

1. PENJELASAN MATRIKS EVALUASI KLASIFIKASI:

- ACCURACY: Proporsi prediksi yang benar dari total prediksi
- PRECISION: Proporsi prediksi positif yang benar ($\text{True Positive} / (\text{True Positive} + \text{False Positive})$)
- RECALL: Proporsi data positif yang berhasil diprediksi ($\text{True Positive} / (\text{True Positive} + \text{False Negative})$)
- F1-SCORE: Harmonic mean dari Precision dan Recall ($2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$)
- AUC (Area Under Curve): Area di bawah kurva ROC, mengukur kemampuan model memisahkan kelas

2. PENJELASAN MODEL:

TensorFlow Model:

- Arsitektur MLP dengan 4 hidden layers (256→128→64→32)
- Batch Normalization untuk stabilitas training
- Dropout layers (0.2-0.3) untuk regularisasi
- L1/L2 regularization untuk mencegah overfitting
- AdamW optimizer dengan weight decay
- ReduceLROnPlateau untuk adaptive learning rate
- Early stopping untuk mencegah overfitting
- Softmax activation untuk multi-class, sigmoid untuk binary

PyTorch Model:

- Arsitektur serupa dengan TensorFlow
- Manual implementation dari regularization techniques
- CosineAnnealingWarmRestarts untuk learning rate scheduling
- CrossEntropyLoss untuk classification

- Model checkpointing untuk menyimpan best model

3. PERBANDINGAN MODEL:

Model Terbaik: TensorFlow

- TensorFlow F1-Score: 0.9994
- PyTorch F1-Score: 0.9975
- TensorFlow unggul dengan selisih F1: 0.0019

4. MATRIKS TERBAIK UNTUK KLASIFIKASI:

F1-SCORE adalah matriks terbaik karena:

- Menggabungkan Precision dan Recall dalam satu metrik
- Baik untuk dataset yang tidak seimbang (imbalanced)
- Memberikan gambaran keseluruhan performa model
- Lebih robust dibanding accuracy untuk kasus imbalanced data

Namun, pemilihan metrik terbaik tergantung pada konteks:

- ACCURACY: Baik untuk balanced dataset
- PRECISION: Penting jika False Positive harus diminimalkan
- RECALL: Penting jika False Negative harus diminimalkan
- AUC: Baik untuk membandingkan model secara keseluruhan

5. ANALISIS CONFUSION MATRIX:

- Diagonal utama menunjukkan prediksi yang benar
- Off-diagonal menunjukkan kesalahan klasifikasi
- Dapat mengidentifikasi kelas mana yang sering salah diprediksi

6. REKOMENDASI:

- Kedua model menunjukkan performa yang baik
- Epochs 30 sudah cukup dengan early stopping
- Teknik regularisasi efektif mencegah overfitting
- Learning rate scheduling membantu konvergensi
- Batch normalization mempercepat dan menstabilkan training

7. KESIMPULAN AKHIR:

Model TensorFlow lebih baik dengan:

- Accuracy: 0.9994 vs 0.9983
- F1-Score: 0.9994 vs 0.9975