

# Dataset

---

Dataset EMNIST Balanced adalah himpunan data citra tulisan tangan yang dikembangkan dari NIST Special Database 19, lalu dikonversi ke format gambar 28×28 piksel grayscale, sehingga mirip dengan dataset MNIST. Bedanya, EMNIST mencakup lebih banyak kelas, tidak hanya angka tetapi juga huruf besar dan huruf kecil tertentu.

Dataset: EMNIST Balanced

- 47 kelas (huruf + angka)
- 112.800 sampel training
- 18.800 sampel testing
- 28×28 gambar grayscale

**Input** : Citra tulisan tangan 28×28 piksel grayscale, di-flatten menjadi vektor berukuran 784 fitur untuk dimasukkan ke dalam MLP (Artificial Neural Network). Nilai tiap piksel yang berada pada rentang 0–255 dinormalisasi menjadi 0–1.

**Output** : Label kelas target berupa angka atau huruf tertentu (total 47 kelas). Label ini kemudian di-one-hot encoding agar bisa diproses oleh model klasifikasi multi-kelas.

# CLASSIFICATION

## ANN MLP

	MLP
Akurasi	87.98%
Precision	88.13%
Recall	87.98%
F1-Score	87.92%

## Akurasi

Nilai akurasi menunjukkan seberapa banyak prediksi model yang benar dibandingkan dengan seluruh data uji. Akurasi = 0.8798 berarti sekitar 87.98% dari semua sampel berhasil diklasifikasikan dengan benar oleh model. Hampir 88% dari semua data huruf/tulisan tangan yang diuji berhasil diklasifikasikan ke kelas yang benar.

## Precision

Precision menunjukkan ketepatan prediksi model, yaitu dari semua prediksi yang dianggap benar oleh model, berapa persen yang memang benar sesuai kelas aslinya. Precision = 88,13% berarti ketika model bilang suatu gambar adalah huruf tertentu, sekitar 88% benar-benar huruf itu.

## Recall

Recall mengukur kemampuan model menemukan semua contoh dari suatu kelas. Recall = 87,98% berarti dari semua data asli untuk setiap huruf, model berhasil menemukan dan mengklasifikasikan sekitar 88% dengan benar, sisanya ada yang terlewat atau salah klasifikasi.

## F1-Score

F1-Score adalah rata-rata harmonis antara precision dan recall. F1-Score = 87,92% menunjukkan keseimbangan yang baik: model tidak hanya tepat (precision tinggi) tapi juga cukup menyeluruh (recall tinggi). Jadi, secara keseluruhan ini sudah cukup baik untuk MLP dengan dataset besar dan 47 kelas seperti EMNIST.

# Analisis ANN MLP Kelompok 16

## Model Architecture

Input Layer: Flatten(28x28 → 784 neurons)

Hidden Layers:

- Dense(1024), aktivasi ReLU + Batch Normalization + Dropout(0.4)
- Dense(768), aktivasi ReLU + Batch Normalization + Dropout(0.4)
- Dense(512), aktivasi Swish + Batch Normalization + Dropout(0.3)
- Dense(256), aktivasi Swish + Dropout(0.2)

Output Layer: Dense(47) + Softmax (47 possible classes)

## Training Strategy

- Optimizer → AdamW (Adaptive weight decay for fast convergence)
- Loss Function → Categorical Crossentropy (for multi-class classification)
- Batch Size → 128 (Balances speed and stability)
- Early Stopping → Stops when validation loss stagnates to avoid overfitting.
- Reduce Learning Rate on Plateau → Adjusts learning rate when validation loss stops improving.



