L2 regularization

- Derive the gradient of L2 regularization
- Implement L2 regularization for Logistic Regression to predict heart disease
- Train and test the model with multiple lambda parameters (0.00001 to 1) and observe how the weights' magnitude change (enough to evaluate this only after the final training iteration)

Weight decay (L2 regularization), Dropout

- Add weight decay with 3 different lambda parameters (e.g. 0.000001, 0.0001,0.1) to the Neural Network training, train until 10 epochs, observe how validation accuracy changes compared to not using weight decay
- Add Dropout layers with p=0.3 after each hidden layer, train until 10 epochs, observe how validation accuracy changes compared to not using dropout (do not use weight decay here)

Regularization, validation (Homework)

- Implement L1 regularization for Logistic Regression and observe the differences compared to L2 regularization
- Optimize the weight decay lambda and the number of neurons hyperparameter using a grid search with 5 lambda (0.000001,0.00001,0.0001,0.001,0.1) and 5 neuron number (200,450,700,950,1200) hyperparameters (that means 25 combinations). Train only until 10 epochs.
- Describe the relationship between the two hyperparameters (look at the test accuracy after each hyperparameter combination)