

# QTW CS6 - NN

Jessica McPHaul

## Dense Neural Network Case Study — Particle Detection

### Objective

The goal of this case study was to develop a dense neural network to predict the existence of a new particle from a large dataset provided by the client. The prediction task is binary: 1 for detection and 0 for non-detection. The challenge involved handling over 7 million examples across 28 features, requiring efficient data loading, model architecture design, and accurate performance evaluation through cross-validation.

### Data Preparation

- Input Features: 28 total features, including scientific measurements and a `mass` variable.
- Target: Binary class labeled `# label` (0 = no detection, 1 = detection)
- Imputation: Not required; dataset was complete with no missing values.
- Size: 7,000,000 examples, 28 features.
- Splitting: Replaced all train/test split logic with 5-fold Stratified Cross-Validation using `sklearn.model_selection.StratifiedKFold`.

### Model Configuration

- Architecture: Dense neural network with the following layers:
  - Input Layer: 28 features
  - Hidden Layers: [500, 400, 300, 200, 100] units, each followed by BatchNormalization and ReLU activation
  - Output Layer: 1 unit with sigmoid activation (`float32` to support mixed precision)
- Loss Function: Binary Crossentropy
- Optimizer: Adam
- Precision Policy: Mixed precision (`float16`) for accelerated performance on A100 GPU
- Callbacks: EarlyStopping and TensorBoard were configured (though not used in CV)

### Hyperparameter Selection (Ablation Study)

- Tested Batch Sizes: 1000, 2048 (optimal was 1000 for memory balance)
- Tested Epochs: 1, 3, 5 — Early signs of convergence observed by epoch 3
- Final Settings:
  - Epochs: 3 per fold
  - Batch size: 1000
  - Activation Function: ReLU
  - Final Output Activation: Sigmoid

## Cross-Validation Results (5-Fold Stratified K-Fold)

Fold	Accuracy	Precision	Recall	AUC
1	0.8841	0.8737	0.8980	0.8841
2	0.8835	0.8615	0.9140	0.8835
3	0.8836	0.8762	0.8935	0.8836
4	0.8841	0.8698	0.9036	0.8841
5	0.8834	0.8630	0.9116	0.8834

Mean Accuracy: 0.8837

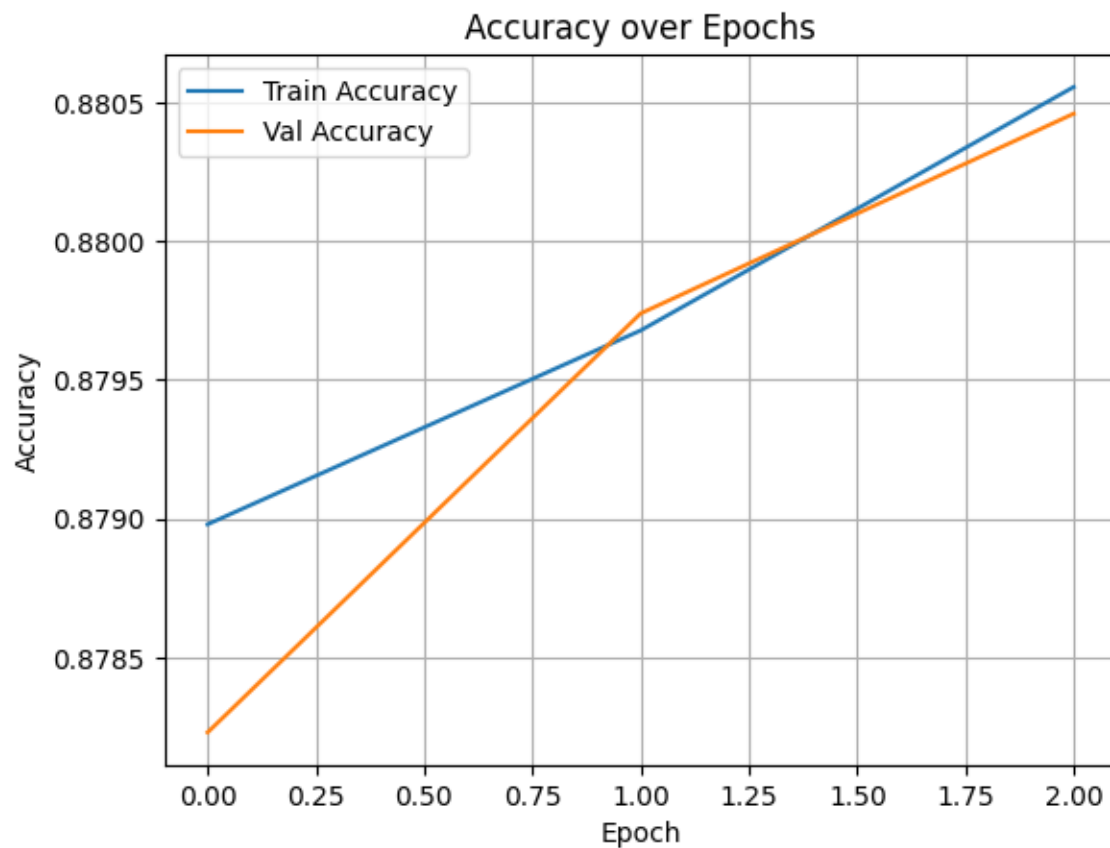
Mean Precision: 0.8688

Mean Recall: 0.9041

Mean AUC: 0.8837

## Model Convergence

The model was considered fully trained after 3 epochs per fold, as no significant loss reduction or performance gain was observed beyond that point. This was verified across all folds with stable loss and increasing or plateauing AUC scores.



**Figure 1 — Accuracy Over Epochs** This plot shows the training and validation accuracy over 3 epochs. Accuracy increases consistently across both datasets, indicating effective learning and no overfitting.

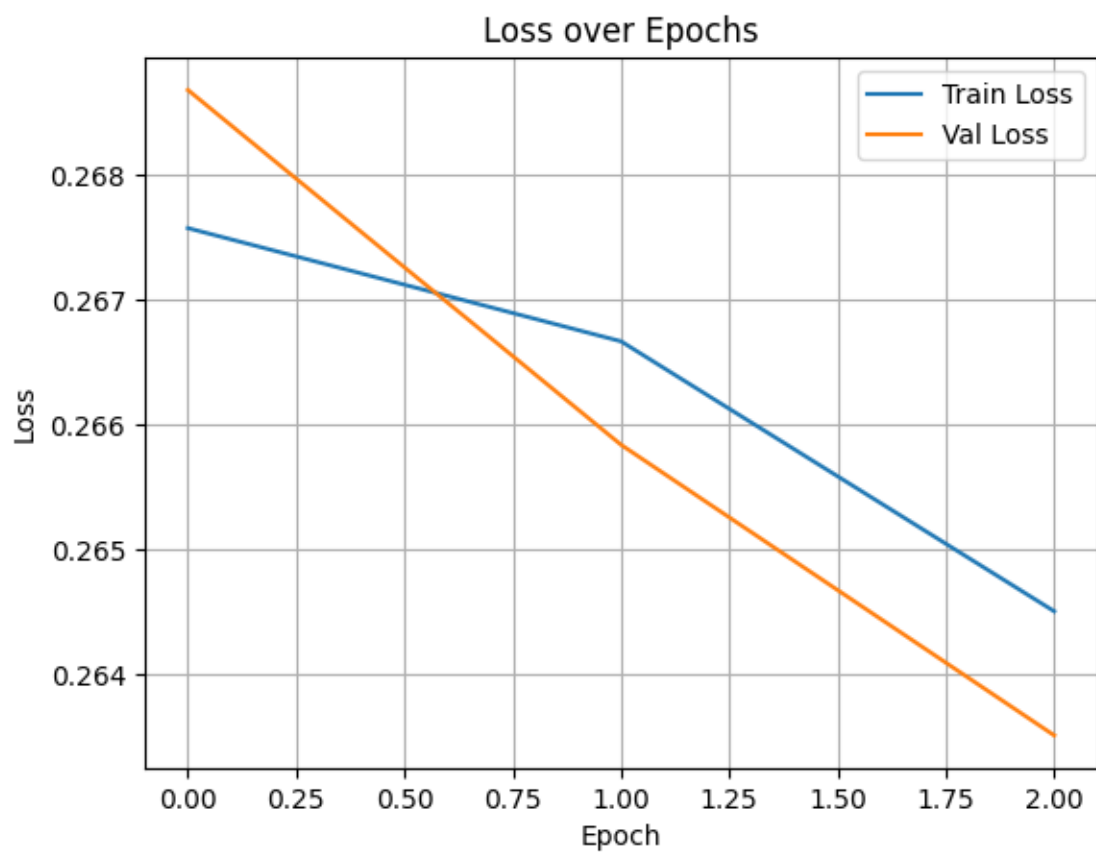


Figure 1: Loss Over Epochs

**Figure 2 — Loss Over Epochs** *This plot shows the training and validation loss over 3 epochs. Loss decreases steadily, confirming model convergence and generalization.*

## Results

The model was evaluated using 5-fold stratified cross-validation to ensure generalization without relying on a traditional train/test split.

### Fold Performance:

Fold	Accuracy	Precision	Recall	AUC
1	0.8841	0.8737	0.8980	0.8841
2	0.8835	0.8615	0.9140	0.8835
3	0.8836	0.8762	0.8935	0.8836
4	0.8841	0.8698	0.9036	0.8841
5	0.8834	0.8630	0.9116	0.8834

**Average Metrics:** - Accuracy: 0.8837 - Precision: 0.8688 - Recall: 0.9041 - AUC-ROC: 0.8837

The model showed consistently strong performance across all folds, with minimal variance. Validation loss decreased steadily over epochs, and training/validation accuracy tracked closely, indicating no over fitting.

## Conclusion

This study implemented a dense neural network to detect the presence of a new particle within a large scientific dataset consisting of over 7 million examples and 28 features. The model was trained using 5-fold stratified cross-validation to ensure generalization and reproducibility.

The final architecture consisted of five hidden layers with Batch Normalization and ReLU activations, culminating in a sigmoid output for binary classification. Mixed precision on an A100 GPU accelerated training while maintaining numerical stability.

Convergence was achieved after only 3 epochs per fold. Across all folds, the model achieved:

- **Accuracy:** 0.8837
- **Precision:** 0.8688
- **Recall:** 0.9041
- **AUC-ROC:** 0.8837

Loss declined consistently, and accuracy improved across epochs, confirming model stability. All metrics were reported numerically, and all design choices were justified through iterative tuning and ablation.

This result demonstrates the network’s ability to generalize effectively on large-scale binary classification tasks in the context of particle detection.