SPLITTING THE DATASET INTO TRAINING AND TEST SET

