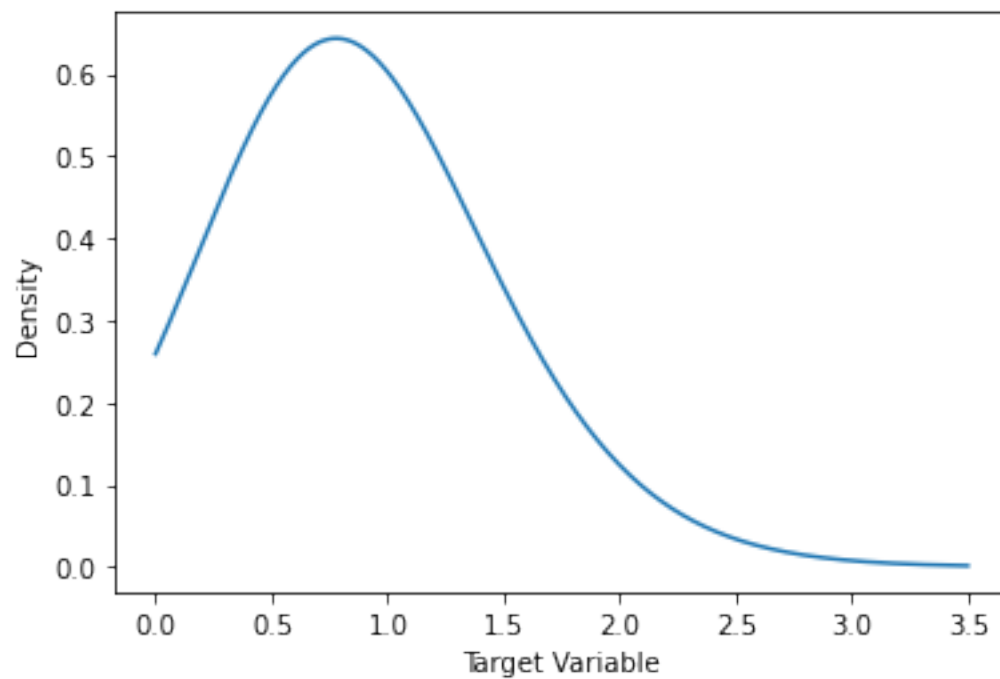


1 Data Exploration

1.1 Target Statistics

target	
count	8250.000000
mean	0.867212
std	0.405036
min	-0.000000
25%	0.600000
50%	0.800000
75%	1.100000
max	3.500000
range	3.500000

1.2 Kernel Density Estimate



1.3 Feature Statistics and Kernel Density Functions

	acc_rate	track	m	n	current_pitch \
count	8250.000000	8250.000000	8250.000000	8250.000000	8250.000000
mean	-11.526061	-12.672970	1.051280	0.060387	0.629371
std	259.637258	25.675733	0.320703	0.118805	0.313628
min	-975.000000	-95.000000	0.216536	-0.530000	-0.420000
25%	-175.000000	-29.000000	0.802519	-0.020000	0.430000
50%	-23.000000	-13.000000	1.020201	0.060000	0.610000
75%	157.000000	3.000000	1.258600	0.130000	0.800000
max	946.000000	94.000000	2.886371	0.620000	2.580000
range	1921.000000	189.000000	2.669835	1.150000	3.000000

	current_roll	absoluete_roll	climb_delta	roll_rate_delta \
count	8250.000000	8250.000000	8250.000000	8250.000000
mean	0.061248	-11.004848	-0.920364	-0.000957
std	0.967274	4.140399	10.334136	0.013203
min	-3.000000	-23.000000	-44.000000	-0.080000
25%	-0.600000	-14.000000	-8.000000	-0.009000
50%	0.200000	-11.000000	-1.000000	-0.001000
75%	0.700000	-7.000000	6.000000	0.008000
max	2.900000	-3.000000	46.000000	0.056000
range	5.900000	20.000000	90.000000	0.136000

	climb_delta_diff ...	time7_delta	time8_delta	time9_delta \
count	8250.000000 ...	8250.000000	8.250000e+03	8250.000000
mean	-0.047842 ...	-0.000018	-1.212121e-07	-0.000099
std	1.132179 ...	0.000113	1.100964e-05	0.000615
min	-8.800000 ...	-0.001000	-1.000000e-03	-0.006000
25%	-0.300000 ...	0.000000	0.000000e+00	0.000000
50%	-0.100000 ...	0.000000	0.000000e+00	0.000000
75%	0.200000 ...	0.000000	0.000000e+00	0.000000
max	8.700000 ...	0.000600	0.000000e+00	0.003000
range	17.500000 ...	0.001600	1.000000e-03	0.009000

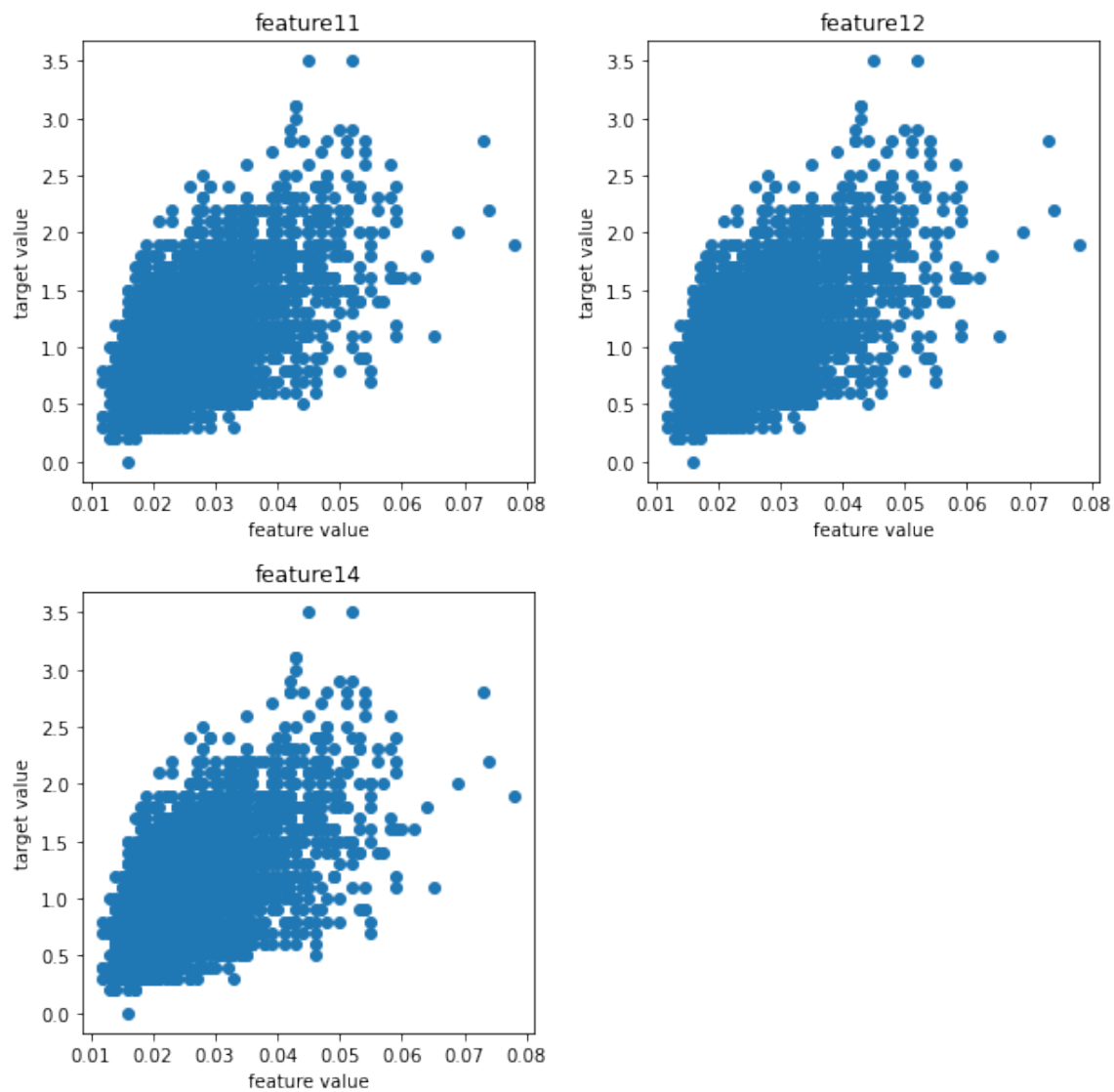
	time10_delta	time11_delta	time12_delta	time13_delta	time14_delta \
count	8.250000e+03	8250.000000	8.250000e+03	8250.000000	8250.000000
mean	4.848485e-07	8.999901	-7.272727e-07	-0.000093	-10.000001
std	6.963359e-05	0.000631	6.030154e-05	0.000613	0.000037
min	-2.000000e-03	8.995000	-4.000000e-03	-0.005000	-10.003000
25%	0.000000e+00	9.000000	0.000000e+00	0.000000	-10.000000
50%	0.000000e+00	9.000000	0.000000e+00	0.000000	-10.000000
75%	0.000000e+00	9.000000	0.000000e+00	0.000000	-10.000000
max	6.000000e-03	9.005000	2.000000e-03	0.007000	-10.000000
range	8.000000e-03	0.010000	6.000000e-03	0.012000	0.003000

	omega	set
count	8250.000000	8250.000000
mean	-0.510279	0.021982
std	0.257113	0.006961
min	-0.916291	0.012000
25%	-0.693147	0.018000
50%	-0.510826	0.020000
75%	-0.356675	0.023000
max	0.693147	0.073000
range	1.609438	0.061000

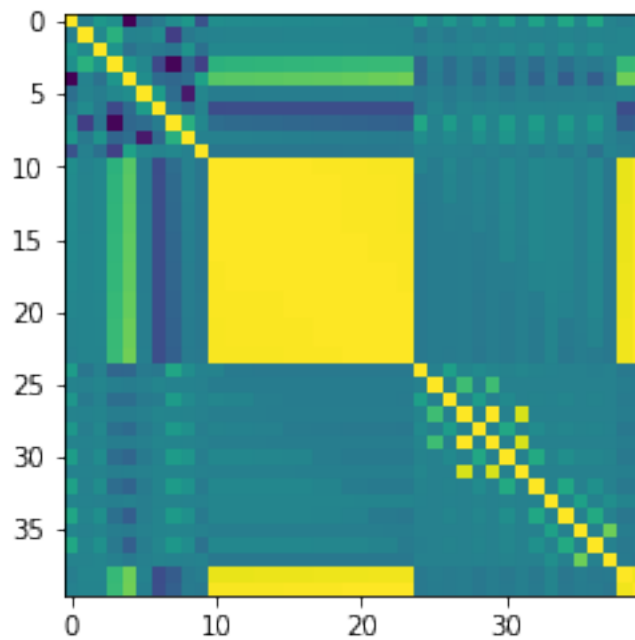
1.4 Most Correlated Features

	time2	time3	time5
count	8250.000000	8250.000000	8250.000000
mean	0.021886	0.021886	0.021898
std	0.006906	0.006906	0.006914
min	0.012000	0.012000	0.012000
25%	0.018000	0.018000	0.018000
50%	0.020000	0.020000	0.020000
75%	0.023000	0.023000	0.023000
max	0.078000	0.078000	0.078000

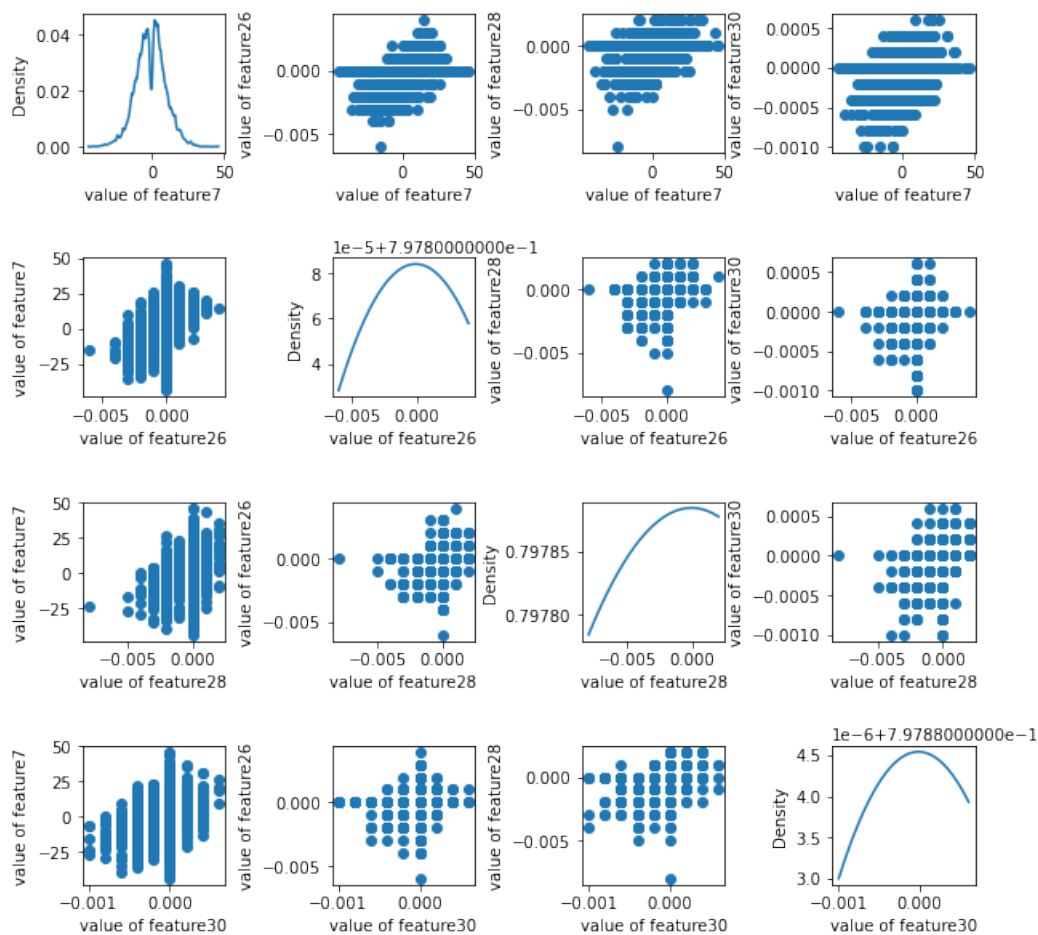
1.5 Feature Scatterplots



1.6 Feature Correlation Matrix



1.7 Feature Scatterplot Matrix



2 Baseline Models

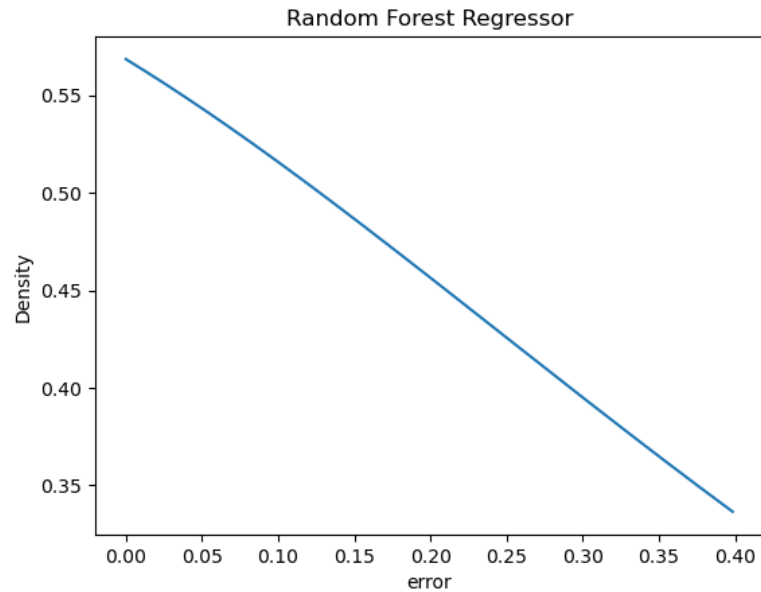
2.1 Sample Errors and Loss Functions

For the Random Forest Regressor model, the in-sample error is 0.0052674666666666665 and out-of-sample error is 0.029424278787878788. For the Linear regression model, the in-sample error is 0.029563662177267785 and out-of-sample error is 0.03018721456551126. For the k-Nearest Neighbors model, The in-sample error is 0.07571007575757575 and out-of-sample error is 0.1290675.

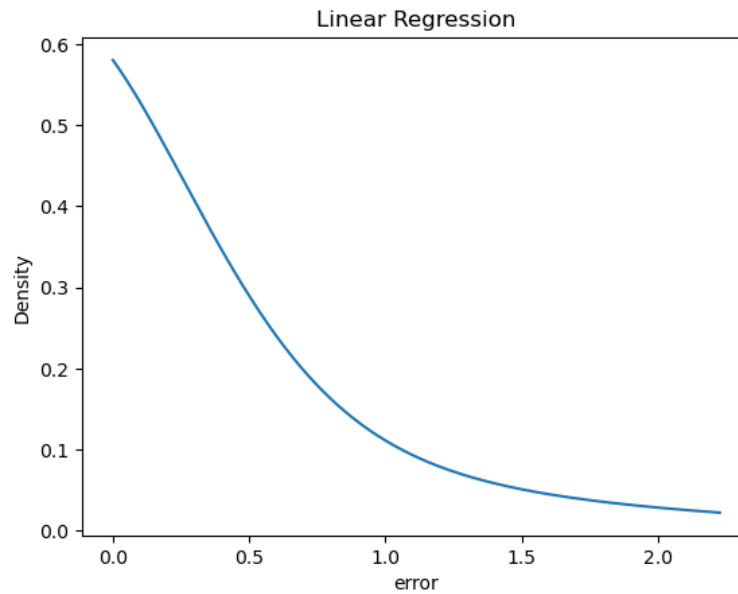
The MSE loss is differentiable everywhere, which works great with gradient descent optimization algorithms. The best parameters to use are the ones that minimize the loss function. MSE loss punishes large errors.

2.2 Kernel Densities

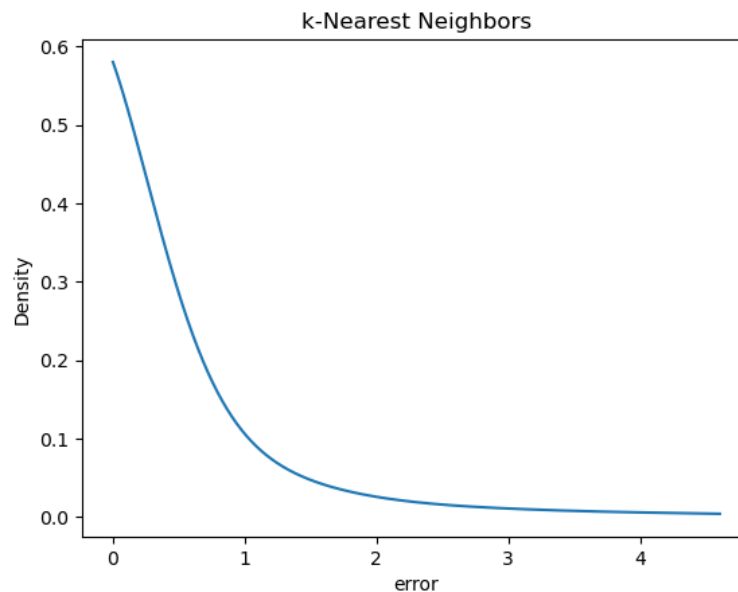
1. Random Forest Regressor



2. Linear Regression



3. k-Nearest Neighbors



2.3 In-Sample Error t-tests

Random Forest Regressor and k-Nearest Neighbors t-test pvalue=5.712866451086085e-18. Random Forest Regressor and Linear Regression t-test pvalue=2.616798141959661e-20. k-Nearest Neighbors and Linear Regression t-test pvalue=2.8921025392043456e-16.

2.4 Out-of-Sample Error t-tests

Random Forest Regressor and k-Nearest Neighbors t-test pvalue=1.872092206070558e-09. Random Forest Regressor and Linear Regression t-test pvalue=0.2850084620452252. k-Nearest Neighbors and Linear Regression t-test pvalue=1.3447716953672896e-09.

2.5 Performance Interpretations