



Reminder from Intro Presentation: Module Assessment

- Final exam (60%)
 - Sample questions will be provided closer to the exam time
- Continuous assessment (40%)
 - ➡ Assignment 1: 20% (due at end of Week 6)
 - Assignment 2: 20% (due at end of Week 11)
 - Details and specific deadlines will be posted on Blackboard
 - The two assignments will be substantial multi-week programming assignments.



Deep Learning Assignment 1 2021

- In this assignment, you will implement a deep neural network from scratch
 - Multi-part assignment, drawing on material covered each week
 - You are expected to start work on it immediately; it is NOT recommended to wait until near the deadline
- You must use a Jupyter Notebook:
 - Python recommended but R is OK too
 - You can use low-level functionality of numpy and scikit-learn, but **not** their implementations of anything core (e.g. optimisers, grad descent etc.)



1-Person or 2-Person Assignment: You Choose

- You can do the assignment on your own or you can join with 1 other person (no more)
 - Final part will involve separate contributions by both people
- Policy on two-person submissions:
 - You are encouraged to work in pairs for Parts 1-4, but you cannot collaborate with other people in the class beyond that.
 - Part 5 must be done twice for a 2-person assignment.
 - Your documentation and code must be clear about what contributions each person made; it is **not** sufficient to say something like “we both worked on everything together”.
 - Across all parts of the assignment, I may award different marks to each person doing a 2-person assignment if I feel that their contributions are not equal.



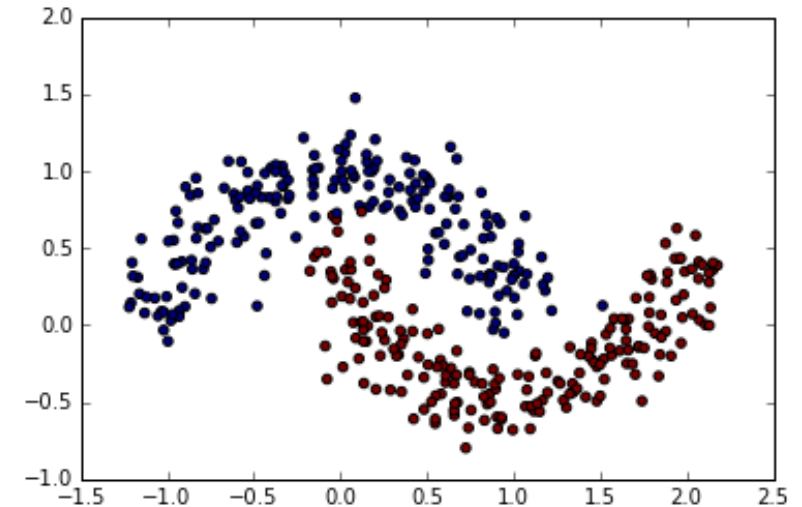
Part 1: Implement Logistic Regression

- Logistic Regression
 - Use Jupyter Notebook (Python or R) to implement a neural network approach to logistic regression (no hidden layers, one output node)
 - Work in pairs: note contributions of both people (larger groups not allowed)
 - Your code must handle different numbers of inputs and different numbers of training cases, but you don't have to support more than one output node
- Recommended to follow my notes closely
 - Post questions on discussion forum if there are gaps/inconsistencies, or if you are unclear about anything
 - No plagiarism! ANY plagiarism will result in 0 for full assignment
 - Reference ALL sources you used
 - **Do not submit code taken from the web**



Part 2: Easy Tasks

- I will supply two fairly small datasets:
 - One will be linearly separable (almost or fully), the other will not
 - I will provide sample Python code to load and plot the datasets; you are allowed to use this code in your own assignment
- Divide each dataset randomly into:
 - Training set (70%): use for main training
 - Validation set (15%): use for tuning, e.g. selecting learning rates
 - Test set (15%): held out set for final performance evaluation
- Train a logistic regressor using your code from Part 1, and see how it performs on both datasets





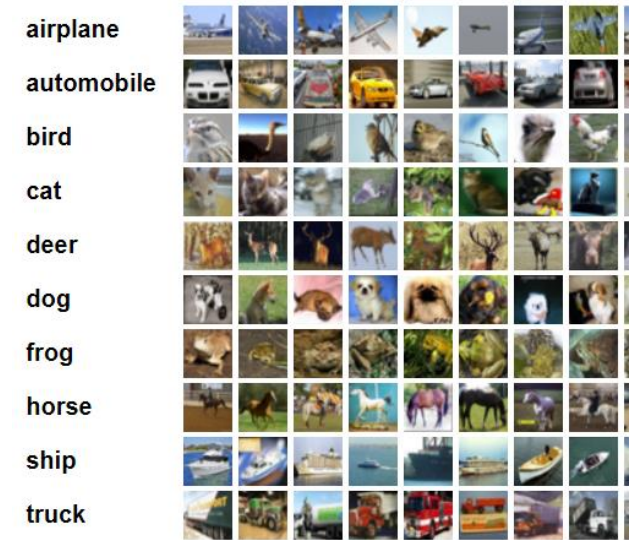
Part 3: Implement and Test a Shallow Neural Network

- Implement a shallow neural network
 - Start with your logistic regression code
 - Again, follow my notes as closely as possible
 - Your code must support different numbers of input nodes and different numbers of nodes in the hidden layer; you don't have to support more than one output node
 - No need to support more than 1 hidden layer
- Test your implementation on the two small datasets:
 - Is it able to handle the linearly separable data? (I hope so)
 - Does it perform better than the logistic regression implementation on the dataset that is not linearly separable?



Part 4: Challenging Task

- Image recognition is one task now tackled with ML, that was considered “very hard” not long ago
 - We will use the CIFAR-10 dataset
<https://www.cs.toronto.edu/~kriz/cifar.html>
 - Each image is 32x32, 3 colour channels (RGB):
Use just 1 colour or convert to greyscale;
Convert to vector of 1024 floats
 - 50k training images, 10k testing images: you can select random subsets (e.g. 1 batch)
 - 1 hidden layer (larger than first case), 1 output
- You must try to distinguish between **2 classes only**:
 - Send me an email message saying the name of the 2 people in your group and I’ll give you a pair of images.
- *PhD students: If you want to use a different dataset related to your research, contact me in advance for agreement*





Part 5: Deep Learning Enhancements

- This part must be done individually:
 - Do not do it with the other member of your group
 - Include 2 separate sections in your report about what each of you did in Part 5
- Each group member: pick one enhancement that is characteristic of deep learning
 - The two group members have to pick different enhancements from each other
 - implement your enhancement
 - Test how well it works relative to the shallow NN on the image dataset



Assignment Submission & Deadline

- Make a submission at or before the deadline, covering all parts of the assignment, via Blackboard
- You must submit it in TWO formats:
 1. The Notebook itself, including all output from running it (do not expect us to run it)
 2. A PDF print of the Notebook (easier for us to grade)

Note: avoid unnecessary outputs in your notebook; for example, don't print out all rows of the training data file, or every single training iteration
- For a 2-person group:
 - Only 1 person must make the submission, with both group members' names and IDs on it.
 - The submission must have two separate sections for Part 5, one for each group member.
- **DEADLINE: Friday 19 March 2021**



Final Plagiarism Note

- *Plagiarism is the act of copying, including or directly quoting from the work of another without adequate acknowledgement, in order to obtain benefit, credit or gain. Plagiarism can apply to many materials, such as words, ideas, images, information, data, approaches or methods. Sources of plagiarism can include books, journals, reports, websites, essay mills, another student, or another person.*
- If you plagiarise **any part** of your assignment, you will get 0 for your full assignment, with no option to resubmit
- It is **your** responsibility to ensure that you understand the university's rules on what constitutes plagiarism
- You are required to act with academic integrity in carrying out your assignments
- There are many sources of information available online on what constitutes plagiarism, so that you can avoid it.