

Pattern Recognition (Semester Two, 2011-2012)

Course Project

Date: March, 2012

The purpose of the project is to enable the student to get hands-on experience in the design, implementation and evaluation of pattern recognition algorithms.

Project will involve using at least 5 classifiers and evaluating their performances on at least a common pattern classification task. The objective is to use preprocessing, feature extraction/selection, classifier design and evaluation to get as high an accuracy as possible on the selected classification task.

You can select following tasks for your project. If you have a better topic and dataset, you can select another task after discussing with Dr. Chan.

1. Spambase

<http://archive.ics.uci.edu/ml/datasets/Spambase>

2. Blood Transfusion Service Center

<http://archive.ics.uci.edu/ml/datasets/Blood+Transfusion+Service+Center>

3. Breast Cancer Wisconsin (Diagnostic) Data Set

<http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29>

4. Yeast Data Set

<http://archive.ics.uci.edu/ml/datasets/Yeast>

5. Glass Identification Data Set

<http://archive.ics.uci.edu/ml/datasets/Glass+Identification>

6. The ORL Database of Faces (difficult)

<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

7. PIE Database (very difficult)

http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261

Emphasis of the project is on solving a practical and interesting pattern recognition problem using the tools that you have learnt in this course and comparing your results to the state-of-the-art performance. You should implement the project by using Matlab program with PRtools package (<http://www.prtools.org/>).

Group Size

A group should contain FIVE students. A group leader should also be selected. You should send an email to Mr. Zhimin He (531981194@qq.com) with the **1) Student ID**, **2) Class**, **3) Name** and **4) mobile number (group leader only)** of each member in your group by 6-April-2011. The subject title of the email should be “**PR Course: Project Group**”.

Grading

Report	75%
Presentation	25%

Report

The project report should clearly explain each detail procedure. The report should include the following information:

- **Table of contents**
- **Introduction**
The objective of the project and the background work on the problem you select.
- **Pre-processing**
A discussion of your pre-processing methods (feature selection, sample selection, normalization, etc) which include a rationale for your choice.
- **Classification Method**
Discussion of the classifiers you use and explain why you think they are appropriate for the task.
- **Experimental Result**
Explain the experimental setting. Discuss and compare classifiers' performances and evaluate the results.
- **Conclusion**
The advantages and disadvantages of each classifier and how you can improve your system.

Each student should write an individual report about what he/she has learned and the challenges about this project. Your individual report should be attached to the end of the project report.

Presentation

Each group will be allocated 10 minutes for oral presentation in English on *9-June-2012*.

Report Submission

The following material should be submitted by *9-June-2012*:

1. A ZIP file which contains a softcopy of your report in WORD format and all program matlab codes should be sent to Mr. Zhimin He (531981194@qq.com). The subject title of the email should be “**PR Course: Project (Group XX)**”. You will be informed your group name later.
2. A hardcopy of your report should be submitted to before your presentation

If you submit the work after the deadline you will incur a deduction of 5% of the full mark for each day of delay.