OUTPUTS:

After **200** iterations, the total error is **18.576629427049554. Activation function = sigmoid**
The final weight vectors are (starting from input to output layers)
```
[[ 0.10243711  0.07088822 -0.67034539 -0.70990276]
 [-0.4814823  -0.96022601 -0.14580456 -0.55294729]
 [ 0.30304385 -0.86910583 -0.41557951 -0.37531986]
 [ 0.83298959 -2.22257941 -2.66182528 -1.02320885]
 [ 0.22434251 -0.37179879 -0.56307917 -0.97590874]
 [ 0.89814057 -0.21385872 -0.603848   -0.32376094]]
[[ 0.46327297  0.42407761]
 [-2.74953266 -2.91253444]
 [-2.9682999  -3.90171211]
 [-2.22102426 -1.64619826]]
[[-1.59692839]
 [-2.37889277]]
```
Test Error: 5.965106748782173

After **200** iterations, the **total error** is **18.90680540378554. Activation function = tanh**
The final weight vectors are (starting from input to output layers)
```
[[-0.24660351  0.58508958 -1.25317849 -0.29892837]
 [-0.83650105 -0.51552529 -0.90193429 -0.14064054]
 [-0.18158727  0.11702119 -0.68290327  0.33061748]
 [-0.42227222  0.72378775 -1.19111795  0.28642055]
 [-0.34982224  0.419144   -0.87188174 -0.37110946]
 [ 0.27920268  0.79338559 -0.66388207  0.37837672]]
[[ 1.03260756  0.81460414]
 [-0.75737324 -0.95208232]
 [-0.85655528  0.94494944]
 [-0.34407853 -0.18722177]]
[[ 0.89301333]
 [-0.46807545]]
```
Test Error: 6.1322227511444645

After **1000** iterations, the **total error** is **94.04786073718809. Activation function = relu**
The final weight vectors are (starting from input to output layers)
```
[[-0.16595599  0.44064899 -0.99977125 -0.39533485]
 [-0.70648822 -0.81532281 -0.62747958 -0.30887855]
 [-0.20646505  0.07763347 -0.16161097  0.370439  ]
 [-0.5910955   0.75623487 -0.94522481  0.34093502]
 [-0.1653904   0.11737966 -0.71922612 -0.60379702]
 [ 0.60148914  0.93652315 -0.37315164  0.38464523]]
[[ 0.7527783   0.78921333]
 [-0.82991158 -0.92189043]
 [-0.66033916  0.75628501]
 [-0.80330633 -0.15778475]]
[[0.91577906]
 [0.06633057]]
```
Test Error: 31.64839106633683

OUTPUT RESULTS:

Dataset used: Car data set from UCI
URL: https://archive.ics.uci.edu/ml/machine-learning-databases/car/car.data

Activation Function	Learning Rate	Number of Iterations	Total error	Test Error
sigmoid	0.05	200	18.576629427049554	5.965106748782173
tanh	0.001	200	18.90680540378554	6.1322227511444645
relu	0.009	1000	94.04786073718809	31.64839106633683

From the results obtained from the output, the sigmoid function works best for the car dataset as it has the lowest total error and test error. This result makes sense as sigmoid activation function works well on classification problems and our data set is a classification problem.