**Name: Salma Abdel Monem Abdel Motaleb**

**Assignment Three – Week4**

1. Use the mini-bath approach to make the algorithm learn faster by dividing the dataset into several powers of 2 sized batches and make the algorithm learn after each batch instead after the whole dataset.
2. Using the minibatch approach may make the cost function oscillate and so we use smoothing techniques like exponentially weighted averages or the Root mean square optimization techniques instead of just pure gradient descent.
3. The ADAM optimization technique combines the both optimizations of exponentially weighted averages and the Root mean square and so it is much faster.
4. The learning rate could be tuned to be high at the beginning and low at the ending by using one of the learning rate decay techniques
5. The problem usually is not the local minimum but with the plateaus
6. Some hyperparameters have higher effect and so it should be tuned first.
7. The hyperparameter selection process might be done randomly through their space
8. We should use appropriate scale to choose the hyperparameters space effectively.
9. If the computational power is high, we can test multiple models with different hyperparameter settings in parallel till we find the optimum choice, otherwise we can adopt one model and tune it step by step.
10. We can use batch normalization to normalize the inputs to each layer so that the optimizer can learn faster
11. Frameworks are used to enable more professional algorithms.