# Assignment 2

**Notes:**
**This assignment to be solved in group of 4.**
**Team member can be from any group.**
**Problem 1**

## Sign Language to Speech:

### About

The data set is a collection of images of alphabets from the American Sign Language, separated in 29 folders which represent the various classes.

### Content

The training data set contains 87,000 images which are 200x200 pixels. There are 29 classes, of which 26 are for the letters A-Z and 3 classes for *SPACE*, *DELETE* and *NOTHING*.
These 3 classes are very helpful in real-time applications, and classification.
The test data set contains a mere 29 images, to encourage the use of real-world test images.

Dataset link : https://archive.ics.uci.edu/ml/datasets/congressional+voting+records

- Train 3 different classifiers to classify the 29 classes [You can use sklearn].
- Use different input for training (RGB , GREY Binary)
- Report Precision & recall for each experiment.

**Problem 2**

## Anomaly detection – Feature Selection

- Dataset:https://datasetsearch.research.google.com/search?query=Breast%20Cancer%20Dataset&docid=L2cvMTFqOWM3ejY5Yw%3D%3D

Features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image. n the 3-dimensional space is that described in: [K. P. Bennett and O. L. Mangasarian: "Robust Linear Programming Discrimination of Two Linearly Inseparable Sets", Optimization Methods and Software 1, 1992, 23-34].

Given the attached dataset of Breast cancer prediction, apply the following:

- Detailed-illustration in a report for the applied techniques during the **feature selection**,
  pre-processind.
- Evaluate on 3 different model and report precision - recall.