

2. Random Forest

Οι κλάσεις `custom_learning_curve`, `Node`, `ID3` έχουν υλοποιηθεί με βάση το εργαστήριο και είναι για την δημιουργία καμπυλών, κόμβων και ενός δέντρου απόφασης αντίστοιχα.

Η κλάση `RandomForest` υλοποιεί τον αλγόριθμο για να έχουμε πολλά δέντρα απόφασης- δάσους.

`def __init__(self, n_trees, max_depth=None)`: Αρχικοποιείται ο αλγόριθμος με τον αριθμό δέντρων και το βάθος που του έχουμε δώσει.

`def fit(self, x, y)`: Κατά την εκπαίδευση του μοντέλου και για κάθε επανάληψη φτιάχνουμε ένα δείγμα δεδομένων και υλοποιούμε το δέντρο σύμφωνα με τον ID3 αλγόριθμο, προσθέτοντας το κάθε εκπαιδευμένο δέντρο στην λίστα του δάσους.

`def predict(self, x)`: Για κάθε δέντρο κάνουμε την αντίστοιχη πρόβλεψη και στο τέλος τις συνδυάζουμε για να βγάλουμε το τελικό αποτέλεσμα με βάση την πλειοψηφία.

Επιλογή Αριθμού Δέντρων

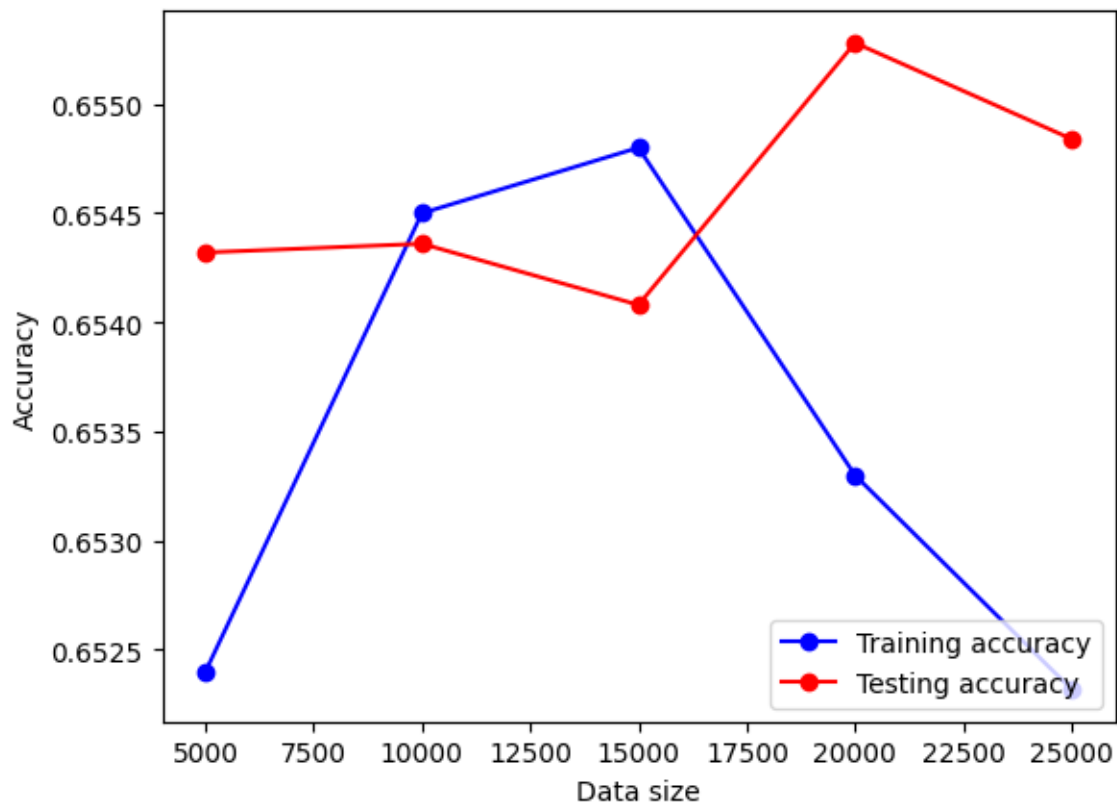
Τρέχοντας τον αλγόριθμο για : 1,5,10,30,50 δέντρα με μέγιστο βάθος = 5 βλέπω για την ακρίβεια του αλγόριθμου τα παρακάτω αποτελέσματα.

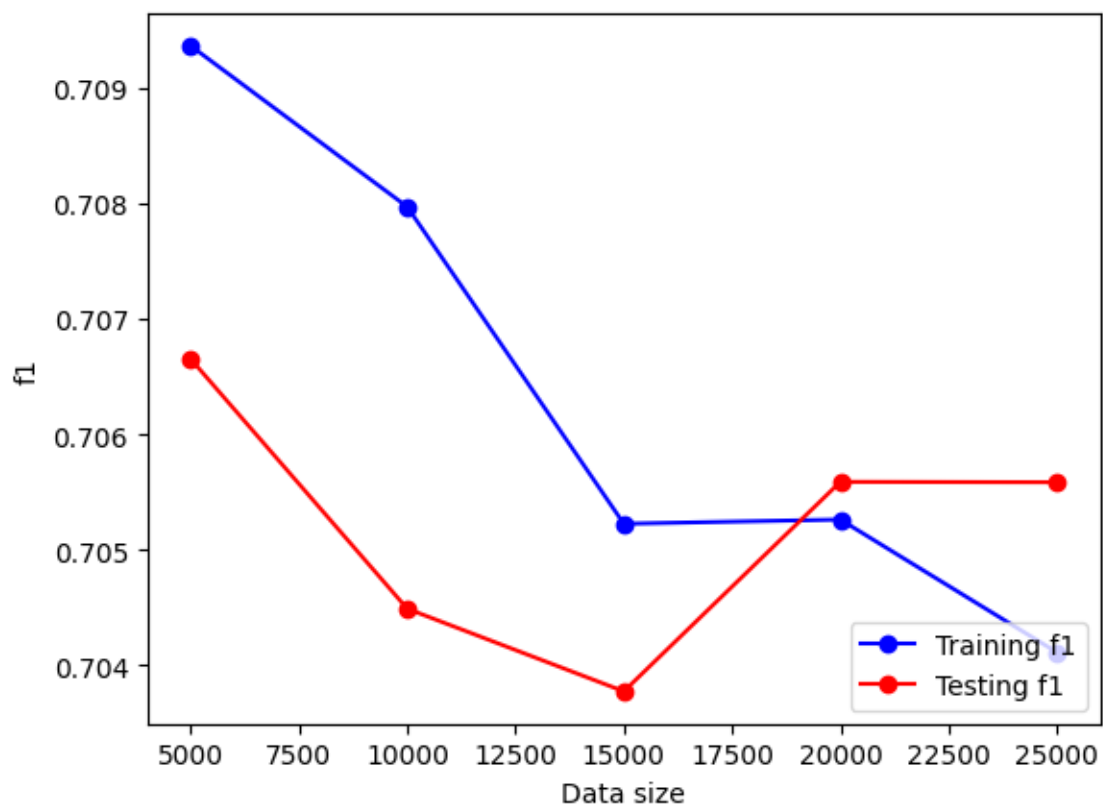
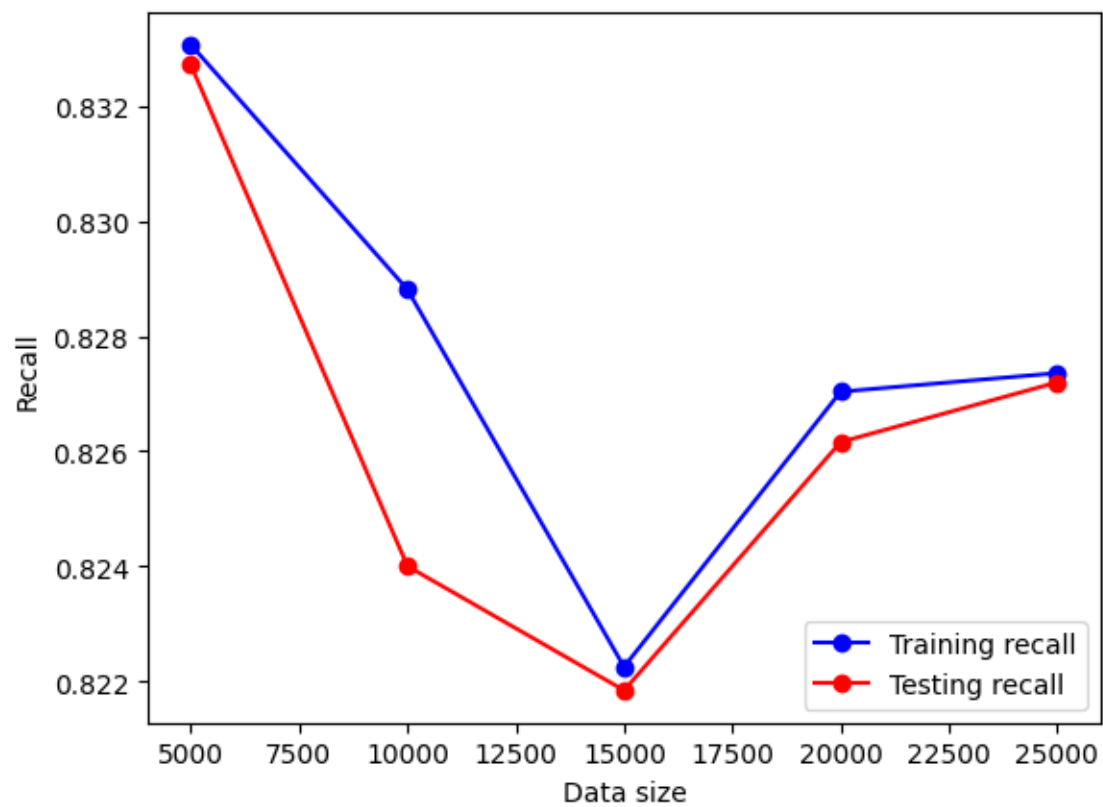
Επιλέγω 5 δέντρα διότι έχει την μεγαλύτερη ακρίβεια με ελάχιστη διαφορά.

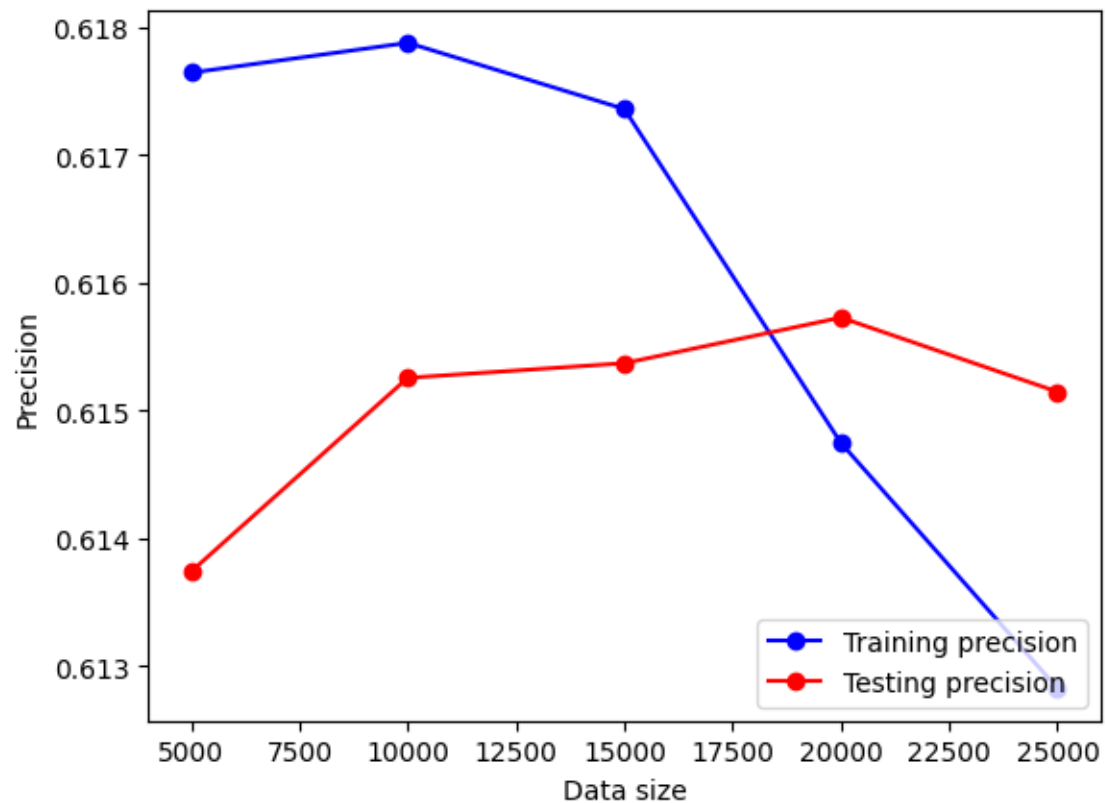
Tree 1					
0.65232					
0.65464					
Tree 5		precision	recall	f1-score	support
0.6732	0	0.73	0.48	0.58	12500
0.67408	1	0.61	0.83	0.70	12500
Tree 10					
0.652	accuracy			0.65	25000
0.65456	macro avg	0.67	0.65	0.64	25000
	weighted avg	0.67	0.65	0.64	25000
Tree 30		precision	recall	f1-score	support
0.65252	0	0.74	0.48	0.58	12500
0.65456	1	0.62	0.83	0.71	12500
Tree 50					
0.6518	accuracy			0.65	25000
0.65452	macro avg	0.68	0.65	0.64	25000
	weighted avg	0.68	0.65	0.64	25000

Καμπύλες με τα

αποτελέσματα για n trees = 5







Β μέρος

Αποτελέσματα έτοιμου αλγόριθμου :

```
100%|██████████| 25000/25000 [01:08<00:00, 366.37it/s]
100%|██████████| 25000/25000 [01:24<00:00, 294.23it/s]
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	12500
1	1.00	1.00	1.00	12500
accuracy			1.00	25000
macro avg	1.00	1.00	1.00	25000
weighted avg	1.00	1.00	1.00	25000

	precision	recall	f1-score	support
0	0.83	0.83	0.83	12500
1	0.83	0.83	0.83	12500
accuracy			0.83	25000
macro avg	0.83	0.83	0.83	25000
weighted avg	0.83	0.83	0.83	25000

