

## Steps for the implementation of Reduction\_Filteration\_GBPSO-SVM:

### Step 1:

- Removing the genes whose variance of expression value is greater than 0.15 and median ratio of both classes is between 0.95 and  $0.95^{-1}$  (using *conditional statements* in MATLAB)
- Removing the genes whose variance of expression value is equal or less than 0.15 and mean ratio of both classes is between 0.95 and  $0.95^{-1}$  (using *conditional statements* in MATLAB).

In this step, the dataset is reduced to one given as **S1\_Dataset.csv**

### Step 2:

- Divide the reduced dataset into the training (85%) and unseen dataset (15%) (using *cvpartition* in MATLAB).
- Employing the three filters (TT, COR and WRS) to rank the genes and select the 200 most discriminative ones accordingly (using *ttest2*, *ranksum* and *coding for COR equation* in MATLAB programming).

In this step, the datasets filtered by TT, COR and WRS are given as **S2\_Dataset.csv**, **S3\_Dataset.csv** and **S4\_Dataset.csv**, respectively.

### Step 3:

- Retrieving the source code of Geometric PSO from (<https://github.com/sebastian-luna-valero/PSOsearch/>) then interfacing it with MATLAB coding to create the GBPSO-SVM wrapper connected with previous steps.

In this step, the datasets selected by GBPSO-SVM (TT branch), GBPSO-SVM (COR branch) and GBPSO-SVM (WRS branch) and Merged GBPSO-SVM Branches are given as **S5\_Dataset.csv**, **S6\_Dataset.csv**, **S7\_Dataset.csv** and **S8\_Dataset.csv**, respectively.

**Non-involved Dataset.csv** file represents the unseen dataset.