Deep Learning Assignment

Babu Kumar (21068)

March 27, 2024

Answer: Q.1

Cross-entropy loss is preferred over mean squared error (MSE) in logistic regression.
Reason:
Probability Interpretation:
Cross-entropy loss aligns well with the probabilistic nature of classification tasks by measuring the difference between predicted probabilities and true class labels, while MSE may not capture this aspect effectively.

Optimization Behavior:
Cross-entropy loss yields more informative gradients, enabling efficient optimization and faster convergence during training, particularly when used with logistic activation functions.
Single Best Answer:

Cross-entropy loss encourages the model to provide a clear decision boundary between classes, ensuring a single best answer - the class with the highest probability, while MSE may not provide such clear decision boundaries.

Robustness to Imbalanced Data:
Cross-entropy loss is more robust to class imbalance, penalizing misclassifications effectively and leading to better performance in datasets with imbalanced class distributions compared to MSE.

Answer: Q-2

In a binary classification task with a deep neural network equipped with linear activation functions, the loss function that guarantees a convex optimization problem is (b) Mean Squared Error (MSE).

Mean Square Error

For binary classification, where y denotes the target label (either 0 or 1) and $\hat{y}$ denotes the predicted output of the model, the output $\hat{y} = WX + b$ is a linear combination of inputs and weights. Since the output is linear, the MSE loss function becomes a quadratic function of the parameters (weights and biases), ensuring convexity in the optimization problem.

Cross Entropy

While Cross Entropy loss is frequently employed in binary classification tasks, it does not assure convex optimization when paired with linear activation functions. The introduction of nonlinearity by the logarithm function renders the loss function non-convex, potentially resulting in multiple local minima in the optimization problem.

Answer: Q-3

We establish a basic feedforward neural network (FNN) consisting of three dense layers. The input images undergo preprocessing using torchvision transforms, which convert them to tensors and normalize their values. The MNIST dataset is utilized for both training and testing purposes. Training of the FNN model is conducted via stochastic gradient descent (SGD) employing cross-entropy loss. Subsequently, the trained model undergoes evaluation on the test set to assess its accuracy. Hyperparameters like learning rate, the number of hidden layers, neurons per layer, and activation functions (ReLU in this scenario) remain pivotal for optimizing the model's performance. Techniques such as grid search or random search can facilitate hyperparameter tuning, while methods like dropout or batch normalization serve to mitigate overfitting and enhance generalization.

Results

Epoch 1, Batch 100, Loss: 2.225
Epoch 1, Batch 200, Loss: 1.974
Epoch 1, Batch 300, Loss: 1.521
Epoch 1, Batch 400, Loss: 1.062
Epoch 1, Batch 500, Loss: 0.805
Epoch 1, Batch 600, Loss: 0.670
Epoch 1, Batch 700, Loss: 0.587
.
.
.
Epoch 4, Batch 700, Loss: 0.284
Epoch 4, Batch 800, Loss: 0.274
Epoch 4, Batch 900, Loss: 0.282
Epoch 5, Batch 100, Loss: 0.244
Epoch 5, Batch 200, Loss: 0.286
Epoch 5, Batch 300, Loss: 0.289
Epoch 5, Batch 400, Loss: 0.293
Epoch 5, Batch 500, Loss: 0.272
Epoch 5, Batch 600, Loss: 0.270
Epoch 5, Batch 700, Loss: 0.260
Epoch 5, Batch 800, Loss: 0.266
Epoch 5, Batch 900, Loss: 0.255
Finished Training
Accuracy on the test set: 92.32

Answer: Q-4

The provided code loads pretrained models such as LeNet-5, AlexNet, VGG, and ResNet-18, 50, and 101, trains them on the Street View House Numbers (SVHN) dataset using cross-entropy loss and stochastic gradient descent (SGD), and evaluates their performance on the test set by computing accuracy.

- LeNet-5 achieved 34.67% accuracy.
- VGG achieved 94.54% accuracy.
- ResNet-18 achieved 92.24% accuracy.

- ResNet-50 achieved 93.11% accuracy.
- ResNet-101 achieved 92.47% accuracy.

From the results:

- VGG demonstrates the highest accuracy of 94.54%, indicating its strong suitability for the SVHN dataset. VGG architectures are known for their deep layers and strong feature extraction capabilities, which likely contributed to its superior performance on this complex dataset.
- ResNet-50 also performs notably well with an accuracy of 93.11%, showcasing the effectiveness of its residual connections in capturing intricate patterns present in the SVHN images.
- LeNet-5, on the other hand, exhibits the lowest accuracy of 34.67%, which could be attributed to its relatively shallow architecture compared to modern deep learning models, limiting its capacity to learn complex features from the SVHN dataset.
- AlexNet's performance was not mentioned, but given its similarity to VGG in terms of architecture, it could perform reasonably well on this task, likely achieving accuracy closer to VGG's performance.
- ResNet-18 and ResNet-101 also achieved competitive accuracies, demonstrating the effectiveness of residual networks in handling intricate patterns present in the SVHN dataset. However, their accuracies are slightly lower compared to VGG and ResNet-50, possibly due to their respective depths and complexities.