UNIVERSITY OF
LINCOLN

Lincoln School of Computer Science

Assessment Item Briefing Document

Title: CMP9137M - Machine Learning Assessment Item 1 (assignment)	Indicative Weighting: 50%

Learning Outcomes:

On successful completion of this assessment item a student will have demonstrated competence in the following areas:

- [LO3] Use machine learning software to solve complex real-world problems in an application domain of interest.

Requirements

This assessment comprises of one task on machine learning as explained below. Your submission should include a concise report (maximum 5 pages, neither including the cover sheet nor the references or appendixes) that describes your work on the above task. The submission should include your source code as an appendix.

The Task:

In this assignment, you are required to use Machine Learning to solve the problem of "***Recognizing handwritten digits***". The MNIST database of handwritten digits with their labels, available from this page (http://yann.lecun.com/exdb/mnist/), has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger set available from NIST. The original black and white (bilevel) images from NIST were size normalized to fit in a 20x20 pixel box while preserving their aspect ratio. The resulting images contain grey levels as a result of the anti-aliasing technique used by the normalization algorithm. The images were centered in a 28x28 image by computing the center of mass of the pixels, and translating the image so as to position this point at the center of the 28x28 field.

You are expected to explore a range of machine learning classifiers, inspired by the various models and categories explored within the module and beyond (i.e. from reading and literature). You will then investigate the performance of those classifiers, compare and critique them to justify your recommended classifier(s). This should include measures/metrics such as TP/FP rates, Precision-Recall, F-measure, and any other relevant metrics.

From the MNIST dataset, three datasets of different size (5000, 1000, 500 samples) will be extracted by the delivery team and will be made available to you from Blackboard for training, validation and testing. As part of your assignment, you will need to use the "Test" dataset to only to test your trained classifier. In addition to reporting test results, your report should include results of the training and validation phases – for all chosen classifiers. You will also need to discuss and justify those results in detail.

In this assignment you are free to train any classifier you find in Python, and to do any pre-processing of the data you want. You are free to implement your own algorithm(s) instead of only using libraries. You will need to clearly mention your resources, acknowledge appropriately, and compare between classifiers and their results in your report.

You have to prepare a technical report indicating, discussing and justifying all the steps and strategies you followed to obtain your classifier, and why you think they were right to arrive to your best classifier. You will need to compare between, at least, THREE different classifiers in your assignment, discussing and justifying any assumptions you may had, why the results are similar/different, ...,etc., to demonstrate your understanding of Machine Learning techniques. The report (in PDF format) and the source code of your classifiers should be submitted to Blackboard. Note that your source code should be included as an appendix within your submitted PDF file. Please do not submit your source code as supporting material.

Useful Information

This assessment is an individual assessment component. Your work must be presented according to the Lincoln School of Computer Science guidelines for the presentation of assessed written work.
Please make sure you have a clear understanding of the grading principles for this component as detailed in the accompanying Criterion Reference Grid.
If you are unsure about any aspect of this assessment component, please seek the advice of a member of the delivery team.

Submission Instructions

The deadline for submission of this work is included in the School Submission dates on Blackboard.

You must make a single electronic submission with your report in PDF format (only), which must be uploaded to the assignment submission area on Blackboard for this component.

DO NOT include this briefing document with your submission.