Guidance

☐ Both `CRC codes` and `IBD codes` are the `same` except for input datasets.

☐ **CRC Files**:
- ○ `CRC_FS.ipynb` → performs feature selection and creates a file containing selected features
  - ☐ there are two parameters (in code) that control feature selection:
    - ○ **N** = number of returned genes by SVFS
    - ○ **feature_selection_repetition** = how many times to run SVFS

- ○ `CRC_Classification.ipynb` → performs classification using features generated by CRC_FS.ipynb
  - ☐ there are two parameters (in code) that control plotting:
    - ○ **_USE_FIRST_N_GENES_CONF** = how many genes to use for plotting confusion matrix
    - ○ **_USE_FIRST_N_GENES_ALL** = how many genes to use for plotting accuracy/f1-score plots

- ○ `CRC_Visualizer` → Visualizes the number of each gene repetition (should be run after performing feature selection)

- ○ `CRC_features_backup.txt` → Backup of features used for paper (To get paper results, change its name to CRC_features.txt)

☐ All picures are saved in **Pictures** folder (there is one picture folder in CRC folder and one in IBD folder)

☐ CommonCodes folder:
- ○ `HelperFunctions.py` → contains functions that are being used in other files.
- ○ reduction.py → Used by SVFS
- ○ mutual_info.py → Used by SVFS