Table 1: using PCA reduce dimension to 500, no max feature, no grid search CV

Model Name	Accuracy	MSE	MAE	RMSE
Linear regression	0.10	5240.922	31.271	5.569
Naive Bayes	0.15	3.46	1.536	1.238
Random Forest	0.31	3.87	1.462	1.209
Decision Tree	0.25	4.41	1.632	1.277

Table 2: after max feature=100, No PCA, random forest and decision tree used grid search CV(using example parameter)

Model Name	Accuracy	MSE	MAE	RMSE	Macro F1 Score	Micro F1 Score	Weighted F1 score	Dummy classifier score
Linear regression(no grid)	0.19	2.94	1.40	1.18	0.13	0.19	0.18	0.23
Naive Bayes(no grid)	0.24	4.13	1.57	1.25	0.21	0.24	0.25	0.23
Random Forest	**0.37**	3.37	1.30	1.14	**0.24**	**0.37**	**0.30**	0.23
Decision Tree	0.33	4.07	1.47	1.21	0.22	0.33	0.28	0.23

linear regression improved a lot by using max feature 100

Table 3: after max feature=50, No PCA, random forest and decision tree used grid search CV(using example parameter)

Model Name	Accuracy	MSE	MAE	RMSE	Macro F1 Score	Micro F1 Score	Weighted F1 score	Dummy classifier score
Linear regression(no grid)	0.15	2.58	1.36	1.16	0.11	0.15	0.13	0.23
Naive Bayes(no grid)	0.25	4.11	1.56	1.25	0.21	0.25	0.26	0.23
Random Forest	**0.36**	3.85	1.41	1.19	**0.22**	**0.36**	**0.30**	
Decision Tree	0.33	3.67	1.38	1.18	0.19	0.33	0.25	

After reduce max feature from 100 to 50, accuracy and score of naive bayes increased, but accuracy and scores of linear regression decreased.

Table 4: after max feature=20, No PCA, random forest and decision tree used grid search CV(using example parameter)

Model Name	Accuracy	MSE	MAE	RMSE	Macro F1 Score	Micro F1 Score	Weighted F1 score	Dummy classifier score
Linear regression(no grid)	0.14	2.47	1.34	1.16	0.09	0.14	0.09	
Naive Bayes(no grid)	0.28	4.30	1.59	1.26	0.20	0.28	0.26	
Random Forest	0.32	3.78	1.43	1.20	**0.21**	0.32	**0.27**	
Decision Tree	**0.33**	4.31	1.50	1.22	0.18	**0.33**	0.26	

in conclusion, random forest has the best performance among 4 models, but when max feature is getting smaller, the decision tree's accuracy will approach to the random forest.