Executive Summary — Fashion-MNIST Classification (ANN & CNN)

Objective:

The goal of this project is to implement and compare **Artificial Neural Networks (MLP)** and **Convolutional Neural Networks (CNN)** for multi-class classification of the Fashion-MNIST dataset (28×28 grayscale images, 10 clothing categories). The comparison focuses on **baseline models** versus **upgraded models with data augmentation and regularization** to achieve higher accuracy.

Methodology:

1. **Dataset Handling**
   - 60,000 training images, 10,000 test images.
   - Normalized pixel values to [0,1]; added channel dimension for CNN.
   - Split training data into 90% training / 10% validation.
2. **Model Architectures**
   - **Basic MLP (ANN)**: Flatten → Dense(128) → Dropout → Dense(64) → Dropout → Dense(10, softmax).
   - **Basic CNN**: Conv32 → MaxPool → Conv64 → MaxPool → Flatten → Dense128 → Dropout → Dense10.
   - **Upgraded CNN**: Added batch normalization, extra convolutional filters, and data augmentation (rotation, shift, flip).
   - **Upgraded MLP**: Applied dropout and optional batch normalization.
3. **Training Setup**
   - Optimizer: Adam
   - Loss: Categorical cross-entropy
   - Batch size: 128
   - Epochs: 10–30 depending on model
   - Early stopping and learning rate reduction for upgraded models.

Results:

Model	Test Accuracy	Notes
Basic MLP	~85.7%	Struggles with visually similar clothing items (Shirt ↔ T-shirt, Pullover ↔ Coat).
Basic CNN	~88.2%	Better spatial feature extraction than MLP; fewer confusions.
Upgraded MLP	~86.1%	Dropout and batchnorm improve generalization slightly.

Model	Test Accuracy	Notes
Upgraded CNN	~92%	Data augmentation, batch normalization, and deeper architecture reduce confusion and overfitting.

Confusion Matrix Insights:

- Most confusions occur between visually similar items (e.g., Shirt ↔ T-shirt/top, Pullover ↔ Coat, Sneaker ↔ Ankle Boot).
- CNN models consistently reduce these confusions compared to MLP.

Key Findings

- **CNNs outperform MLPs** due to their ability to capture spatial features and local patterns in images.
- **Upgraded models** with data augmentation, dropout, and batch normalization significantly improve test accuracy and generalization.
- The **largest gains** come from adding convolutional layers, batch normalization, and augmentation, while MLP improvements are modest.

Recommendations

1. Use **data augmentation** and **dropout** for better generalization.
2. Increase CNN depth or filters to improve feature extraction.
3. Experiment with learning rate schedules, optimizer variants (AdamW), and ensembling for further performance gains.
4. Analyze confusion matrices to target similar-looking classes with specialized augmentation or attention modules.

Conclusion

- **Baseline models** provide a solid foundation for Fashion-MNIST classification.
- **Upgraded CNNs** achieve the highest accuracy (~92%) and demonstrate the advantage of convolutional feature extraction combined with regularization and augmentation.
- MLPs are limited in performance, but improvements with dropout and batch normalization are still beneficial.
- Systematic model upgrades and training strategies can maximize classification performance for similar image datasets.