

## MNIST PIPELINE

1. Data: Total\_data : (60000 , 42) entries

Every filtration except grayscale has been performed on binarized image with threshold 40%.

Filtration Type	Parameter	Vectorization	Column
Height	1,0	Persistent Entropy	1 (dim 0), 2 dim(1)
Height	0,1	Persistent Entropy	3,4
Height	1,1	Persistent Entropy	5,6
Height	-1,1	Persistent Entropy	7,8
Height	1,-1	Persistent Entropy	9,10
Height	-1,-1	Persistent Entropy	11,12
Height	-1,0	Persistent Entropy	13,14
Height	0,-1	Persistent Entropy	15,16
Radial	6,6	Persistent Entropy	17,18
Radial	13,6	Persistent Entropy	19,20
Radial	20,6	Persistent Entropy	21,22
Radial	20,13	Persistent Entropy	23,24
Radial	13,13	Persistent Entropy	25,26
Radial	6,13	Persistent Entropy	27,28
Radial	6,20	Persistent Entropy	29,30
Radial	13,20	Persistent Entropy	31,32
Radial	20,20	Persistent Entropy	33,34
Grayscale	NIL	Persistent Entropy	35,36
Density	2	Persistent Entropy	37,38
Density	4	Persistent Entropy	39,40
Density	6	Persistent Entropy	41,42

2. Data split into test and train: 50,000 and 10,000 rows respectively

Random Forest classifier with 1000 trees applied to train data

Confusion Matrix for prediction: Labels = [0,1,2,3,4,5,6,7,8,9]

	967	0	3	1	0	6	3	1	9	1
0	1042	13	0	1	0	0	7	0	1	
2	1	953	13	0	6	3	4	5	3	
0	0	15	995	0	7	0	7	0	6	
1	5	5	0	931	1	4	1	1	34	
4	0	7	10	5	870	8	0	7	4	
0	3	3	0	7	9	929	0	13	3	
0	3	10	6	1	5	0	1064	0	1	
3	2	6	5	3	10	16	0	947	17	
4	1	11	6	18	5	2	7	5	902	

