

S.No	Features - used	Algorithm	Precision	Recall	Accuracy	PCA		Precision	Recall		Final Result	
	Note :All metrics calculated for stratified shuffle split , iterated overSIX times ,therefore there might be a variations depending on the iteration value.											
1	Basic email + Financial data	ExtraTreesClassifier	0.36551	0.2215	0.84493	FALSE		TRUE	FALSE		FALSE	
2	Basic email + Financial data	ExtraTreesClassifier	0.3385	0.142	0.8486	TRUE		TRUE	FALSE		FALSE	
3	Basic email + Financial data	Logistic Regression - Liblinear	0.56655	0.1575	0.8716	FALSE		TRUE	FALSE		FALSE	
4	Basic email + Financial data	Logistic Regression - Liblinear(with dual)	0.28553	0.7885	0.70873	FALSE		FALSE	TRUE		FALSE	
5	Basic email + Financial data	Logistic Regression - Newton-CG	0.49035	0.127	0.866	FALSE		TRUE	FALSE		FALSE	
6	Basic email + Financial data	Logistic Regression - Liblinear	0.29165	0.8665	0.7016	TRUE		FALSE	TRUE		FALSE	
7	Basic email + Financial data	Logistic Regression - Liblinear(with dual)	0.29182	0.8165	0.71133	TRUE		FALSE	TRUE		FALSE	
8	Basic email + Financial data	Logistic Regression - Newton-CG	0.47399	0.205	0.86367	FALSE		TRUE	FALSE		FALSE	
9	Basic email + Financial data	Linear SVM	0.46726	0.132	0.8642	FALSE		TRUE	FALSE		FALSE	
10	Basic email + Financial data	Linear SVM	0.46364	0.204	0.8624	TRUE		TRUE	FALSE		FALSE	
	Created from_poi_to_total_emails ratio , to_poi_total_emails , shared_receipt_with POI / Total emails ratios											
11	Financial data + emails + emails ratio data	ExtraTreesClassifier	0.39381	0.2355	0.84973	FALSE		TRUE	FALSE		FALSE	
12	Financial data + emails + emails ratio data	Logistic Regression - Liblinear	0.56724	0.1645	0.87187	FALSE		TRUE	FALSE		FALSE	
13	Financial data + emails + emails ratio data	Logistic Regression - Liblinear(with dual)	0.57301	0.155	0.87193	FALSE		TRUE	FALSE		FALSE	
14	Financial data + emails + emails ratio data	Logistic Regression - Newton-CG	0.4375	0.0735	0.86387	FALSE		TRUE	FALSE		FALSE	
15	Financial data + emails + emails ratio data	ExtraTreesClassifier	0.41988	0.2175	0.8556	TRUE		TRUE	FALSE		FALSE	
16	Financial data + emails + emails ratio data	Logistic Regression - Liblinear	0.28173	0.808	0.69973	TRUE		FALSE	TRUE		FALSE	
17	Financial data + emails + emails ratio data	Logistic Regression - Liblinear(with dual)	0.28539	0.752	0.71587	TRUE		FALSE	TRUE		FALSE	
18	Financial data + emails + emails ratio data	Logistic Regression - Newton-CG	0.47807	0.218	0.864	TRUE		TRUE	FALSE		FALSE	
19	Financial data + emails + emails ratio data	Linear SVM	0.50405	0.1865	0.86707	FALSE		TRUE	FALSE		FALSE	
20	Financial data + emails + emails ratio data	Linear SVM	0.6631	0.186	0.87887	TRUE		TRUE	FALSE		FALSE	
	Added feature by taking log of financial features											
21	Log(Financial data) + emails + emails ratio data	ExtraTreesClassifier	0.51754	0.332	0.86967	FALSE		TRUE	TRUE		TRUE	f1-score 0.42477
22	Log(Financial data) + emails + emails ratio data	Logistic Regression - Liblinear	0.06667	0.0005	0.8658	FALSE		FALSE	FALSE		FALSE	
23	Log(Financial data) + emails + emails ratio data	Logistic Regression - Liblinear(with dual)	0.06667	0.0005	0.8658	FALSE		FALSE	FALSE		FALSE	
24	Log(Financial data) + emails + emails ratio data	Logistic Regression - Newton-CG	0.4684	0.063	0.86553	FALSE		TRUE	FALSE		FALSE	
25	Log(Financial data) + emails + emails ratio data	ExtraTreesClassifier	0.47018	0.3075	0.86147	TRUE		TRUE	TRUE		TRUE	f1-score 0.38884
26	Log(Financial data) + emails + emails ratio data	Logistic Regression - Liblinear	0.25448	0.9015	0.87313	TRUE		FALSE	TRUE		FALSE	
27	Log(Financial data) + emails + emails ratio data	Logistic Regression - Liblinear(with dual)	0.258	0.891	0.6438	TRUE		FALSE	TRUE		FALSE	
28	Log(Financial data) + emails + emails ratio data	Logistic Regression - Newton-CG	0.25653	0.909	0.6366	TRUE		FALSE	TRUE		FALSE	
29	Log(Financial data) + emails + emails ratio data	Linear SVM	0.50646	0.196	0.86733	FALSE		TRUE	FALSE		FALSE	
30	Log(Financial data) + emails + emails ratio data	Linear SVM	0.5281	0.249	0.8702	TRUE		TRUE	FALSE		FALSE	
	Added all features created so far											
31	Log(finances)+finances + emails + emails ratio	ExtraTreesClassifier	0.41524	0.2425	0.85347	FALSE		TRUE	FALSE		FALSE	
32	Log(finances)+finances + emails + emails ratio	Logistic Regression - Liblinear	0.63958	0.181	0.8772	FALSE		TRUE	FALSE		FALSE	
33	Log(finances)+finances + emails + emails ratio	Logistic Regression - Liblinear(with dual)	0.71828	0.167	0.8802	FALSE		TRUE	FALSE		FALSE	

34	Log(finances)+finances + emails + emails ratio	Logistic Regression - Newton-CG	0.64007	0.1885	0.87767	FALSE		TRUE	FALSE		FALSE		
35	Log(finances)+finances + emails + emails ratio	ExtraTreesClassifier	0.49844	0.3205	0.8664	TRUE		TRUE	TRUE		TRUE	f1-score	0.40947
36	Log(finances)+finances + emails + emails ratio	Logistic Regression - Liblinear	0.47513	0.3105	0.86233	TRUE		TRUE	TRUE		TRUE	f1-score	0.392815
37	Log(finances)+finances + emails + emails ratio	Logistic Regression - Liblinear(with dual)	0.32336	0.778	0.75333	TRUE		TRUE	TRUE		TRUE	f1-score	0.55068
38	Log(finances)+finances + emails + emails ratio	Logistic Regression - Newton-CG	0.32195	0.7965	0.7492	TRUE		TRUE	TRUE		TRUE	f1-score	0.559225
39	Log(finances)+finances + emails + emails ratio	Linear SVM	0.46736	0.2255	0.86247	FALSE		TRUE	FALSE		FALSE		
40	Log(finances)+finances + emails + emails ratio	Linear SVM	0.56283	0.159	0.8714	TRUE		TRUE	FALSE		FALSE		
	Added (Financial feature/Salary) and applied to all financial data												
41	Normalized(finances) + emails	ExtraTreesClassifier	0.32056	0.1715	0.80133	FALSE		TRUE	FALSE		FALSE		
42	Normalized(finances) + emails	Logistic Regression - Liblinear	0.10798	0.0115	0.81942	FALSE		FALSE	FALSE		FALSE		
43	Normalized(finances) + emails	Logistic Regression - Liblinear(with dual)	0.11765	0.001	0.83225	FALSE		FALSE	FALSE		FALSE		
44	Normalized(finances) + emails	Logistic Regression - Newton-CG	0.35849	0.038	0.82833	FALSE		TRUE	FALSE		FALSE		
45	Normalized(finances) + emails	ExtraTreesClassifier	0.30544	0.1825	0.79458	TRUE		TRUE	FALSE		FALSE		
46	Normalized(finances) + emails	Logistic Regression - Liblinear	0.625	0.1325	0.84217	TRUE		TRUE	FALSE		FALSE		
47	Normalized(finances) + emails	Logistic Regression - Liblinear(with dual)	0.11765	0.001	0.83225	TRUE		FALSE	FALSE		FALSE		
48	Normalized(finances) + emails	Logistic Regression - Newton-CG	0.63342	0.127	0.84225	TRUE		TRUE	FALSE		FALSE		
49	Normalized(finances) + emails	Linear SVM	0.16608	0.0235	0.81758	FALSE		FALSE	FALSE		FALSE		
50	Normalized(finances) + emails	Linear SVM	0.63387	0.1385	0.84308	TRUE		TRUE	FALSE		FALSE		
	Added normal financial data to the above ^												
51	Normalised(finances) + Finances + emails	ExtraTreesClassifier	0.43682	0.2195	0.8582	FALSE		TRUE	FALSE		FALSE		
52	Normalised(finances) + Finances + emails	Logistic Regression - Liblinear	0.43725	0.108	0.86253	FALSE		TRUE	FALSE		FALSE		
53	Normalised(finances) + Finances + emails	Logistic Regression - Liblinear(with dual)	0.51129	0.1585	0.8676	FALSE		TRUE	FALSE		FALSE		
54	Normalised(finances) + Finances + emails	Logistic Regression - Newton-CG	0.51129	0.1585	0.8676	FALSE		TRUE	FALSE		FALSE		
55	Normalised(finances) + Finances + emails	ExtraTreesClassifier	0.48936	0.23	0.86533	TRUE		TRUE	FALSE		FALSE		
56	Normalised(finances) + Finances + emails	Logistic Regression - Liblinear	0.45667	0.166	0.86247	TRUE		TRUE	FALSE		FALSE		
57	Normalised(finances) + Finances + emails	Logistic Regression - Liblinear(with dual)	0.28122	0.6952	0.6952	TRUE		FALSE	TRUE		FALSE		
58	Normalised(finances) + Finances + emails	Logistic Regression - Newton-CG	0.28122	0.8265	0.6952	TRUE		FALSE	TRUE		FALSE		
59	Normalised(finances) + Finances + emails	Linear SVM	0.22656	0.2525	0.73175	FALSE		FALSE	FALSE		FALSE		
60	Normalised(finances) + Finances + emails	Linear SVM	0.23197	0.259	0.73358	TRUE		FALSE	FALSE		FALSE		
	Added all the features available												
61	Normalised(finances) + Finances + emails + email-ratio	ExtraTreesClassifier	0.38772	0.2305	0.84887	FALSE		TRUE	FALSE		FALSE		
62	Normalised(finances) + Finances + emails + email-ratio	Logistic Regression - Liblinear	0.48404	0.1365	0.86547	FALSE		TRUE	FALSE		FALSE		
63	Normalised(finances) + Finances + emails + email-ratio	Logistic Regression - Liblinear(with dual)	0.49416	0.148	0.8662	FALSE		TRUE	FALSE		FALSE		
64	Normalised(finances) + Finances + emails + email-ratio	Logistic Regression - Newton-CG	0.56167	0.1685	0.8716	FALSE		TRUE	FALSE		FALSE		
65	Normalised(finances) + Finances + emails + email-ratio	ExtraTreesClassifier	0.27886	0.244	0.81507	TRUE		FALSE	FALSE		FALSE		
66	Normalised(finances) + Finances + emails + email-ratio	Logistic Regression - Liblinear	0.48966	0.1775	0.86567	TRUE		TRUE	FALSE		FALSE		
67	Normalised(finances) + Finances + emails + email-ratio	Logistic Regression - Liblinear(with dual)	0.27935	0.821	0.69373	TRUE		FALSE	TRUE		FALSE		
68	Normalised(finances) + Finances + emails + email-ratio	Logistic Regression - Newton-CG	0.4862	0.185	0.86527	TRUE		TRUE	FALSE		FALSE		
69	Normalised(finances) + Finances + emails + email-ratio	Linear SVM	0.448	0.168	0.86147	FALSE		TRUE	FALSE		FALSE		

70	Normalised(finances) + Finances + emails + email-ratio	Linear SVM	0.47526	0.1585	0.86447	TRUE		TRUE	FALSE		FALSE		
	Chose the best features out of all and final iteration												
											#VALUE!		
71	Log(finances) + Normalized(finances)(2) + emails + ratios	ExtraTreesClassifier	0.40205	0.274	0.84887	FALSE		TRUE	FALSE		FALSE		
72	Log(finances) + Normalized(finances)(2) + emails + ratios	Logistic Regression - Liblinear	0.61513	0.187	0.876	FALSE		TRUE	FALSE		FALSE		
73	Log(finances) + Normalized(finances)(2) + emails + ratios	Logistic Regression - Liblinear(with dual)	0.61513	0.187	0.876	FALSE		TRUE	FALSE		FALSE		
74	Log(finances) + Normalized(finances)(2) + emails + ratios	Logistic Regression - Newton-CG	0.5575	0.1115	0.86973	FALSE		TRUE	FALSE		FALSE		
75	Log(finances) + Normalized(finances)(2) + emails + ratios	ExtraTreesClassifier	0.34805	0.268	0.83547	TRUE		TRUE	FALSE		FALSE		
76	Log(finances) + Normalized(finances)(2) + emails + ratios	Logistic Regression - Liblinear	0.31584	0.8695	0.73147	TRUE		TRUE	TRUE		TRUE	f1-score	0.59267
77	Log(finances) + Normalized(finances)(2) + emails + ratios	Logistic Regression - Liblinear(with dual)	0.31584	0.8695	0.73147	TRUE		TRUE	TRUE		TRUE	f1-score	0.59267
78	Log(finances) + Normalized(finances)(2) + emails + ratios	Logistic Regression - Newton-CG	0.31395	0.864	0.73013	TRUE		TRUE	TRUE		TRUE	f1-score	0.588975
79	Log(finances) + Normalized(finances)(2) + emails + ratios	Linear SVM	0.41779	0.3405	0.8488	FALSE		TRUE	TRUE		TRUE	f1-score	0.379145
80	Log(finances) + Normalized(finances)(2) + emails + ratios	Linear SVM	0.56022	0.3605	0.877	TRUE		TRUE	TRUE		TRUE	f1-score	0.46036