

train.csv, N_train = 40000, N_test = 8000

