MNIST DATA SET ANALYSIS

Ian Gorrell , Matija Vuckovic , and Angel Montes

Dept. of Computer Science, North Central College, Naperville, USA

Abstract: This project investigates four machine learning models; MLP, CNN, RNN, and KNN on the MNIST digit classification task. Each model was evaluated for accuracy, memory usage, training time, and implementation complexity. CNN achieved the highest accuracy (97.82%), followed by RNN (95.11%) and MLP (94.78%). KNN offered quick prototyping but slightly lagged in accuracy (92.00%). These findings highlight how different model structures impact performance on image classification tasks.

1. Introduction

In image classification, different neural architectures offer trade-offs between complexity, memory usage, and accuracy. To explore these trade-offs, we implemented and compared four popular models; MLP, CNN, RNN, and KNN on the MNIST dataset. The dataset consists of 28x28 grayscale images of handwritten digits and provides a robust benchmark for evaluating model performance. Our aim was to identify strengths and weaknesses of each model in terms of training efficiency and real-world applicability.

2. Methodology

A. Model Implementation

Each model was implemented using PyTorch or scikit-learn, trained for 10 epochs with a batch size of 64 and a learning rate of 0.01. CNN consisted of two convolutional layers followed by pooling and fully connected layers. MLP used three fully connected layers, RNN included one recurrent and one dense layer, and KNN used no training phase, classifying by proximity.

B. Evaluation Metrics

Models were compared using test accuracy, training time, and memory usage. Metrics were tracked over all epochs. Accuracy trends and memory use were visualized, and final values were used for direct comparison.

CNN achieved the highest accuracy (97.82%), with RNN and MLP closely behind (95.11% and 94.78%, respectively). KNN achieved 92.0% with no training loop. In terms of memory usage, CNN required the most (19.06 MB), followed by RNN and MLP at approximately 17.5 MB each. KNN, despite having no training phase, used 21.91 MB during inference due to storing and comparing all input vectors.

Regarding training time, CNN had the longest duration at 159.30 seconds, followed by RNN at 152.27 seconds and MLP at 149.79 seconds. KNN had the shortest total runtime at 29.13 seconds since it performs no training. Despite its resource demands, CNN demonstrated the best overall performance, making it the most effective model for this image classification task.

4. Conclusions

This study confirmed CNN's dominance in image recognition tasks due to its ability to detect spatial features. MLP and RNN also performed well, with RNN surprisingly effective despite image data not being sequential. KNN, while not trainable, offered a quick and lightweight alternative. Overall, model choice depends on the task and available resources.

References

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

Y. LeCun and C. Cortes, "MNIST handwritten digit database," [Online]. Available: http://yann.lecun.com/exdb/mnist/