Multiclass Classification of Wheat Seeds Using a Feed-Forward Artificial Neural Network

Renan Da Silva Sousa - NF1012001

Master of Data Analytics (MDA) Program, University of Niagara Falls Canada

Fall 2025 Machine Learning (DAMO-640-8)

Professor: William Pourmajidi

December 2025

Table of Contents

1. Introduction

This project applies to the Seeds dataset, which contains 210 samples of wheat kernels described by seven geometric features. Each sample belongs to one of three wheat varieties. The goals of this project are to perform exploratory data analysis (EDA), prepare the dataset, train an ANN, compare dropout hyperparameters, and evaluate performance on a test set.

2. Methods

2.1 Data Loading and Preprocessing

The dataset was loaded from the local data/raw/ folder. Columns were assigned according to the UCI Seeds description. No missing values were found. The class labels (1–3) were shifted to 0–2 to match TensorFlow requirements.

The dataset was split into 80% training and 20% testing using stratify=y, resulting in 42 test samples (14 per class). StandardScaler was applied to all features, fit on the training set.

2.2 ANN Architecture

The ANN architecture consisted of:

- Input layer: 7 features
- Hidden Layer 1: 16 units, ReLU, L2 regularization
- Dropout
- Hidden Layer 2: 8 units, ReLU, L2 regularization
- Dropout
- Output layer: 3 units, softmax

Two models were compared:

- Model A: Dropout = 0.2
- Model B: Dropout = 0.5

Training was performed using categorical cross-entropy loss, Adam optimizer, batch size 16, validation split 0.2, and 100 epochs.

3. Results

3.1 Training and Validation Curves

Training curves showed that the model with dropout 0.2 achieved smoother and higher validation accuracy compared to dropout 0.5. Validation loss also converged better with dropout 0.2.

3.2 Test-Set Performance

Model A (Dropout 0.2):

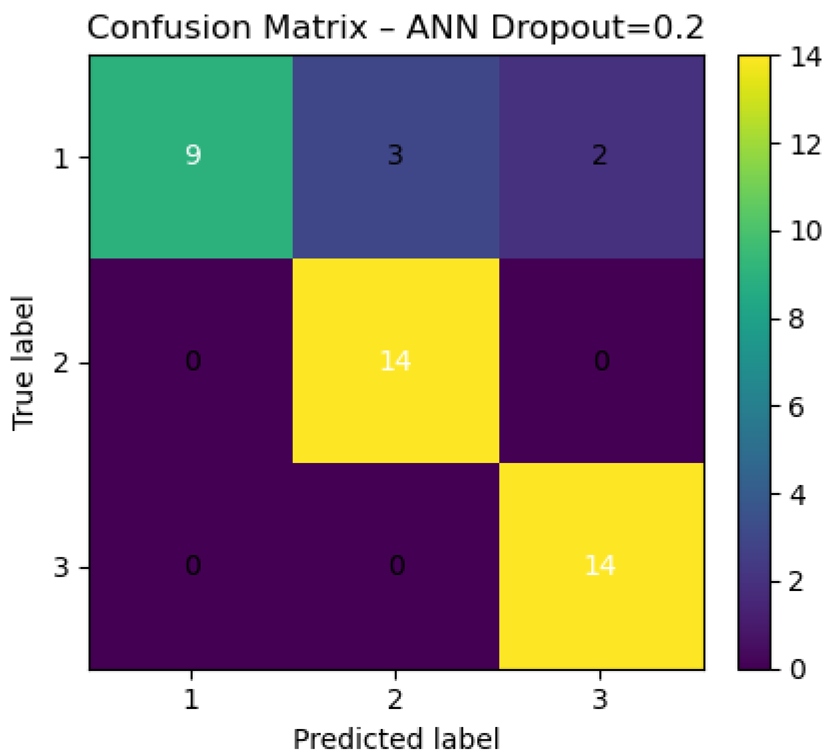
```
=== Model ===
Accuracy : 0.8810
Precision: 0.8995
Recall   : 0.8810
F1-score : 0.8731

Classification report (per class):
              precision    recall  f1-score   support

     Class 1     1.0000    0.6429    0.7826        14
     Class 2     0.8235    1.0000    0.9032        14
     Class 3     0.8750    1.0000    0.9333        14

    accuracy                         0.8810        42
   macro avg     0.8995    0.8810    0.8731        42
weighted avg     0.8995    0.8810    0.8731        42
```

Model B (Dropout 0.5):

```
=== Model ===
Accuracy : 0.8571
Precision: 0.8824
Recall   : 0.8571
F1-score : 0.8446

Classification report (per class):
              precision    recall  f1-score   support

     Class 1     1.0000    0.5714    0.7273        14
     Class 2     0.8235    1.0000    0.9032        14
     Class 3     0.8235    1.0000    0.9032        14

    accuracy                         0.8571        42
   macro avg     0.8824    0.8571    0.8446        42
weighted avg     0.8824    0.8571    0.8446        42
```

Model A outperformed Model B across all metrics.

3.3 Confusion Matrix

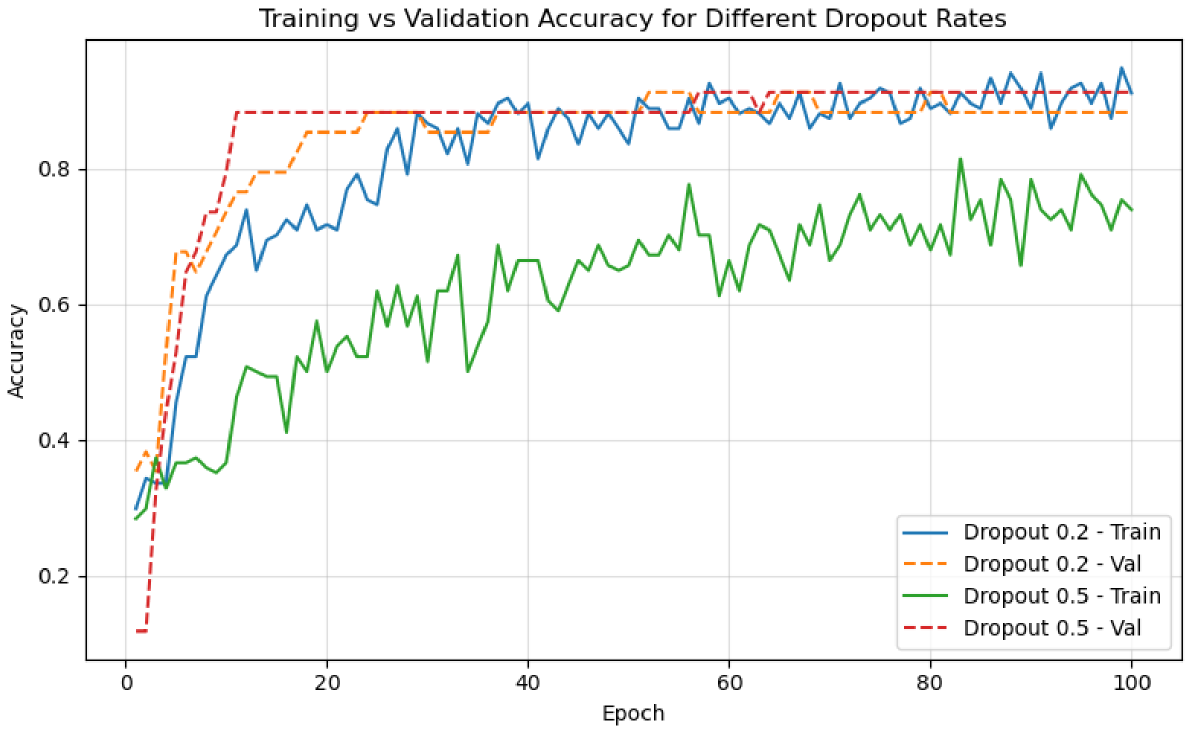


The confusion matrix for the best model (dropout 0.2) showed:

Class 2 and Class 3 were perfectly classified (14/14 correct each).

Class 1 had 5 misclassifications: 3 predicted as Class 2 and 2 as Class 3.

This indicates that classes 2 and 3 are more distinct in feature space, while class 1 overlaps more with others.



4. Discussion

The evaluation clearly shows that **Model A (dropout 0.2)** delivers stronger and more consistent generalization compared to the more heavily regularized Model B. The moderate dropout level enables the model to learn meaningful structure in the data while still preventing overfitting. In contrast, the 0.5 dropout rate imposes excessive regularization, limiting learning capacity and resulting in slower convergence and weaker predictive performance.

Class-level behavior reinforces this finding: while classes 2 and 3 are consistently well separated, class 1 exhibits greater overlap with the other varieties, explaining its lower recall. Nonetheless, overall precision and recall remain high across all classes, indicating that the model performs reliably and without systematic bias.

5. Conclusion

The final evaluation confirms **Model A (dropout = 0.2)** as the superior configuration for this classification task. It achieves a strong balance between accuracy, robustness, and generalization, outperforming the more constrained dropout-0.5 model across all key test metrics. Given the relatively small dataset size, a moderate dropout rate proves optimal, enabling the ANN to capture important geometric patterns without overfitting.

Overall, the final model is accurate, stable, and well-suited for further experimentation, comparative studies, or potential deployment in analytical workflows.

6. GitHub Repository

https://github.com/RenanSdeSilva/-UNFC-seeds-ann-project.git

This project is fully documented and version-controlled in a dedicated GitHub repository. The repository includes:

- **All source code** used for data preprocessing, model training, evaluation, and visualization.
- A **structured folder layout** containing raw data, notebooks, and report.
- The complete **Jupyter Notebook** used to run the analysis end-to-end.
- A **README** file summarizing the project objectives, methodology, and execution steps.

The GitHub repository ensures transparency, reproducibility, and maintainability of the project. It also provides a framework for future enhancements, such as extending the model, integrating hyperparameter optimization, or adapting the workflow for similar classification tasks.