In this project we will experiment with random hyperplanes for classification. Your program will take a dataset as input and produce new features following the procedure below. The input is in the same format as for previous assignments.

Input
1. training dataset with labels (same format as project 1)
2. test dataset with labels that may be 0 (same format as project 1)
3. number of new features k

For i = 0 to k do:

a. Create random vector w where each w_j is uniformly sampled between -1 and 1.

b. Let x_j be our training data points. Determine the largest and smallest $w^T x_j$ across all x_j. Select w0 randomly between [smallest $w^T x_j$, largest $w^T x_j$].

c. Project training data X (each row is datapoint x_j) onto w.
Let projection vector z_i be $Xw + w0$ (here X has dimensions n by m and w is m by 1).
Append $(1+\text{sign}(z_i))/2$ as new column to the right end of Z. Remember that z_i is a vector and so $(1+\text{sign}(z_i))/2$ is 0 if the sign is -1 and 1 otherwise.

d. Project test data X' (each row is datapoint x_j) onto w.
Let projection vector z'_i be $X'w + w0$. Append $(1+\text{sign}(z'_i))/2$ as new column to the right end of Z'. We create the test data in exactly the same way as we do the training except that we do it on X' the test data instead of X the training data.

1. Run hinge loss on Z and predict on Z' after standardizing the data. Remember to standardize
using the column lengths from the training data only and not the test data.
2. Do values of k=10, 100, 1000, and 10000.
3. How does the error compare to hinge loss on original data X and X' for each k? Don't forget to standardize the data before applying hinge loss. This will improve the accuracy and speed of
your search.

Submit your assignment by copying it into the directory
/afs/cad/courses/ccs/f20/cs/675/101/<ucid>.
For example if your ucid is abc12 then copy your solution into
/afs/cad/courses/ccs/f20/cs/675/101/abc12.

Your completed assignment is due by midnight ber November 25th 2020