**Due: October 20, 2021, 12:00 PM**

# Problem 1 - SVM

Suppose we learn the optimal $w$ and $b$ for a dataset by linear SVM. The decision boundary is $w^T \cdot x + b = 0$. Now we apply a linear transformation to all $x$ using $x' = a \cdot x + c$, where $a \neq 0$ is a scalar.

- Can we use or modify the current $w$ and $b$ to perform well on this data? If yes, calculate the new decision boundary ($w'$ and $b'$)?
- What if $x' = A \cdot x + c$, where $A$ is an invertible matrix?
- What if $A$ is non-invertible? Explain why SVM could fail.

# Problem 2 - Ensemble Methods

- Given a training set, one way of creating an ensemble classifiers is to sample different subsets from the original dataset. Often less training instances can lead to unstable results and lower accuracy. Under what conditions do ensembles help improve the overall accuracy?
- In AdaBoost, why do we increase the weight of the misclassified instances? In each round, do we expect $\alpha$ to increase or decrease? Why?
- Can AdaBoost eventually achieve a 0 error rate if it runs for sufficient iterations?

# Problem 3 - Programming

- Download 'covtype.binary' (6.7M) from libsvm official site:
  https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/
- Load data, first column is binary labels, others are attributes as sparse matrix (category:value).
- Pre-proprocess the data by subtracting mean and scale them in [-1,1].
- Use Liblinear, apply SVM (primal), SVM (dual) (-s 0,1 and 2) and compare results and running time.
- Use Libsvm, apply kernel methods, see if the accuracy can beat linear SVMs.
- Visualize results in 2 dimensional space. (You may sample the data once it has too many points)
- Apply PCA before classification, see whether the results can be improved.