# Main Assignment

Re-submit Assignment

---

**Due**   11 Dec 2020 by 17:00     **Points**   10     **Submitting**   a file upload
**File types**   pdf and docx     **Available**   after 9 Nov 2020 at 9:00

---

In this assignment, you will compare and evaluate different classification methods. The dataset you will be working with is the `fashion-mnist` dataset. This dataset consists of a training set of 60,000 examples and a test set of 10,000 examples. This dataset is designed to be significantly more challenging than the MNIST dataset but is structured in the same way so can be used as a drop-in replacement.



*Some examples from the fashion MNIST dataset. Each class takes up three rows. Image linked from* **https://github.com/zalandoresearch/fashion-mnist/blob/master/doc/img/fashion-mnist-sprite.png**

---

**The dataset, and a description of how to use it can be found here:**

**https://github.com/zalandoresearch/fashion-mnist**

---

Your task is as follows:

1. Select three classification algorithms to evaluate on this dataset.
2. Perform a systematic optimisation of the hyperparameters of your chosen algorithms.
3. Evaluate the performance of your optimised algorithms.
4. Analyse your results to understand the behaviour of your chosen algorithms.

You can use any software platform or library to complete this assignment. Options you may consider include:

- Python with **scikit-learn**
- **Matlab**
- Java with **WEKA, or via the WEKA GUI.**
- Your own implementation (but beware of spending too much time on this).

**It is expected that you should spend no more than 10-12 hours working on this assignment.**

You should submit one report to the following specifications:

- A4 paper, 2cm margins, Arial font, 11pt minimum font size, single spaced.
- No more than 2 pages in length. **Reports of longer than two pages will not be marked.**
- Add only your student ID to the document header.

Here is a report template in Word format: **report-template.docx** ↓

The report should contain the following sections:

# Algorithm Selection

- State the three classification algorithms you have chosen.
- Write a brief justification for why you have chosen them.
- For each algorithm, briefly describe the hyperparameters that you will explore and why. *You do not have to explore all the hyperparameters; select the ones that you think will be the most important.*

# Methodology

- Explain your training procedure, including any further subdivision of the data.
- Describe how you have optimised the hyperparameters.
- Explain how you chose to measure and test algorithm performance.

# Results

- Briefly summarise the results you have obtained.
- You may include graphs or tables, but these must be included within the page count.
- Identify and explain any notable features of the results. For example, are there any classes that are frequently confused?

# Conclusion

- Analyse and explain your results, noting, for example, any particular sensitivity or insensitivity of an algorithm to the value of a particular hyperparameter.
- Suggest why you think your best performing algorithm was superior to your other choices.

Your submission will be graded according to the rigour of the methodology that you have followed, and the analysis of your outcomes. The performance of the models is a secondary criterion.

You do not need to submit your code. The assignment will be graded based only on the report.

---

**This assignment is worth 25% of your mark for the module.**

**Late submissions will not be accepted unless approved by welfare.**

---

## Main Assignment

| Criteria | Ratings | Pts |
|---|---|---|
| Algorithm Selection<br><br>* State the three classification algorithms you have chosen.<br>* Write a brief justification for why you have chosen them.<br><br>* For each algorithm, briefly describe the hyperparameters that you will explore and why. You do not have to explore all the hyperparameters; select the ones that you think will be the most important. | | 3 pts |
| Methodology<br><br>* Explain your training procedure, including any further subdivision of the data.<br>* Describe how you have optimised the hyperparameters.<br>* Explain how you chose to measure and test algorithm performance. | | 2 pts |
| Presentation and Analysis of Results<br><br>* Briefly summarise the results you have obtained.<br>* You may include graphs or tables, but these must be included within the page count. * Identify and explain any notable features of the results. For example, are there any classes that are frequently confused? | | 3 pts |
| Conclusions<br><br>* Analyse and explain your results, noting, for example, any particular sensitivity or insensitivity of an algorithm to the value of a particular hyperparameter.<br>* Suggest why you think your best performing algorithm was superior to your other choices. | | 2 pts |
| | | Total points: 10 |