

## ▼ Kasus - Klasifikasi Tulisan Tangan dengan Dataset MNIST

### Deskripsi Umum

Anda diminta untuk melakukan klasifikasi dengan menggunakan algoritma Naive Bayes dan SVM untuk merekognisi tulisan tangan dari dataset MNIST

### Tentang Dataset MNIST

Dataset MNIST (Modified National Institute of Standards and Technology) merupakan dataset berupa citra grayscale dengan ukuran 28x28 yang berisi tulisan tangan dari digit angka 0-9. Jumlah data dalam dataset ini adalah 70.000 data.

### Detail Kuis

1. Pastikan jumlah data setiap label angka (0-9). Apakah terdapat data imbalance pada dataset MNIST yang digunakan? Jelaskan jawaban kelompok Anda dan buktikan!
2. Tampilkan 15 buah citra pertama dari dataset MNIST berserta labelnya. Anda dapat menggunakan referensi dari Job Sheet 03.
3. Lakukan proses ekstraksi fitur pada data MNIST. Pada proses ini, Anda diperbolehkan untuk
  - Langsung menggunakan nilai asli dari citra (nilai asli setiap pixel dari dataset) sebagai fitur.
  - Melakukan ekstraksi fitur lain seperti histogram, PCA, atau yang lainnya. Anda diperbolehkan melakukan ekspolarasi pada proses ini.
4. Buatlah data training dan data testing menggunakan rasio 70:30, 80:20, 90:10.
5. Lakukan proses klasifikasi dengan menggunakan algoritma Naive Bayes dan SVM.
  - Anda diperbolehkan untuk melakukan tunning parameter.
  - Anda diperbolehkan mengeksplorasi jenis kernel yang digunakan pada SVM.
6. Evaluasi model yang Anda buat terhadap data training dan data testing.
  - Gunakan metric *accuracy* untuk mengetahui tingkat akurasi data training dan data testing.
  - Gunakan fungsi *classification\_report* untuk mengetahui performa model secara holistik.
  - Gunakan confusion matrix untuk mengetahui tingkat ketepatan pelabelan.
7. Tampilkan citra data testing berserta hasil pelabelannya. Anda dapat menggunakan Job Sheet 03 sebagai acuan.
8. Apa model terbaik yang Anda dapatkan? Bagaimana konfigurasinya? Berapa tingkat akurasi yang didapatkan? Jelaskan!

## ▼ Bantuan

Berikut merupakan kode bantuan untuk proses mengunduh dataset MNIST

```
# Unduh Dataset MNIST
from sklearn.datasets import fetch_openml

mnist = fetch_openml('mnist_784')

# Baca deskripsi dataset
# Tahap ini mungkin diperlukan untuk lebih memahami dataset
print(mnist.DESCR)

# Inspeksi semua key dalam dataset
# Cara ini digunakan untuk memahami struktur dari dataset MNIST
mnist.keys()
```