



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

COS 314: Artificial Intelligence

Assignment 3: Neural Networks for Image Classification

Total: 50

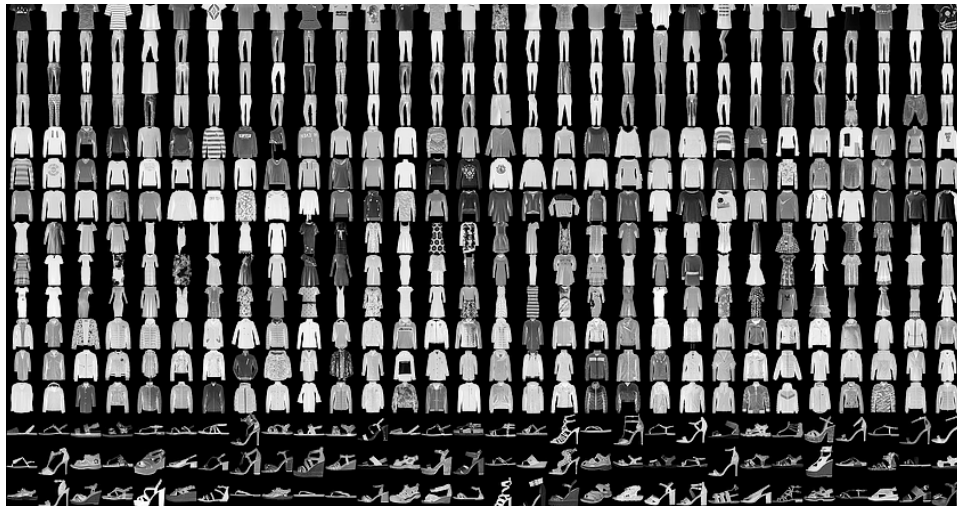
Due Date: 22 May 2019 23:59

1 Constraints

For this assignment you may **only** use C++ or Java. Furthermore, you may **not** use external libraries (built in ones are fine). You must demo your submission in the Thursday practical session on the 23rd May.

2 General Task

One of the most popular applications of neural networks is image classification. You will need to build a simple neural network model to classify images from the MNIST-Fashion data set



Fashion-MNIST is a dataset of Zalando's article images consisting of a training set of 60000 examples and a test set of 10000 examples. Each example is a 28×28 grayscale image, associated with a label from 10 classes. For ease of use the training set and test set have been uploaded in CSV format, with each example stored per row. The row contains the example label (0 to 10) and 28^2 pixel values. You should scale these pixel values to $[0, 1]$ as a preprocessing step in your code (don't alter the files).

3 Model Details

For this task you must build a neural network model as follows:

- Have 2-hidden layers using the **ReLU** activation function
 - Hidden layer 1 must have 100 units (the layer closer to the input layer)
 - Hidden layer 2 must have 50 units (the layer closer to the output layer)
- Have the output layer use the **softmax** activation function
- Use the **negative log likelihood** cost function.
- The weights must be optimized using mini-batch gradient descent.
 - Mini-batch size of 50
 - You will need to determine a good learning rate. (You will need to motivate this choice)
 - You will need to determine a good stopping condition (You will need to motivate this choice)

4 Submission

You will need to submit the code for two programs.

- The first of which to demonstrate the training of your neural network, which displays the error after each epoch.
- The second of which must be able to load a csv file, containing unseen patterns and classify them. This program must use your **pre-trained** neural network.

You will be asked questions about your model during the demo.