CMPUT 367 Assignment 1 Question 3

Scientific Question: Does regularization improve the generalization ability of the Softmax Regression classifier on a given dataset?

Hypothesis: Incorporating L2 regularization into the Softmax Regression classifier will improve its generalization ability, thereby increasing validation accuracy and reducing overfitting compared to the same classifier without regularization.

Analysis Consideration: Constant Learning Rate

- **Controlled Experiment**: Keeping the learning rate constant at 0.1 throughout the experiment ensures that any observed differences in model performance are due to the effects of L2 regularization. This control is crucial for attributing changes in generalization ability directly to the regularization technique.
- **Simplification of Analysis**: A constant learning rate simplifies the experimental setup, making it easier to isolate and understand the impact of L2 regularization on the classifier's performance. It removes the learning rate as a variable, focusing the analysis on the effects of regularization.

Regularization rate	Best validation accuracy along training
1	0.7732
0.1	0.8722
0.01	0.9006

Graphs are in the codebase file with trainlossq3-1,2,3.png and valid_accq3-1,2,3.png.
From the results, we see that

- If the regularization rate is too large, the learning curve is noisy and the training does not converge well.
- If the learning rate is too small, the learning curve is smooth but the training is slow.
- A moderate learning rate is desired.

Conclusion:

- **Regularization Rate Impact**: The experiment indicates that L2 regularization does have an impact on the validation accuracy of the Softmax Regression classifier. However, contrary to the hypothesis, the highest validation accuracy is observed with the lowest regularization rate tested (0.01), suggesting that a lower degree of regularization is more beneficial for the dataset and model configuration used in this experiment.
- **Over-regularization**: A regularization rate of 1 appears to have over-penalized the model complexity, resulting in lower validation accuracy. This could be due to the model being too constrained to capture the underlying patterns in the data, leading to underfitting.

Based on the results provided where the learning rate was held constant at 0.1 and the regularization rate was varied, the hypothesis that "Incorporating L2 regularization into the Softmax Regression classifier will improve its generalization ability, thereby increasing validation accuracy and reducing overfitting compared to the same classifier without regularization" can be partially satisfied, with important nuances:

- The hypothesis is confirmed for lower regularization rates (0.01), where the model achieved the highest validation accuracy (0.9006). This suggests that at this level of regularization, the model's generalization ability improved compared to higher regularization rates.
- However, when the regularization rate is increased to 1, the validation accuracy drops significantly to 0.7732. This indicates that excessive regularization can harm the model's

performance, likely due to underfitting, where the model is too constrained to learn the underlying patterns in the data.

The hypothesis holds true under the condition that the regularization rate is appropriately tuned. In this experiment, a lower regularization rate (0.01) rather than a higher rate (1) led to improved validation accuracy, which supports the hypothesis within the optimal regularization range. However, it also reveals that the benefit of regularization has a threshold: beyond a certain point, increasing the regularization rate can negatively affect model performance.
Therefore, the conclusion is that L2 regularization does indeed benefit the generalization of the Softmax Regression classifier on the given dataset when applied with a suitable regularization strength. The optimal regularization rate appears to be less than 1 for the constant learning rate of 0.1, with the best performance observed at a regularization rate of 0.01 in this experiment.