

In [52]: *#checking the results dataframe*
results

Out[52]:

	Methodology	Model	Accuracy	roc_value	Threshold
0	RepeatedKFold Cross validation	Logistic Regression with L1 regularisation	0.999122	0.918748	0.054333
1	RepeatedKFold Cross validation	Logistic Regression with L2 regularisation	0.999175	0.986403	0.002160
2	RepeatedKFold Cross validation	KNN	0.999350	0.907030	0.200000
3	RepeatedKFold Cross validation	Tree Model with gini criteria	0.999034	0.906891	1.000000
4	RepeatedKFold Cross validation	Tree Model with entropy criteria	0.999228	0.937863	1.000000
5	RepeatedKFold Cross validation	Random Forest	0.999631	0.973048	0.010000
6	RepeatedKFold Cross validation	XGBoost	0.999666	0.983353	0.000244
7	RepeatedKFold Cross validation	SVM	0.998385	0.513285	0.003033
8	StratifiedKFold Cross validation	Logistic Regression with L1 regularisation	0.998754	0.889076	0.021087
9	StratifiedKFold Cross validation	Logistic Regression with L2 regularisation	0.998771	0.983423	0.001542
10	StratifiedKFold Cross validation	KNN	0.999192	0.805746	0.200000
11	StratifiedKFold Cross validation	Tree Model with gini criteria	0.998841	0.826249	1.000000
12	StratifiedKFold Cross validation	Tree Model with entropy criteria	0.999017	0.821244	1.000000
13	StratifiedKFold Cross validation	Random Forest	0.999438	0.946472	0.010000
14	StratifiedKFold Cross validation	XGBoost	0.999438	0.972205	0.000038
15	StratifiedKFold Cross validation	SVM	0.998280	0.401770	0.004330
16	Random Oversampling with stratifiedKFold CV	Logistic Regression with L1 regularisation	0.998280	0.500000	1.500000
17	Random Oversampling with stratifiedKFold CV	Logistic Regression with L2 regularisation	0.986324	0.980645	0.430392
18	Random Oversampling with stratifiedKFold CV	KNN	0.998069	0.805604	0.200000
19	Random Oversampling with stratifiedKFold CV	Tree Model with gini criteria	0.999017	0.811057	1.000000
20	Random Oversampling with stratifiedKFold CV	Tree Model with entropy criteria	0.999263	0.846833	1.000000
21	Random Oversampling with stratifiedKFold CV	Random Forest	0.999508	0.930504	0.030000
22	Random Oversampling with stratifiedKFold CV	XGBoost	0.999456	0.972159	0.000178

It's interesting to note that Logistic Regression with L2 regularization and XGBoost, both utilizing the RepeatedKFold cross-validation method, outperform the StratifiedKFold cross-validation method. Additionally, they even outperform the approach of random