# **Assignment 2 (CA2: 40%)**

The objective of the assignment is to help you build a conversational agent or chatbot that is able to perform some useful tasks.

## **Guidelines**

1. You are to work on the problem set in **groups of 1-3**.
   1. Please include Appendix I to detail out the contributions by respective students.
   2. Appendix I is not necessary if you are doing the project solo.
2. Submit your code and in a compressed package (zip file).
3. Students are required to submit their assignment using the assignment link under the Assignment folder.
4. The normal SP’s academic policies on Copyright and Plagiarism applies. Please note that you are to cite all sources. You may refer to the citation guide available at: <http://eliser.lib.sp.edu.sg/elsr_website/Html/citation.pdf>

## Submission Details

Deadline: 15 Feb, 2019, 11:59pm

Submit through: Blackboard

## Late Submission

50% of the marks will be deducted for assignments that are received within ONE (1) calendar day after the submission deadline. No marks will be given thereafter.

Exceptions to this policy will be given to students with valid LOA on medical or compassionate grounds. Students in such cases will need to inform the lecturer as soon as reasonably possible. Students are not to assume on their own that their deadline has been extended.

# **PART A: UNSUPERVISED LEARNING (40 marks)**

## Background

1. Given the iris dataset, if we knew that there were k types of iris, but did not have access to a taxonomist to label them: we could try a clustering task: split the observations into well-separated group called clusters.

## Dataset

Use the iris dataset from scikit-learn

## Tasks

1. Write the code to solve the clustering task. Normally you would be using scikit-learn, but if you'd prefer to work with your own implementation of learning algorithms, or some other toolkit, that is fine.
2. Write a short report (1-2 pages) detailing your implementation, your experiments and analysis.
3. Test your clustering with different possible values of k
4. Determine the best possible value of k. And show how you are able to determine that this is the best value for k.

# **PART B: DEEP LEARNING (60 marks)**

## Background

Implement an image classifier using a deep learning network. You may use any of the well-known Deep Learning (DL) network models such as AlexNet, ResNet, etc.

## Dataset

You are to use the MNIST dataset.

## Tasks

1. Write the code to solve the prediction task. Normally you would be using TensorFlow/Keras, but if you'd prefer to work with your own implementation of learning algorithms, or some other toolkit, that is fine.
2. Write a short report (1-3 pages) detailing your implementation, your experiments and analysis. In particular, we'd like to know:

* How is your prediction task defined? And what is the meaning of the output variable?
* How do you represent your data as features?
* Did you process the features in any way?
* Did you bring in any additional sources of data?
* How did you select which DL model to use?
* Did you try to tune the hyperparameters of the learning algorithm, and in that case how?
* How do you evaluate the quality of your system?
* Can you say anything about the errors that the system makes? For a classification task, you may consider a confusion matrix.
* Is it possible to say something about which features the model considers important? (Whether this is possible depends on the type of classifier you are using)
* Provide a reference section for any papers, online articles, books, publications that you have referenced.

## Submission requirements

1. Submit a zip file containing all the project files (python files or notebook), all data sets used, and the report.
2. Submit online via the Assignment link.

# **PART C: CHATBOT Mini Project (Bonus — 20 marks)**

**This part is optional and is not a requirement**

## Background

Build a chatbot to perform some useful functions such as providing information or recommendations.

## Dataset

You are to choose your own data sets or use existing data sets.

## Model

You can develop a new model for the data set or use an existing model.

## Tasks

1. Write the code for the chatbot.
2. Integrate the chatbot with the data sets & model

## Submission requirements

1. Submit a zip file containing all the project files (all source files — e.g. html, js, css, py files), all data sets used, and mini project summary (in powerpoint format with all the main screens captured and annotated).

— *End of Assignment* —

**APPENDIX I**

Summary table of contributions (to be included in your report submission in an appendix)

|  |  |  |  |
| --- | --- | --- | --- |
| Unsupervised Learning | | | |
| Student Name/ID | Item | Description | Reference/Comments |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  | Other | Any other contributions |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Deep Learning | | | |
| Student Name/ID | Item | Description | Reference/Comments |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  | Other | Any other contributions. |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Overall compilation of report  (Please include name/id of students who are doing this) | | | |
| Student Name/ID | Evaluation item | Description | Comments |
|  | Report | [a section of report, or whole report] |  |
|  |  |  |  |
|  | Other |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Bonus: Chatbot Mini Project | | | |
| Student Name/ID | Item | Description | Reference/Comments |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  | Other | Any other contributions. |  |