Final project, part A

Instructions

Download the data from Moodle.

Comment the code (except self-explanatory assignments e.g. MEAN=SUM/SIZE ☺).

All codes, saved models and the report must be handed in Sun 12th Feb, 2023the latest.

Problem description

In the final project you will analyze data from the image segmentation domain. The database of feature vectors was created in the Vision Group of University of Massachusetts.

The full description of the data is in `segmentation_description.txt`.

Task 1 Load the data `segmentation_data.mat`. This should create 3 variables in your workspace: `feature_names`, `segmentation_features` and `segmentation_labels`.

Task 2 Analyze the features. Identify features that are necessary. Lower the dimensionality of the data.

Task 3 Pick three (3) supervised classification methods. Use cross-validation to set the parameters of the models. Then train the classifiers on the full dataset. Save the trained models.

Task 4 Use an unsupervised method to identify clusters in the data. Find the optimal number of clusters.

Task 5 Write a report (in pdf format) where you

- describe what you did in each previous task,
- justify the selection of the methods,
- report the parameters and results and
- put down your comments (e.g. what went wrong, whether the classifiers behave as expected, etc.).

You can format the report as a conference paper.

Task 6 Prepare a wrapper function. The function expects the parameters `test_features` and `test_labels` in the same format as the original data. The function than performs the feature selection/transformation, loads your trained models and classifies the data. For each method it will report the confusion matrix and the macro-averaged precision and recall values. Plot the results for the three models into a precision-recall space.

Final project, part A

Help

Labels are stored in a categorical array. Built-in Matlab classifiers can handle this type. If you code your own classifier that outputs an index of a category in a variable `category_index`, you can use the following code to check whether the the categories match.

```
train_categories=categories(train_labels);
test_labels==train_categories(category_index)
```