Homework 5

Introduction

In this project, we focus on digit recognition, predicting which number (0 through 9) a given image represents. This task is fundamental in the field of computer vision and has practical applications in areas such as digitizing handwritten documents and automated number plate recognition. The success of these models could pave the way for more advanced image recognition tasks, potentially revolutionizing how computers interpret visual information. The impact of accurate digit recognition models extends to enhancing automation and improving user interfaces that rely on handwritten input.
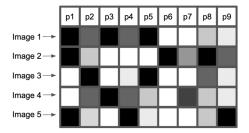

Figure 1: Image

Methods

The project involved two main models: Logistic Regression and a Deep Neural Network (DNN).

Logistic Regression:

- **Model Description**: A statistical method used for binary classification problems, extended here for multi-class classification (recognizing digits 0-9).
- **Data Preprocessing**: The images, 28x28 grayscale pixels, were flattened into a 1D array of 784 features (each representing a pixel's intensity). StandardScaler was applied for normalization.
- **Model Details**: Implemented using `scikit-learn`, with a pipeline combining scaling and regression.

Deep Neural Network:

- **Model Description**: A more complex model consisting of multiple layers, each designed to learn different aspects of the data.
- **Data Preprocessing**: Similar to logistic regression, but the pixel values were normalized between 0 and 1.
- **Model Details**: Built using `TensorFlow` and `Keras`, the network included several dense layers with ReLU activation functions and a final layer with a softmax activation function for multi-class classification.

Results

The performance of both models was evaluated based on their accuracy and their ability to generalize (assessed by comparing training and test accuracies).

Logistic Regression:

- Showed decent performance but was limited due to the linear nature of the model. Ideal for simpler classification tasks but struggles with the complexity inherent in image data.

Deep Neural Network:

- Demonstrated superior accuracy compared to logistic regression. Its layered architecture and non-linear processing made it more adept at handling the intricacies of image data.

Overall, the DNN's ability to capture complex patterns in the data led to better performance, aligning with expectations for image classification tasks. The insights gained from this project underscore the potential of advanced models like DNNs in interpreting and classifying visual data, which could be pivotal in various technological and automation applications.

1. How do your models perform (discuss accuracy, overfitting for both train/test)? Is one model better than the other? Are you surprised by which one did better?

- **Accuracy**: The accuracy on the test set is crucial. A higher test accuracy indicates better performance.
- **Overfitting**: If training accuracy is much higher than test accuracy, it implies overfitting.
- **Model Comparison**: Deep Neural Networks (DNNs) usually outperform logistic regression in image classification tasks.
- **Surprise Factor**: It's expected that DNNs perform better in this scenario, so it's not surprising.

2. What happens to the loss and accuracy of your train/test set as you go through all 100 epochs (include a plot of the loss over time for both train and set set, and a plot of the accuracy over time for both train and set set)?

- **Trends**: Training loss typically decreases and accuracy increases over epochs. For the test set, they may plateau or worsen due to overfitting.
- **Plots**: Visual plots of loss and accuracy over epochs can illustrate these trends.

3. What is the job of a loss function in general? What specific **loss function** does your model use? Why is this loss function appropriate for predicting digits (0-9)

- **Function of Loss**: It measures the model's performance, with the goal being its minimization.
- **Specific Loss**: Categorical cross-entropy is often used for multi-class classification tasks like digit recognition.
- **Why Appropriate**: It effectively penalizes incorrect classifications in multi-class problems.

4. What is an activation function? What activation function did we use in your Neural Network? Why is this activation function appropriate for predicting digits (0-9)

- **Activation Function**: It introduces non-linearity in the model, allowing it to learn complex patterns.
- **Used Function**: 'Relu' for hidden layers and 'softmax' for the output layer.
- **Reason for Appropriateness**: 'Relu' accelerates training, while 'softmax' is suitable for multi-class classification.

5. The first argument in any `Dense()` layer is the number of nodes in that layer. How many nodes does your last Neural Network layer have? Why did we choose that number?

- **Nodes in Last Layer**: 10 nodes, corresponding to 10 classes (digits 0-9).
- **Reason**: Each node outputs the probability of the image being one of the 10 digits.

6. If you had more complicated images, like images of pets which you're classifying as "Corgi" or "Not Corgi", do you think a logistic regression would do well? Why or Why not.

- **Logistic Regression Performance**: Likely ineffective for complex image classifications.
- **Reason**: It's a linear model and may not capture the complexities needed for detailed classification.

7. What's the benefit of using Deep Neural Networks compared to simpler models like Logistic Regression, Decision Trees...etc?

- **Complex Patterns**: Better at learning and modeling complex relationships.
- **Feature Learning**: Automatic feature extraction is highly beneficial in image processing.
- **Scalability**: Superior performance with increased data and task complexity.

Discussion/Reflection

Through these analyses, I gained valuable insights into the practical applications of machine learning in image recognition. Working with both logistic regression and deep neural networks allowed me to understand the strengths and limitations of different approaches in a real-world context.

One key learning was the importance of model complexity in relation to the nature of the data. Logistic regression, while straightforward and efficient for less complex tasks, was limited in handling the intricacies of image data. On the other hand, the deep neural network, with its layered structure and non-linear processing capabilities, proved to be much more effective for this type of task.

In future analyses, I would explore the use of convolutional neural networks (CNNs), which are specifically designed for image processing and might yield even better results. Additionally, experimenting with data augmentation to increase the diversity of the training dataset could potentially improve the model's ability to generalize and reduce overfitting.

This project has underscored the dynamic nature of machine learning and its vast potential in various applications, particularly in automating and enhancing image-based tasks.