Individual Assignment questions

The questions below are related to the group assignment and lectures but are not suitable for group work. Please answer these questions individually in writing, and submit the result on Blackboard by the end of October 5th.

Describe the principles of overfitting and how dropout can reduce this (Question 1, 5 points)

Read the research paper "Performance-optimized hierarchical models predict neural responses in higher visual cortex", available from:
http://www.pnas.org/content/pnas/111/23/8619.full.pdf
Write a short (~500 word) summary of the experimental approach and results. (Question 2, 10 points)

It should be clear from the Group Assignment that even a relatively simple deep convolutional learning network is quite computationally intensive to run on a personal computer. Here we will move to a web-based interface for deep learning, at:
http://playground.tensorflow.org/

Here, you can classify the object positions in different data sets (left panel) using deep convolutional network of differing complexity, different numbers of feature maps (number of 'neurons' in each hidden layer), different numbers of layers and different inputs ('features' column). At the top, you can also change the activation function and add normalisation (regularization). In the left column, you can change the ratio of training and test data, and add noise to the network to improve generalization to simulate imperfect inputs.

Play around with these settings and see how they affect your ability to learn classification of different data sets. Write down what you found and how you interpret the effects of these settings. This question is intentionally open to allow you to explore the process. (Question 3, 8 points)

What is the minimum you need in the network to classify the spiral shape with a test set loss of below 0.1? (Question 4, 7 points)

Backpropagation of error:
So far, we have avoided explaining backpropagation in detail. We have discussed what backpropagation does, and used Keras's implementation to train our networks. But we have not looked at how it works because the mathematics are complex and do not fit well with the goals of this course. However, this is a major principle in machine learning with neural networks, so the final part of the assignment will attempt to explain it.

To start this part of the assignment, we suggest watching videos 3 and 4 from the playlist at the following link. You may like to start by watching videos 1 and 2, as

these set things up for videos 3 and 4:
https://www.youtube.com/playlist?list=PLiaHhY2iBX9hdHaRr6b7XevZtgZRa1PoU

Please to answer the following questions (after viewing the content above), each in about 500 words (plus equations):

Explain the principle of backpropagation of error in plain English in about 500 words. This can be answered with minimal mathematical content and should be **IN YOUR OWN WORDS**. What is backpropagation trying to achieve, and how does it do so? (Question 5, 8 points)

BONUS QUESTION:
If you are also attempting to answer question 6, you should also look at the page here: https://github.com/stephencwelch/Neural-Networks-Demystified/blob/master/.ipynb_checkpoints/Part%204%20Backpropagation-checkpoint.ipynb

Describe the process of backpropagation in mathematical terms. Here, explain (in English, in about 500 words) what each equation you give does, and relate this to the answers given in Question 5. You are welcome to express equations in R code (not python) rather than using equation layout (Question 6, 5 points).