EEE 443 — Tutorial 5
PDF of Codes

Code Cell-3 (update_parameters_with_gd):

```
parameters['W' + str(l+1)] -= learning_rate * grads['dW' + str(l+1)]
parameters['b' + str(l+1)] -= learning_rate * grads['db' + str(l+1)]
```

Code Cell-5 (random_mini_batches):

```
mini_batch_X = shuffled_X[:, k * mini_batch_size:(k+1) * mini_batch_size]
mini_batch_Y = shuffled_Y[:, k * mini_batch_size:(k+1) * mini_batch_size]

mini_batch_X = shuffled_X[:, num_complete_minibatches * mini_batch_size:]
mini_batch_Y = shuffled_Y[:, num_complete_minibatches * mini_batch_size:]
```

Code Cell-7 (initialize_velocity):

```
v['dW' + str(l+1)] = np.zeros(parameters['W' + str(l+1)].shape)
v['db' + str(l+1)] = np.zeros(parameters['b' + str(l+1)].shape)
```

Code Cell-9 (update_parameters_with_momentum):

```
v['dW' + str(l+1)] = beta * v['dW' + str(l+1)] + (1 - beta) * grads['dW' + str(l+1)]
v['db' + str(l+1)] = beta * v['db' + str(l+1)] + (1 - beta) * grads['db' + str(l+1)]
parameters['W' + str(l+1)] -= learning_rate * v['dW' + str(l+1)]
parameters['b' + str(l+1)] -= learning_rate * v['db' + str(l+1)]
```

Code Cell-11 (initialize_adam):

```
v['dW' + str(l+1)] = np.zeros(parameters['W' + str(l+1)].shape)
v['db' + str(l+1)] = np.zeros(parameters['b' + str(l+1)].shape)
s['dW' + str(l+1)] = np.zeros(parameters['W' + str(l+1)].shape)
s['db' + str(l+1)] = np.zeros(parameters['b' + str(l+1)].shape)
```

Code Cell-13 (update_parameters_with_adam):

```
v['dW' + str(l+1)] = beta1 * v['dW' + str(l+1)] + (1 - beta1) * grads['dW' + str(l+1)]
v['db' + str(l+1)] = beta1 * v['db' + str(l+1)] + (1 - beta1) * grads['db' + str(l+1)]

v_corrected['dW' + str(l+1)] = v['dW' + str(l+1)] / (1 - beta1**t)
v_corrected['db' + str(l+1)] = v['db' + str(l+1)] / (1 - beta1**t)

s['dW' + str(l+1)] = beta2 * s['dW' + str(l+1)] + (1 - beta2) * grads['dW' + str(l+1)]**2
s['db' + str(l+1)] = beta2 * s['db' + str(l+1)] + (1 - beta2) * grads['db' + str(l+1)]**2

s_corrected['dW' + str(l+1)] = s['dW' + str(l+1)] / (1 - beta2**t)
s_corrected['db' + str(l+1)] = s['db' + str(l+1)] / (1 - beta2**t)

parameters['W' + str(l+1)] -= learning_rate * v_corrected['dW' + str(l+1)] /
(np.sqrt(s_corrected['dW' + str(l+1)]) + epsilon)
parameters['b' + str(l+1)] -= learning_rate * v_corrected['db' + str(l+1)] /
(np.sqrt(s_corrected['db' + str(l+1)]) + epsilon)
```