# DA5401 Assignment 5

We have already seen how LogisticRegression, DecisionTrees and kNN methods work towards classifying data points into their respective categories.

## Task 1: [40 points]

Let's consider the classification problem in https://archive.ics.uci.edu/dataset/76/nursery, which is a 8-features, 3-classes dataset. It is mentioned in the link that the expected performance of over 90% accuracy (See Baseline Model Performance). Let's add the following model performance outcomes to the baselines, shall we?

1. Decision Tree (categorical features)
2. Decision Tree (categorical features in one-hot encoded form)
3. Logistic Regression with L1 regularization
4. k-Nearest Neighbors

You are expected to split the data into train, val & test. Use the val partition to tune the hyperparameters such as (but not limited to) k of kNN, height of DT, or lambda of L1 reg. Remember, there are several other hyper parameters.

Report the performance of the test-data. Create a similar visualization with 9 methods now, with your additional 4 methods. The plot shows the mean and variance, FYI. Use a suitable visualization method to get them. You may wonder; to compute variance, you need more than 2 samples. Right. Repeat this task 5 times to get the mean and variance. :)

## Task 2: [10 points]

You may notice that the shape of logistic regression decision boundary and a sigmoid are a look-alike. We know that range of sigmoid is 0 to 1, which means, we can use sigmoid only when outputs are unipolar. Here are some simple extensions, we may try.

1. Construct a `bipolar_sigmoid(x)` using unipolar sigmoid.
2. A popular bipolar normalizer is `tanh(x)`. Compare the reponse of `tanh(x)` vs your `bipolar_sigmoid(x)`.
3. Parameterize it as `bipolar_sigmoid(ax), tanh(ax)`; You may plot the shapes of the response at different values of 'a' in [-5, -1, -.1, -.01, .001, .01, .1, 1, 5].
4. Now comes the interesting part. Can you evaluate the linear range of 'x' for each value of 'a' in `bipolar_sigmoid(ax)`? Usually, when 'a' is small, the linearity range is high.