MNIST Classification:
A Comparative Study

DS 4002, Project 3: Images Data, 04/11/2025

Group 2 (**MEE**): Emily McMahon (Leader), Maggie Crowner, Ella Thomasson

Project Context

Motivation

Handwritten digit recognition is key for automation tasks like check processing, postal sorting, and apps such as mobile banking and ATMs.

Research Question

How accurately can a computer identify digits from images of handwriting? How does accuracy vary across different machine learning models?

Goals

Classify handwritten digits and determine whether RF, LDA, NN, or CNN performs best on the handwriting image dataset.

Hypothesis

We will be able to classify MNIST handwritten digit image data with 95% accuracy using RF, LDA, NN, and CNN models.

Modelling Approach

Objective & Models

Compare the classification accuracy of four models on the MNIST handwritten digits dataset, highlighting strengths and tradeoffs.

Data & Preparation

Use the pre-split MNIST dataset from Keras to reshape and preprocess the data for modeling.

Evaluation

Fit and test each model, then assess performance using accuracy scores and confusion matrix heatmaps.

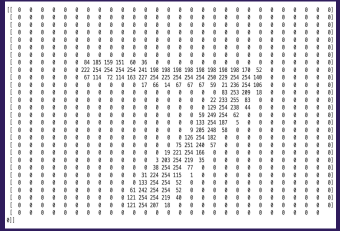
Data Acquisition and Preprocessing

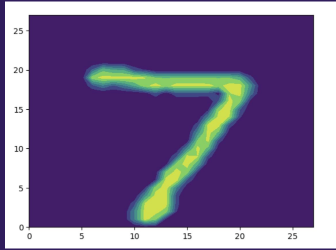
DATA ACQUISITION

- **Format**: 28x28 grayscale images of handwritten digits (0–9)
- **Size**: 60,000 training, 10,000 test images
- **Type**: Image classification; pixel values range 0–255
- **Use**: Suitable for traditional ML and deep learning models
- **Source**: Provided via Keras; originally from NIST

PREPROCESSING

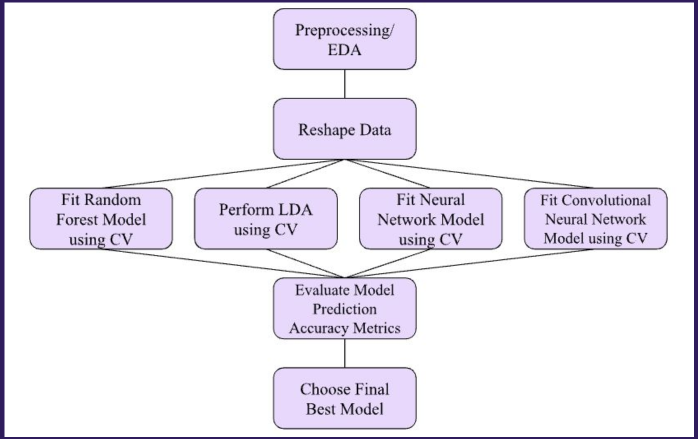
- **Merge sets**: Combine train/test → X: (70000, 28, 28), y: (70000,)
- **Reshape**:
  - RF, LDA, NN → (70000, 784)
  - CNN → (70000, 28, 28, 1)
- **Normalize**: Pixel values scaled to [0, 1]
- **Labels**: Converted to categorical





Analysis Plan and Justification

Analysis Plan



Justification

- Approach
  - **Cross-validation**: Ensures reliable, unbiased results.
  - **Data Reshaping**: Necessary for different models.
  - **Model Variety**: Helps find the best performing model.
- Challenges:
  - **Tuning**: Adjusting hyperparameters for optimal results.
  - **Time**: Cross-validation and training are time-consuming.
- Validation
  - **Confusion Matrix**: Identifies misclassifications.
  - **Accuracy**: Compares model performance clearly.

Challenges and Considerations

Tricky Analysis Decision

- **Problem**:
  Each model requires different input formats. RF, LDA, and NN need flattened (1D) data, while CNN needs 4D tensors with a channel dimension.

- **Significance**:
  Without consistent preprocessing, results wouldn't be comparable—risking biased performance metrics.

- **Solution**:
  - Reshaped data appropriately
  - Normalized pixel values
  - Labels were treated as categorical
  - Applied cross-validation across all models

Bias/Uncertainty

- **Biases**:
  1. **Handwriting:** MNIST digits are cleaner than real-world handwriting.
  2. **Class Imbalance:** Minor discrepancies in the number of observations for each digit.
  3. **Pre-split Train/Test:** Could introduce bias if not randomized

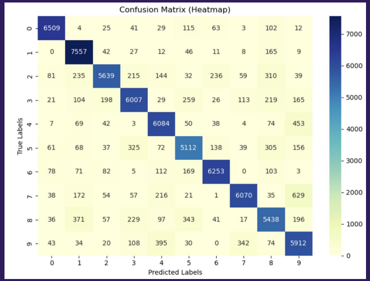
- **Solutions**:
  3. Combined train/test sets and sed 5-fold CV

- **Uncertainty Estimation**:
  - Examined confusion matrices to assess where predictions failed

Results and Conclusions

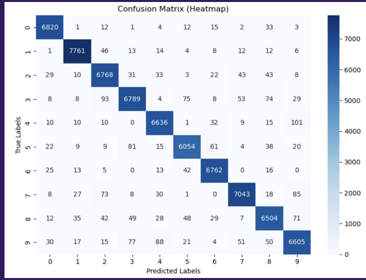
LDA

87%



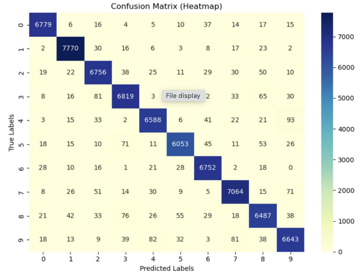
RF

97%



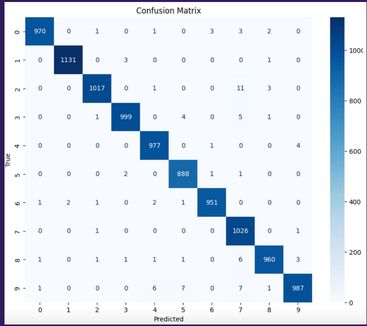
NN

97%



CNN

99%



Conclusion: Three models—RF, NN, and CNN—reached 95% accuracy, supporting our hypothesis. CNN’s strong performance highlights its value for image-based classification.

Next Steps

- Test on real-world handwriting datasets
  - To evaluate model generalizability beyond MNIST’s clean digit images.
- Explore additional models
  - Try more advanced architectures such as CNN Hybrids.
- Augment training data
  - Add noise, rotation, or distortion to simulate real-world variations.

References

Github: https://github.com/EllaThomasson/DS4002-Project3.git

References

[1] R. Yang, “Classifying Hand Written Digits with Deep Learning,” Intelligent Information Management, vol. 10, no. 02, pp. 69–78, 2018, doi: https://doi.org/10.4236/iim.2018.102005.

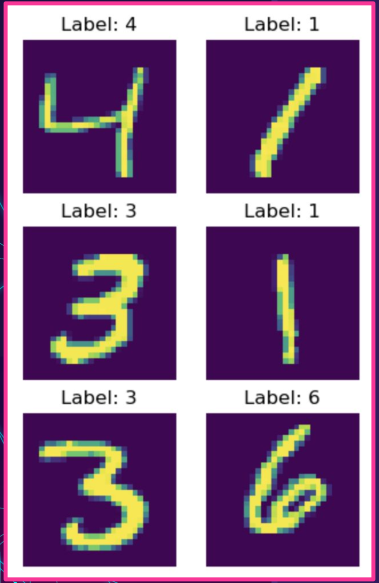
[2] J. Holdsworth and M. Scapicchio, “Deep learning,” Ibm.com, Jun. 17, 2024. https://www.ibm.com/think/topics/deep-learning

[3] G. Boesch, “A Complete Guide to Image Classification in 2021,” viso.ai, Aug. 24, 2021. https://viso.ai/computer-vision/image-classification/

[4] P. Roßbach, “Neural Networks vs. Random Forests -Does it always have to be Deep Learning?,” 2018. Available: https://blog.frankfurt-school.de/wp-content/uploads/2018/10/Neural-Networks-vs-Random-Forests.pdf

[5] K. Team, “Keras documentation: MNIST digits classification dataset,” keras.io. https://keras.io/api/datasets/mnist/

[6] S. Verma, “Understanding Input Output shapes in Convolution Neural Network | Keras,” Medium, Aug. 31, 2019. https://medium.com/data-science/understanding-input-and-output-shapes-in-convolution-network-keras-f143923d56ca (accessed Apr. 09, 2025).



THANKS!

Do you have any questions?