Machine Learning (521289S) Programming Assignment

Deadline: 30.3.2019 16.00

Physiological Signal Analysis Team
Center for Machine Vision and Signal Analysis (CMVS)
University of Oulu

Programming Assignment

- Students will make a MATLAB program to solve a pattern classification task
  - Training and your own testing with given data
  - Evaluation with unknown (similar) data on a ranking server

- The classification accuracy of your program will determine your grade related to the programming assignment
  - Course grade is then calculated by 2/3 * (exam grade) + 1/3 * (programming assignment grade)

- Work should be done solo
  - Plagiarism incidents will be handled according to university guidelines

- **Deadline**: 30.3.2019 16.00

Task

- Your task is to implement and test a classifier on the given data using MATLAB.
- In your solution, you should use more than one method trying to find the best possible solution for the case!
  - This means that your solution has to consist of several steps and each one of these consist at least one method (for example one step for pre-processing, one for feature extraction and one for classification).
  - Using more than one method does not mean that you should make several solutions (each one consisting of one method).

Material

- The learning data used in this task are available on the course Noppa page.
  - You must use the downloaded m-file template as a basis of your code, because your code has to be compatible with our testing interface.
  - You must include all your code within the single m-file based on the template.

- In addition, download the “Template” found on the same page. Rename it to **classify.m**.
  - The learning data, with 8 features, is given in the matrix **trainingData**.
  - The correct class (i.e. the training target) is given by the value 0, 1 or 2 in the vector **class_trainingData**.

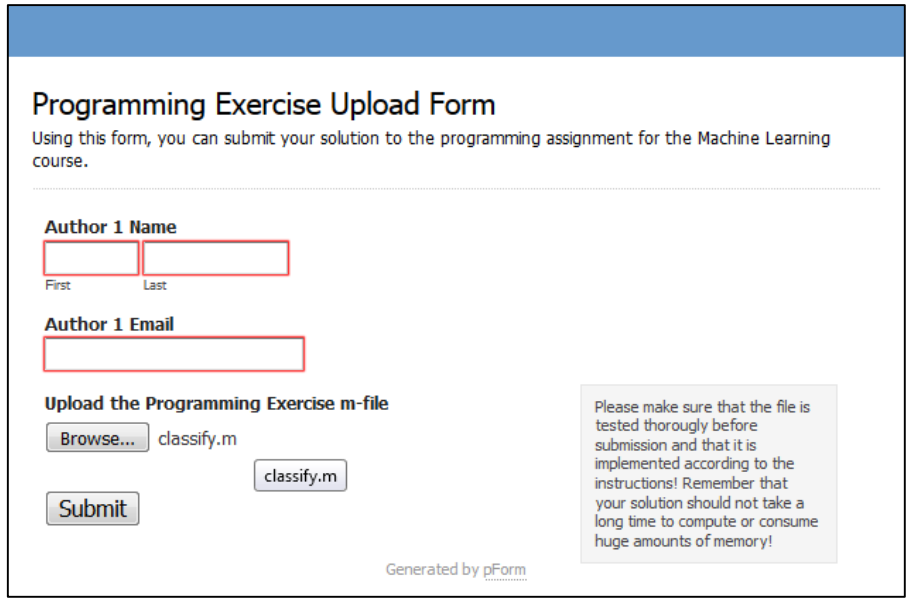
Template

- You must include all the code within the single m-file based on the template.
- The template defines the required interface consisting of a nick name function (getNickName), a training subfunction (trainClassifier) and a classifier evaluation subfunction (evaluateClassifier).
  - You must implement the code for these functions.
  - The number and names of the input and output variables of these functions are fixed and should not be altered in any way.
  - You should not alter the names of the functions themselves, either.
  - The input data is expected to be in the format of the downloaded training data.
- You are free to choose the data format and contents of the output (parameters) of the training subfunction, used to carry the parameters learned from the training data to the classifier.

Uploading your m-file

- The performance of your code is evaluated by submitting the code to our test server using the link provided by the assistant.
- **You will need an account name and password for uploading** (contact Antti Isosalo, antti.Isosalo@oulu.fi)**.**
- You can get them by sending an email to the assistant.
- **You also have to have a nickname** which you can freely choose (use only one throughout the course).
  - The nickname **is used on the ranking page to identify your results** among others.
- **Before** you submit your script to the test server, make sure that your program is properly working in your own machine.

- There can be several students testing their scripts at the same time, so one hour is the time limit for your scripts to run



Results

- As a result from submitting you will get an email containing your solution's performance reading (%).
  - Note, that the email notification gets easily classified as spam
- After getting the result, you can choose to make your solution better and submit a new solution as many times as you wish.
  - The classification accuracy given by our testing server to your work should be at least 67% for you to get the partial credit
- There is a "Ranking list", which shows the best performance for each nickname.

Current rankings:

Ranking	Nickname	Accuracy	Submissions
1.	MrAntti	86.36884 %	1
2.	RimanYli	85.00000 %	1
3.	RimanAli	66.99999 %	1
4.	TestUser	33.17897 %	3

Deliverables

- Create a MATLAB m-file based on the template.
- Remember to clean and properly comment your code.
- Write a short (~1-page) report of your work.
- Remember to include all the computed values to it, especially your solution's classification accuracy 1) from your own testing and 2) from our testing server to the report.
- In the report, explain what methods you have chosen and why you have chosen those methods.
  - If you have used a method not included in the lectures of the course, you must also explain the method more thoroughly in the report.

Deliverables (cont.)

- Attach the code as an appendix to the report.
- Remember to also include your student number and the nickname you used to the report.
- **Email** the **report (in PDF format) and completed code (as m-file)** to the assistant (Anja Keskinarkaus, anja.keskinarkaus@oulu.fi).
- **Include the code also an appendix to the report** (PDF format).
  - **The assistant will inform you whether your programming exercise is accepted or if it still needs more work.**

Grading

- The work is graded according to its performance (should be > 67%), and the result will form one-third of the course grade.

Classification accuracy [%] as reported by the server	Grade
85.00000 -	5
82.00000 - 84.99999	4
78.00000 – 81.99999	3
73.00000 – 77.99999	2
67.00000 – 72.99999	1
below 67	Fail