp_alpha=0.0, monotonic_cst=None)
```

[\[source\]](#)

- What are default values?
- How do you specify a different hyperparameter?
- Will hyperparameters improve the model or not?

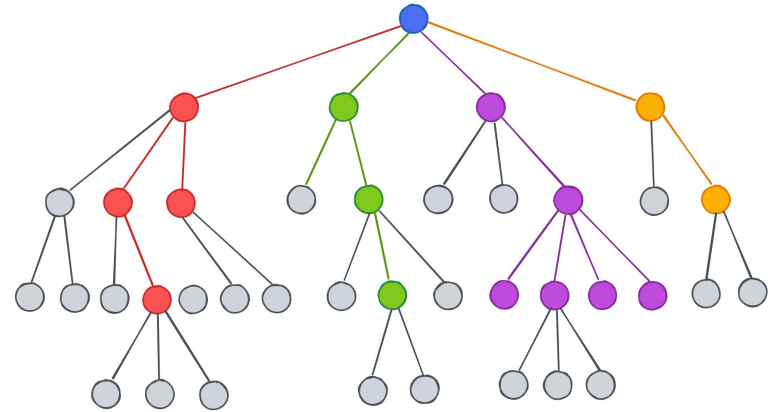
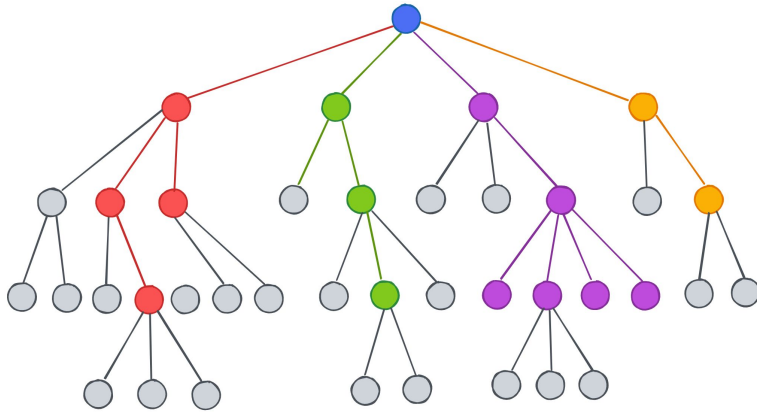
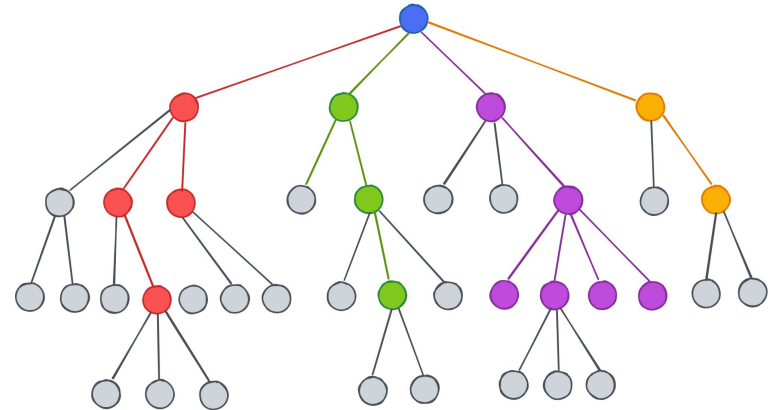
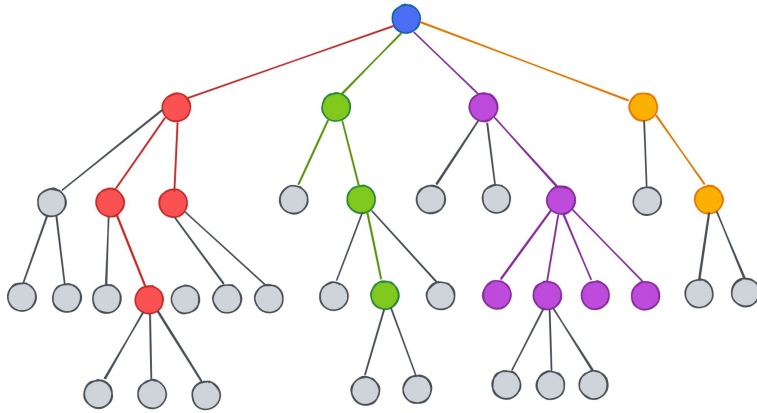
[Source](#)



CONFIDENTIAL: Oxmedica Ltd. 10347756



What about in random forest?



[Source](#)

What do these look like in Sklearn?

RandomForestClassifier

```
class sklearn.ensemble.RandomForestClassifier(n_estimators=100, *,
criterion='gini', max_depth=None, min_samples_split=2, min_samples_leaf=1,
min_weight_fraction_leaf=0.0, max_features='sqrt', max_leaf_nodes=None,
min_impurity_decrease=0.0, bootstrap=True, oob_score=False, n_jobs=None,
random_state=None, verbose=0, warm_start=False, class_weight=None,
ccp_alpha=0.0, max_samples=None, monotonic_cst=None) # \[source\]
```

[Source](#)



CONFIDENTIAL: Oxmedica Ltd. 10347756



Optimising Hyperparameters

General idea...

See whiteboard



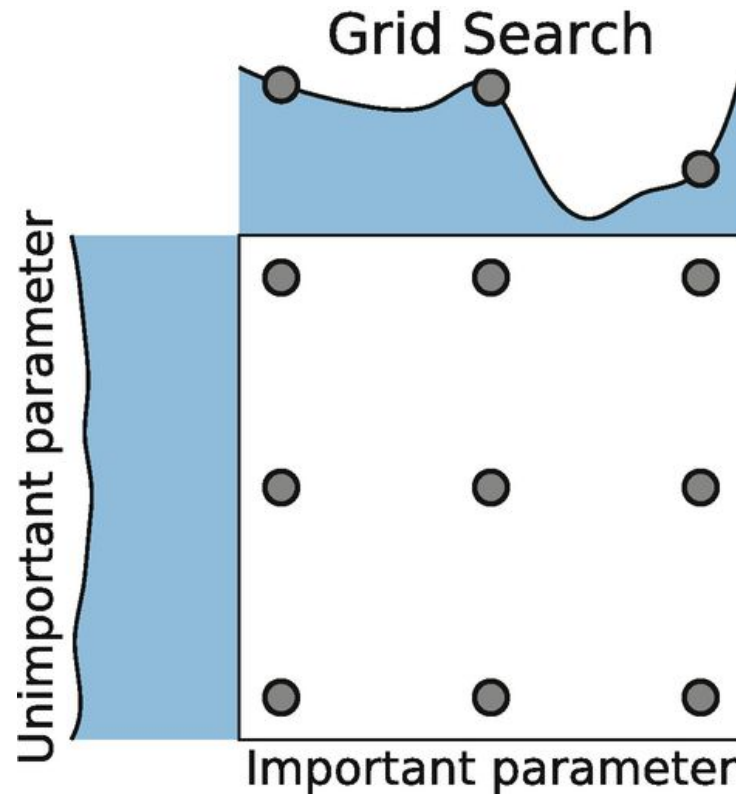
Any ideas of how we might do this?

→ What is the **most basic way possible**?

→ Improvements on that?

Grid Search

Grid search visualised



[source](#)



CONFIDENTIAL: Oxmedica Ltd. 10347756



How might we do this in code...



Notebook



harrymayne.com

