

## Practical Machine Learning with Tensorflow

### Assignment 3

1. (50, 25, 25, 5, 100) as the number of examples is generally the first dimension of a tensor. As there are 5 quantities, each consisting of 100 values for each of the cells of the 25x25 grid, the order should be such that the highest level of abstraction should be on the left. Hence, it will be (50, 25, 25, 5, 100).
2. `X[20:, 10:25, :5, :100, :50]` will result in tensor of shape (30, 15, 5, 5, 50). This can be verified by running the slicing operation using numpy.
3. (50, 25, 25, 5) is the solution as the average is computed over the 5th dimension (index 4) and hence 100 values are averaged to get a single value.
4. We are required to build a model classify each sentence using this data, so `x_train` should contain each the sentences as a data-points from all documents. As there are 10 documents with 100 sentences. It means we have 1000 sentences to classify where each sentence contains 20 words (on avg). Hence, (1000, 20) is the correct option.
5. By simply making a dataset running `tf.data.Dataset.from_tensor_slices(t)` and then running the reduce operation on the dataset using the given parameter values, we can get the solution.
6. By simply making a dataset running `tf.data.Dataset.from_tensor_slices(t)` and then running the map and reduce operations on the dataset using the given parameter values, we can get the solution.
7. Here, x and y will be: [5, 6, 7, 8, ..., 14] and [-1, -2, -3, ..., -49]. Using batch size of 7 makes first batch to contain first 7 elements from each array. As we are just left with 3 elements in the first array, the second batch will take 3 elements from each array and store them in **b**. Hence, the answer is :

```
[array([ 5,  6,  7,  8,  9, 10, 11]), array([-1, -2, -3, -4, -5, -6, -7])]  
[array([12, 13, 14]), array([-8, -9, -10])]
```

8. X consists of 9 elements: [1,2,3,...,9]. When we call `repeat(4)`, it will repeat the same list 4 times to make it of size 36. Calling with batch size 20, takes the first 20 elements of this new list. Hence the output will be:  
[1,2,3,4,5,6,7,8,9,1,2,3,4,5,6,7,8,9,1,2]
9. Characteristics of the given data:
  - Contains missing values
  - Categorical columns with String type
  - Time index like year, month

To be able to use this data for training NN, we need to first make sure that the data types are numerical and there are no null values.

10. Given data size: (100, 50)
  - With 70% training data: `x_train = (70, 40)`, `y_train = (70, 10)`
  - With 10% validation data: `x_valid = (10, 40)`, `y_valid = (10, 10)`
  - With 20% testing data: `x_test = (20, 40)`, `y_test = (20, 10)`
11. Try running this code:

```

import math
import numpy as np

def sigmoid(x):
    return 1 / (1 + math.exp(-x))

def relu(x):
    return np.maximum(x, 0)

x = np.array([1, 0.2, 0.4, 0.9, 1, 0, 0.6, 0.3])
w11 = np.array([-0.05, 0.1, 0.1, 0.2, 0.35, 0.6, -0.9, -0.1])
b11 = -0.8
w12 = np.array([-0.5, 0.1, 0.1, 0.02, 0.3, 0.36, 0.9, 0.1])
b12 = -0.1
w13 = np.array([-0.05, 0.1, 0.1, 0.2, 0.35, 0.6, 0.9, -0.1])
b13 = -0.1
w14 = np.array([-0.5, 0.1, 0.1, 0.02, 0.3, 0.36, -0.9, 0.1])
b14 = -0.1
o11 = np.dot(x, w11) + b11
o12 = np.dot(x, w12) + b12
o13 = np.dot(x, w13) + b13
o14 = np.dot(x, w14) + b14
a1 = relu(np.array([o11, o12, o13, o14]))
print(a1)
w21 = np.array([-0.05, 0.1, 0.1, 0.2])
b21 = 0
o21 = sigmoid(np.dot(a1, w21) + b21)
print(o21)

```

12. Refer to the solution of Q11.
13. Dropout helps in regularization by removing the overdependence of output on certain weights by removing weights randomly. Hence, the network doesn't overfit to training data. L2 regularization penalizes the large weights. Increasing the size of the training set by augmentation helps in reducing the variance by inducing more knowledge in the training data as changing the scale, flipping or rotating the images does not change the label of the image.
14. `tf.image.resize` applies resize operation on only the axis specified in its shape argument
15. As we do not know anything about the maximum length of a sentence in the test data, out of the given options, it is better to choose larger `batch_size`.