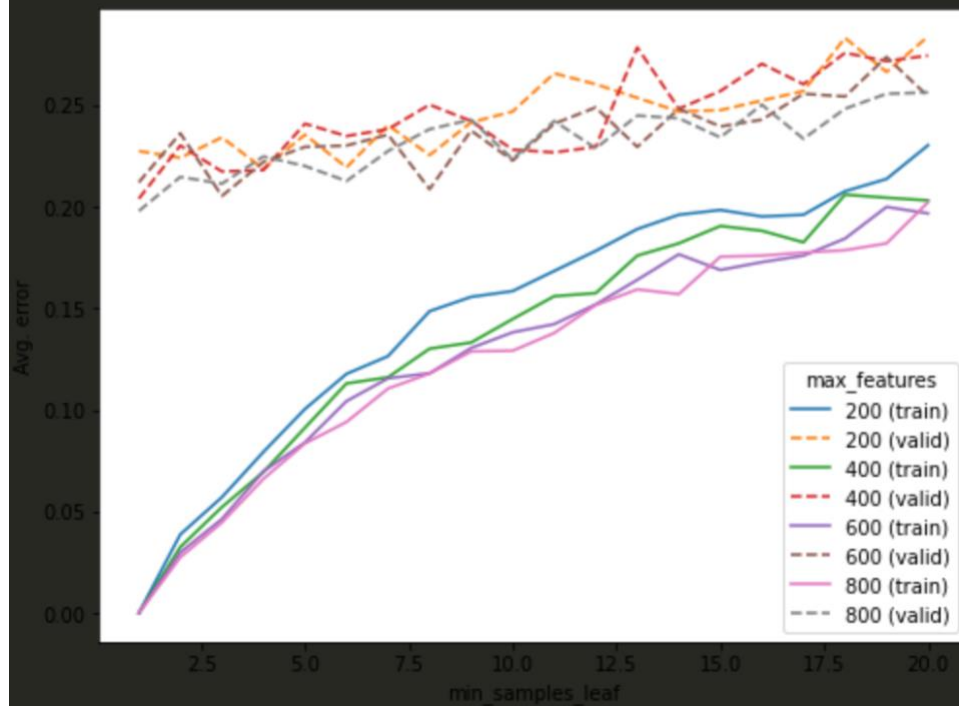
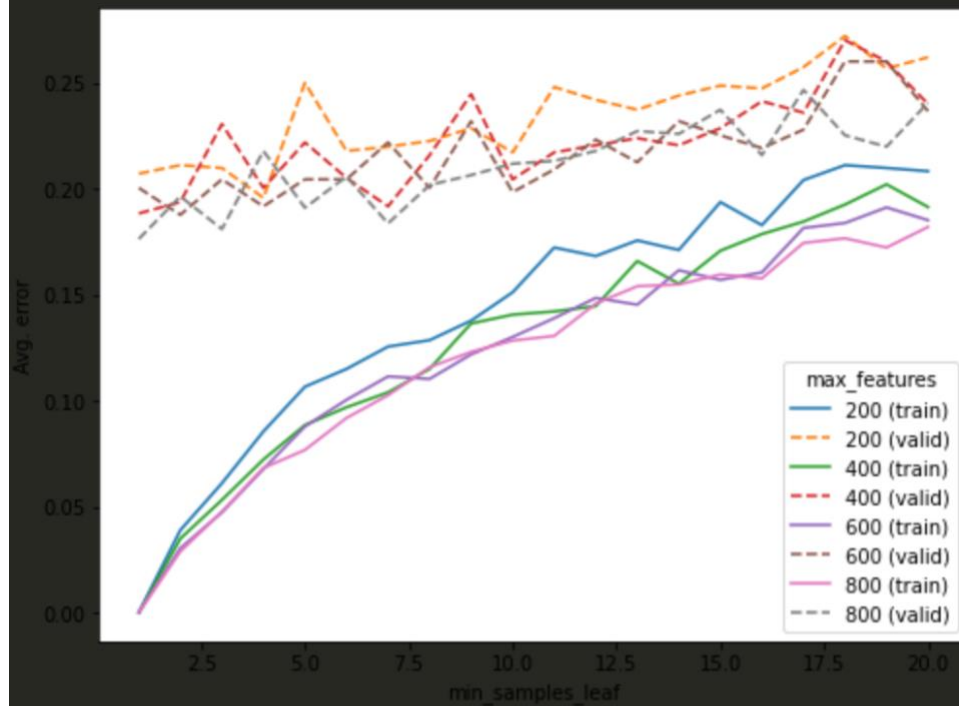


# Decision Tree Classifier

```
(class) DecisionTreeClassifier(*, criterion="gini", splitter="best", max_depth=None,  
min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0, max_features=None,  
random_state=None, max_leaf_nodes=None, min_impurity_decrease=0, class_weight=None, ccp_alpha=0)
```

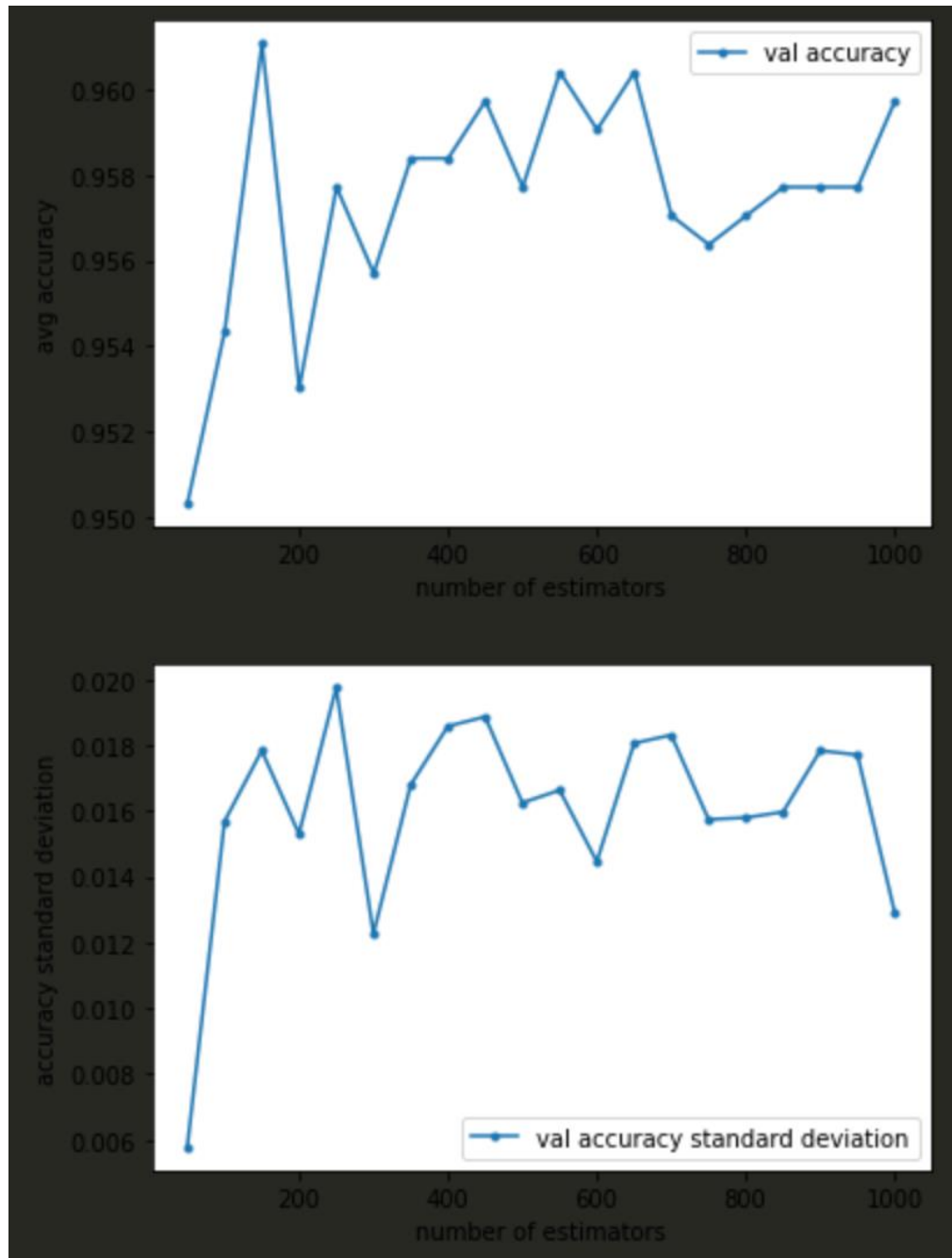


	criterion	max_features	min_samples_leaf	mean_train_score \
0	gini	200	1	1.000000
1	gini	200	2	0.961409
2	gini	200	3	0.930705
3	gini	200	4	0.916275
4	gini	200	5	0.903188
..	...	...	...	...
155	entropy	800	16	0.828859
156	entropy	800	17	0.822651
157	entropy	800	18	0.825503
158	entropy	800	19	0.810403
159	entropy	800	20	0.803356

	mean_test_score
0	0.799329
1	0.796644
2	0.801342
3	0.795973
4	0.760403
..	...
155	0.758389
156	0.754362
157	0.742953
158	0.744966
159	0.738926

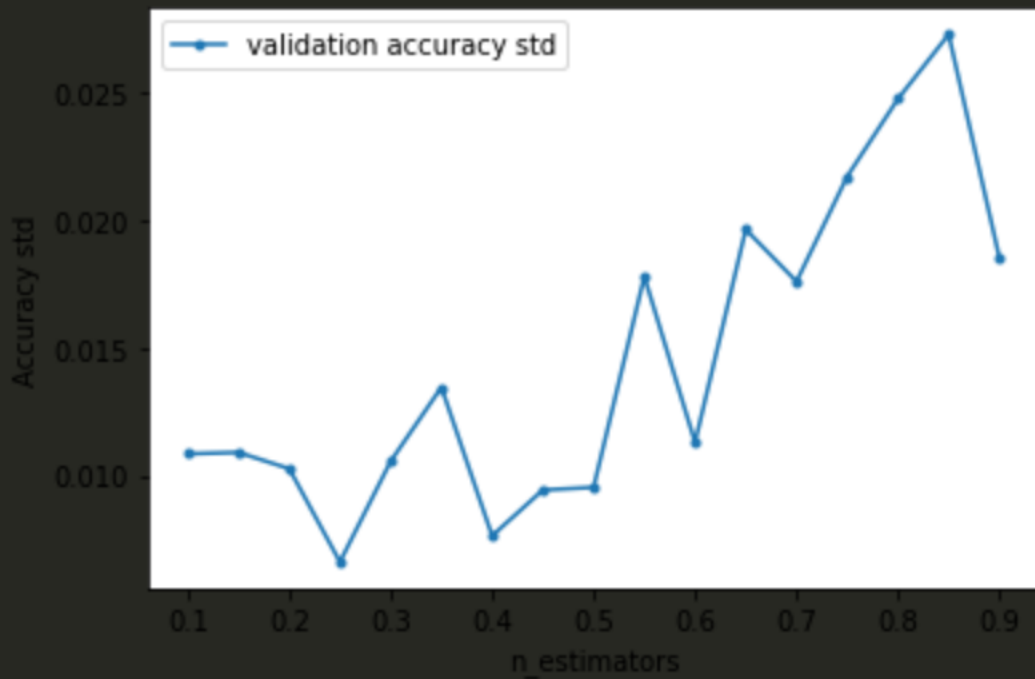
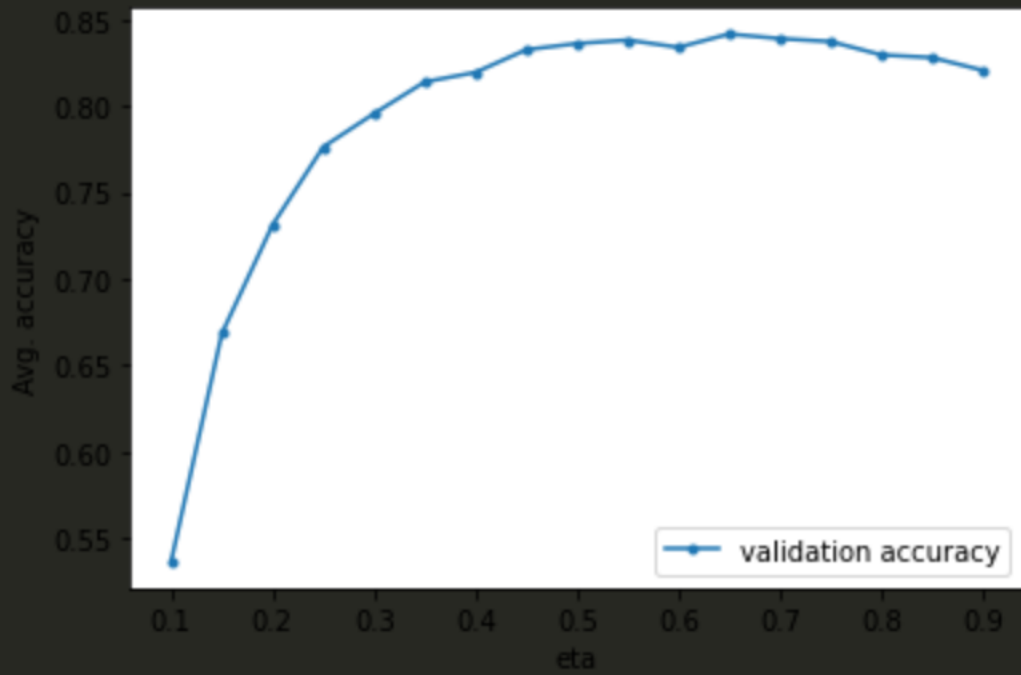
# Random Forest Classifier

```
(class) RandomForestClassifier(n_estimators=100, *, criterion="gini", max_depth=None,
min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0, max_features="auto",
max_leaf_nodes=None, min_impurity_decrease=0, bootstrap=True, oob_score=False, n_jobs=None,
random_state=None, verbose=0, warm_start=False, class_weight=None, ccp_alpha=0,
max_samples=None)
```



N estimators	Avg accuracy	Std Accuracy
50	0.94563758	0.01649424
100	0.95033557	0.01189265
150	0.95302013	0.0171108
200	0.95704698	0.01818284
250	0.95704698	0.01565356
300	0.9590604	0.0167651
350	0.96107383	0.0171896
400	0.95838926	0.01509694
450	0.95369128	0.01550902
500	0.9590604	0.01565356
550	0.9557047	0.01090475
600	0.96107383	0.01211777
650	0.95838926	0.01301391
700	0.95973154	0.01470396
750	0.95771812	0.01596695
800	0.95704698	0.01867172
850	0.95838926	0.01448794
900	0.96174497	0.01679194
950	0.95838926	0.01448794
1000	0.9590604	0.01607939

# Gradient Boosting



	train-mlogloss-mean	train-mlogloss-std	test-mlogloss-mean	\
0	0.278183	0.003545	0.463401	
1	0.144048	0.002822	0.331574	
2	0.080748	0.002885	0.268764	
3	0.049655	0.001800	0.223669	
4	0.032534	0.001577	0.204385	
5	0.023008	0.001068	0.186058	
6	0.017400	0.000904	0.180564	
7	0.013752	0.000626	0.167432	
8	0.011481	0.000403	0.163868	
9	0.009754	0.000347	0.161864	
10	0.008746	0.000317	0.166063	
11	0.008076	0.000312	0.158370	
12	0.007423	0.000312	0.161010	
13	0.006830	0.000222	0.162724	
14	0.006523	0.000197	0.170418	
15	0.006188	0.000127	0.172007	
16	0.005914	0.000168	0.179345	
	test-mlogloss-std	eta		
0	0.010855	0.10		
1	0.010908	0.15		
2	0.010274	0.20		
3	0.006626	0.25		
4	0.010563	0.30		
5	0.013458	0.35		
6	0.007648	0.40		
7	0.009439	0.45		
8	0.009543	0.50		
9	0.017788	0.55		
10	0.011303	0.60		
11	0.019633	0.65		
12	0.017594	0.70		
13	0.021707	0.75		
14	0.024738	0.80		
15	0.027260	0.85		
16	0.018522	0.90		

Using Random Forest Classifier was the most accurate so therefore for my final part I went with the random forest classifier.

```
tester = RandomForestClassifier(criterion='gini', n_estimators=1000, max_features='0.5', min_samples_leaf=1)
```