**IDEAL CASE:**

The Below parameters have been used for:

**Q-learning**

* Optimizer=Adam,
* epochs=1,
* batch\_size=32,
* gamma=0.95
* loss=categorical crossentropy

**Policy Gradient:**

* + - learning\_rate = 0.001
    - epochs=1
    - batch\_size = 32
    - gamma = 0.99
    - Optimizer=Adam
    - Loss= categorical crossentropy

**Neural Network Model:**

* Activation\_hidden\_layer = relu
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 32
* Optimizer=Adam

**LSTM Model:**

* Activation\_hidden\_layer = relu
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 32
* Optimizer=Adam

**Output:**

* Accuracy\_q\_learning\_NN: 0.952783
* Accuracy\_policy\_gradient\_NN: 0.951336
* F1\_score\_q\_learning\_NN: 0.949463
* F1\_score\_policy\_gradient\_NN: 0.948056

* Accuracy\_q\_learning\_LSTM: 0.950779
* Accuracy\_policy\_gradient\_LSTM: 0.947327
* F1\_score\_q\_learning\_LSTM: 0.947593
* F1\_score\_policy\_gradient\_LSTM: 0.944056

**TESTCASE-1:**

The Below parameters have been used for:

**Q-learning:**

* Optimizer=Adam,
* epochs=1,
* batch\_size=32,
* gamma=0.95
* loss= categorical crossentropy

**Policy Gradient:**

* + - learning\_rate = 0.001
    - epochs=1
    - batch\_size = 32
    - gamma = 0.99
    - Optimizer=Adam
    - Loss= categorical crossentropy

**Neural Network Model:**

* Activation\_hidden\_layer = tanh
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 128
* Optimizer=Adam

**LSTM Model:**

* Activation\_hidden\_layer = tanh
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 128
* Optimizer=Adam

**Output:**

* Accuracy\_q\_learning\_NN: 0.955122
* Accuracy\_policy\_gradient\_NN: 0.951559
* F1\_score\_q\_learning\_NN: 0.953480
* F1\_score\_policy\_gradient\_NN: 0.949891

* Accuracy\_q\_learning\_LSTM: 0.955011
* Accuracy\_policy\_gradient\_LSTM: 0.953898
* F1\_score\_q\_learning\_LSTM: 0.953284
* F1\_score\_policy\_gradient\_LSTM: 0.952183

**TESTCASE-2:**

The Below parameters have been used for:

**Q-learning:**

* Optimizer=RMSProp,
* epochs=1,
* batch\_size=32,
* gamma=0.95
* loss= categorical crossentropy

**Policy Gradient:**

* + - learning\_rate = 0.001
    - epochs=1
    - batch\_size = 32
    - gamma = 0.99
    - Optimizer=RMSProp
    - Loss= categorical crossentropy

**Neural Network Model:**

* Activation\_hidden\_layer = relu
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 64
* Optimizer=RMSProp

**LSTM Model:**

* Activation\_hidden\_layer = relu
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 64
* Optimizer=RMSProp

**Output:**

* Accuracy\_q\_learning\_NN: 0.945880
* Accuracy\_policy\_gradient\_NN: 0.9452561
* F1\_score\_q\_learning\_NN: 0.943091
* F1\_score\_policy\_gradient\_NN: 0.949752

* Accuracy\_q\_learning\_LSTM: 0.943875
* Accuracy\_policy\_gradient\_LSTM: 0.942762
* F1\_score\_q\_learning\_LSTM: 0.941053
* F1\_score\_policy\_gradient\_LSTM: 0.939643

**TESTCASE-3: #edwin**

The Below parameters have been used for:

**Q-learning:**

* Optimizer=Adam,
* epochs=3,
* batch\_size=64,
* gamma=0.95
* loss=categorical crossentropy

**Policy Gradient:**

* + - learning\_rate = 0.01
    - epochs=3
    - batch\_size = 64
    - gamma = 0.99
    - Optimizer=Adam
    - Loss= categorical crossentropy

**Neural Network Model:**

* Activation\_hidden\_layer = relu
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 64
* Optimizer=Adam

**LSTM Model:**

* Activation\_hidden\_layer = relu
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 64
* Optimizer=Adam

**Output:**

* Accuracy\_q\_learning\_NN: 0.946881
* Accuracy\_policy\_gradient\_NN: 0.944209
* F1\_score\_q\_learning\_NN: 0.944950
* F1\_score\_policy\_gradient\_NN: 0.942554

* Accuracy\_q\_learning\_LSTM: 0.946547
* Accuracy\_policy\_gradient\_LSTM: 0.942984
* F1\_score\_q\_learning\_LSTM: 0.9441860
* F1\_score\_policy\_gradient\_LSTM: 0.9402567

**TESTCASE-4:**

The Below parameters have been used for:

**Q-learning:**

* Optimizer=Adam,
* epochs=2,
* batch\_size=128,
* gamma=0.95
* loss=mean\_square\_error

**Policy Gradient:**

* + - learning\_rate = 0.001
    - epochs=2
    - batch\_size = 128
    - gamma = 0.95
    - Optimizer=Adam
    - Loss= mean\_square\_error

**Neural Network Model:**

* Activation\_hidden\_layer = tanh
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 32
* Optimizer=Adam

**LSTM Model:**

* Activation\_hidden\_layer = tanh
* Activation\_output\_layer = softmax
* number of neurons in the hidden layer = 32
* Optimizer=Adam

**Output:**

* Accuracy\_q\_learning\_NN: 0.953786
* Accuracy\_policy\_gradient\_NN: 0.954009
* F1\_score\_q\_learning\_NN: 0.951445
* F1\_score\_policy\_gradient\_NN: 0.951803

* Accuracy\_q\_learning\_LSTM: 0.948441
* Accuracy\_policy\_gradient\_LSTM: 0.946882
* F1\_score\_q\_learning\_LSTM: 0.946131
* F1\_score\_policy\_gradient\_LSTM: 0.944657