Accuracy of random choice: 5.26%

**5000 with 3 fold cross validation**

**Hidden: 10x10x5, Bias fed into output layer, sigmoid activation and step size of 0.5**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 9.42% | 9.72% | 9.57% |
| 2 | 4.89% | 5.52% | 5.21% |
| 3 | 2.04% | 1.68% | 1.86% |
| Average Fold | 5.54% | 5.64% | 5.55% |

**Hidden: 25x10x5, Bias fed into output layer, sigmoid activation and step size of 0.5**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 4.29% | 4.50% | 4.40% |
| 2 | 10.83% | 10.62% | 10.73% |
| 3 | 3.54% | 2.64% | 3.09% |
| Average Fold | 6.22% | 5.92% | 6.07% |

**Hidden: 10x10x5, Bias fed into every layer, sigmoid activation and step size of 0.5**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 5.73% | 5.34% | 5.54% |
| 2 | 4.11% | 3.06% | 3.59% |
| 3 | 5.01% | 3.72% | 4.37% |
| Average Fold | 4.95% | 4.04% | 4.50% |

**Hidden: 10x10x5, Bias fed into every layer, sigmoid activation and step size of 10**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 7.02% | 7.32% | 7.17% |
| 2 | 7.29% | 6.78% | 7.04% |
| 3 | 7.71% | 7.37% | 7.54% |
| Average Fold | 7.34% | 7.16% | 7.25% |

**Hidden: 10x10x5, Bias fed into every layer, sigmoid activation and step size of 50**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 7.02% | 7.32% | 7.17% |
| 2 | 7.86% | 7.08% | 7.47% |
| 3 | 7.71% | 7.37% | 7.54% |
| Average Fold | 7.53% | 7.26% | 7.39% |

**Hidden: 125, Bias fed into every layer, sigmoid activation and step size of 50**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 1.02% | 0.90% | 0.96% |
| 2 | 1.14% | 0.66% | 0.90% |
| 3 | 0.78% | 1.38% | 1.08% |
| Average Fold | 0.98% | 0.98% | 0.98% |

**Hidden: 10x10x5, Bias fed into output layer, sigmoid activation and step size of 50**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 7.02% | 7.32% | 7.17% |
| 2 | 7.86% | 7.08% | 7.47% |
| 3 | 7.71% | 7.37% | 7.54% |
| Average Fold | 7.53% | 7.26% | 7.39% |

**Hidden: 5x10x25x100, Bias fed into output layer, sigmoid activation and step size of 50**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 3.15% | 3.42% | 3.29% |
| 2 | 10.83% | 10.62% | 10.73% |
| 3 | 6.45% | 6.59% | 6.52% |
| Average Fold | 6.81% | 7.10% | 6.85% |

**Hidden: 25x10x10x5, Bias fed into output layer, sigmoid activation and step size of 50**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 7.26% | 7.32% | 7.29% |
| 2 | 8.07% | 10.20% | 9.14% |
| 3 | 7.05% | 6.53% | 6.79% |
| Average Fold | 7.46% | 8.02% | 7.74% |

**Hidden: 10x10x10x10, Bias fed into output layer, sigmoid activation and step size of 10**

Looks like more layers of consistent size is best

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 9.42% | 9.72% | 9.57% |
| 2 | 9.96% | 8.64% | 9.30% |
| 3 | 10.29% | 11.69% | 10.99% |
| Average Fold | 9.89% | 10.02% | 9.95% |

**Hidden: 10x10x10x10x10, Bias fed into output layer, sigmoid activation and step size of 10**

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 7.02% | 7.32% | 7.17% |
| 2 | 7.86% | 7.08% | 7.47% |
| 3 | 7.71% | 7.37% | 7.54% |
| Average Fold | 7.53% | 7.26% | 7.39% |

**Hidden: 10x10x10x10, Bias fed into output layer, tanh activation and step size of 10**

Tan made it worse

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 3.15% | 3.42% | 3.29% |
| 2 | 1.86% | 2.16% | 2.01% |
| 3 | 2.04% | 1.68% | 1.86% |
| Average Fold | 2.35% | 2.42% | 2.39% |

**100000 with 3 fold cross validation**

**Hidden: 10x10x10x10, Bias fed into every layer, sigmoid activation and step size of 10**

100k had different accuracies

|  |  |  |  |
| --- | --- | --- | --- |
| Validation Fold | Training Accuracy | Validation Accuracy | Average Accuracy |
| 1 | 9.89% | 9.95% | 9.92% |
| 2 | 9.92% | 9.89% | 9.91% |
| 3 | 6.88% | 6.69% | 6.79% |
| Average Fold | 8.90% | 8.84% | 8.87% |