Queen Mary
University of London

MTH786U & MTH786P Machine Learning with Python, Semester C, 2020/21

Final assessment **M. Poplavskyi, M.Benning**

---

1 Recognition of fashion items

The goal of this final coding assessment is to implement, describe and present (multi-class) classification models of your choice for the recognition of fashion items. More precisely, the goal is to implement (multi-class) classifiers that take an image of size 28×28 of a fashion item from one of ten classes as their input and automatically determine which class the input image most likely belongs to. The dataset to be used is the so-called Fashion MNIST dataset and will be provided on the QM+ module page.

This assessment is formed of three parts: 1) For the first part you are required to fill in the missing parts of the provided Jupyter notebook template. 2) You then apply learned concepts from the first part and the module MTH786 in general to the Fashion MNIST classification problem. 3) Last but not least, you present your results in a short, written report (written in LaTeX) of no more than 8 pages length.

1.1 Complete the Jupyter notebook

Download the corresponding Jupyter notebook template and follow the instructions therein. Make sure you fill in any place that says "YOUR CODE HERE". Test your implementations with the provided test cases. Different tests are awarded different amounts of marks. **The total number of marks for this part of the assessment is 40/100**. Note that this part of the assessment will be graded automatically with the *nbgrader* Python package.

1.2 Implement & validate classification strategies for Fashion MNIST

Once you have completed the first part of the assessment (or have completed as much as you were able to complete), you can use any concept that you have encountered during the first part of the assessment and throughout this module, in order to obtain a classifier that can determine which fashion item is seen in a 28×28-pixel picture. While you are allowed to use different libraries for visualisation purposes, you can only use NumPy to implement your classifier; tools from libraries such as SciPy are not allowed.

Experimenting with different model approaches, regularisation models & parameters as well as hyperparameter-tuning strategies such as cross validation is recommended. This part will be graded together with the next part. Please feel free to use the provided template notebook for this freestyle part.

1.3 Write a report on the problem formulation and your results

You conclude the assessment by writing a report about the problem description and your findings. The report should be no longer than eight pages and be written preferably in LaTeXwith your favourite editor. If no editor is at hand, please feel free to use online editors such as Overleaf. This part is marked together with the previous part and **the total number of marks for both parts is 60/100**. We refer to the following table for marking guidance on the report and the previous coding part.

Approximate weight	**Description**
60 %	Quality of the project work itself (as described in the report), taking into account its difficulty. This includes the methodology and choice of resources, the application of these resources, explanation/exposition and correctness of mathematical notation and computer programming and analysis of results.
30 %	Quality of presentation: Logical structure of the document, clarity and coherence of exposition, correct use of English (punctuation, spelling and grammar) and precise mathematical writing, layout and style, sensible use of sections and subsections, lack of typing mistakes, choice of sensible (and standard) notation and other conventions, sensible use of equation numbering, appropriate use of figures, tables, charts, diagrams, etc.
10 %	Initiative and ambition (as described in the report): How far did you go, or attempt to go? Did you use only tools from the first part of the assessment, or did you use other methodologies from the lecture notes?