



# **Analisis Kesejahteraan Mental Menggunakan Machine Learning pada Data DigitalExosome**

Ujian Tengah Semester Machine Learning

Fauzan Azhima | 105222003



STUDI KASUS 3

# Latar Belakang

- Kesejahteraan mental dipengaruhi oleh banyak faktor, termasuk lingkungan perkotaan.
- Kualitas udara dan respon fisiologis dapat mencerminkan tingkat stres atau kebahagiaan individu.
- Dataset DigitalExosome menyediakan data lingkungan + fisiologi + label kesejahteraan secara real-time.



> Fauzan Azhima

# Permasalahan dan Tujuan



Fauzan Azhima

## Permasalahan

Bagaimana memprediksi tingkat kesejahteraan mental dari data lingkungan dan fisiologis secara real-time?

## Tujuan

Mengklasifikasikan skor kesejahteraan mental (1–5) dengan menggunakan model machine learning supervised.



# Deskripsi Dataset



Total :  
42.436 entri dari 40  
partisipan

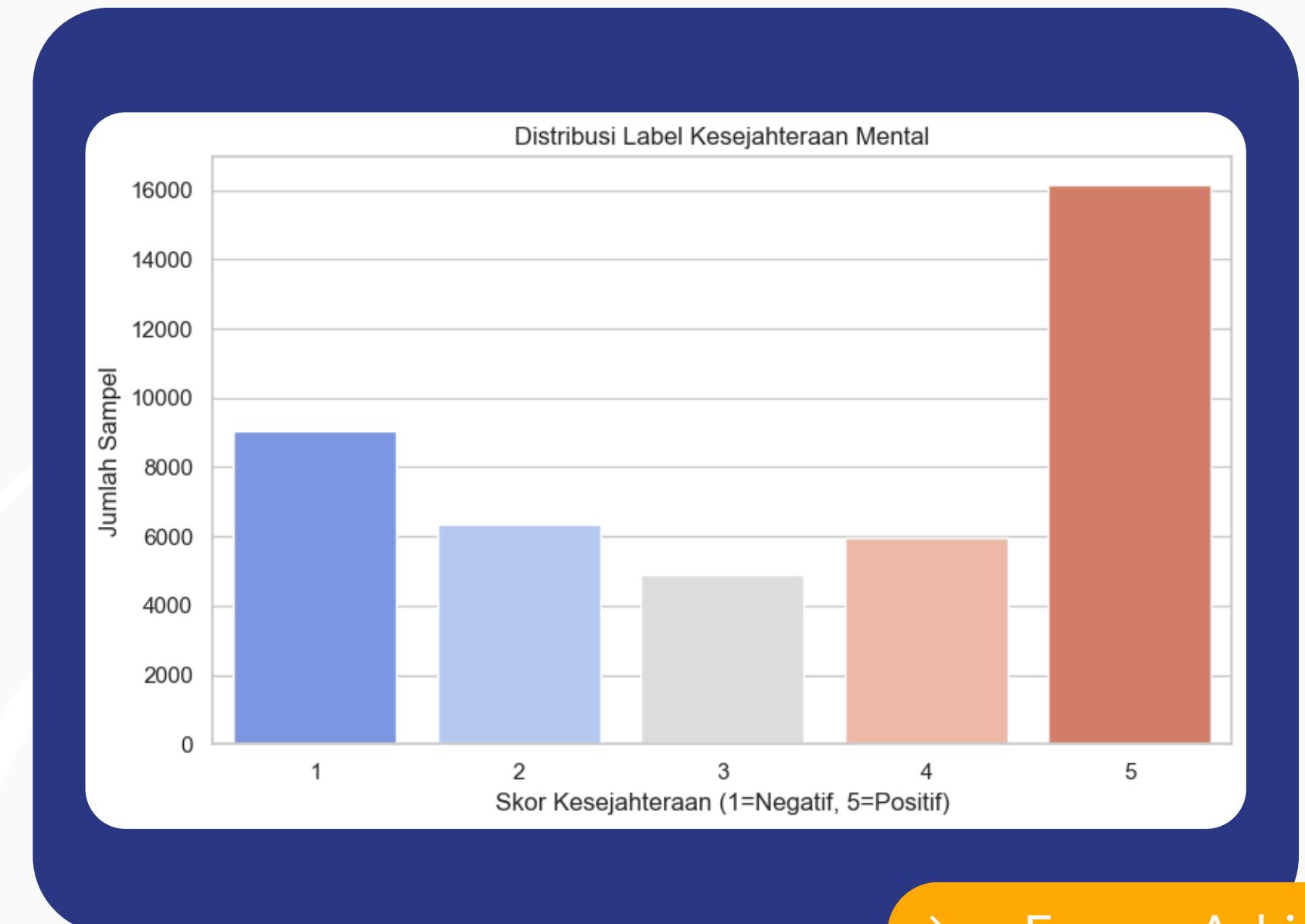
- Fitur:**
- Fisiologis: HR, HRV (IBI), EDA, BVP
  - Lingkungan: PM1, PM2.5, PM10, CO, NO2, NH3, Noise
  - Target: Label kesejahteraan (1-5)

> Fauzan Azhima

Semua data sudah  
dinormalisasi (0-1) dan  
disamakan sampling rate-nya (1Hz)

# EDA dan Kolerasi

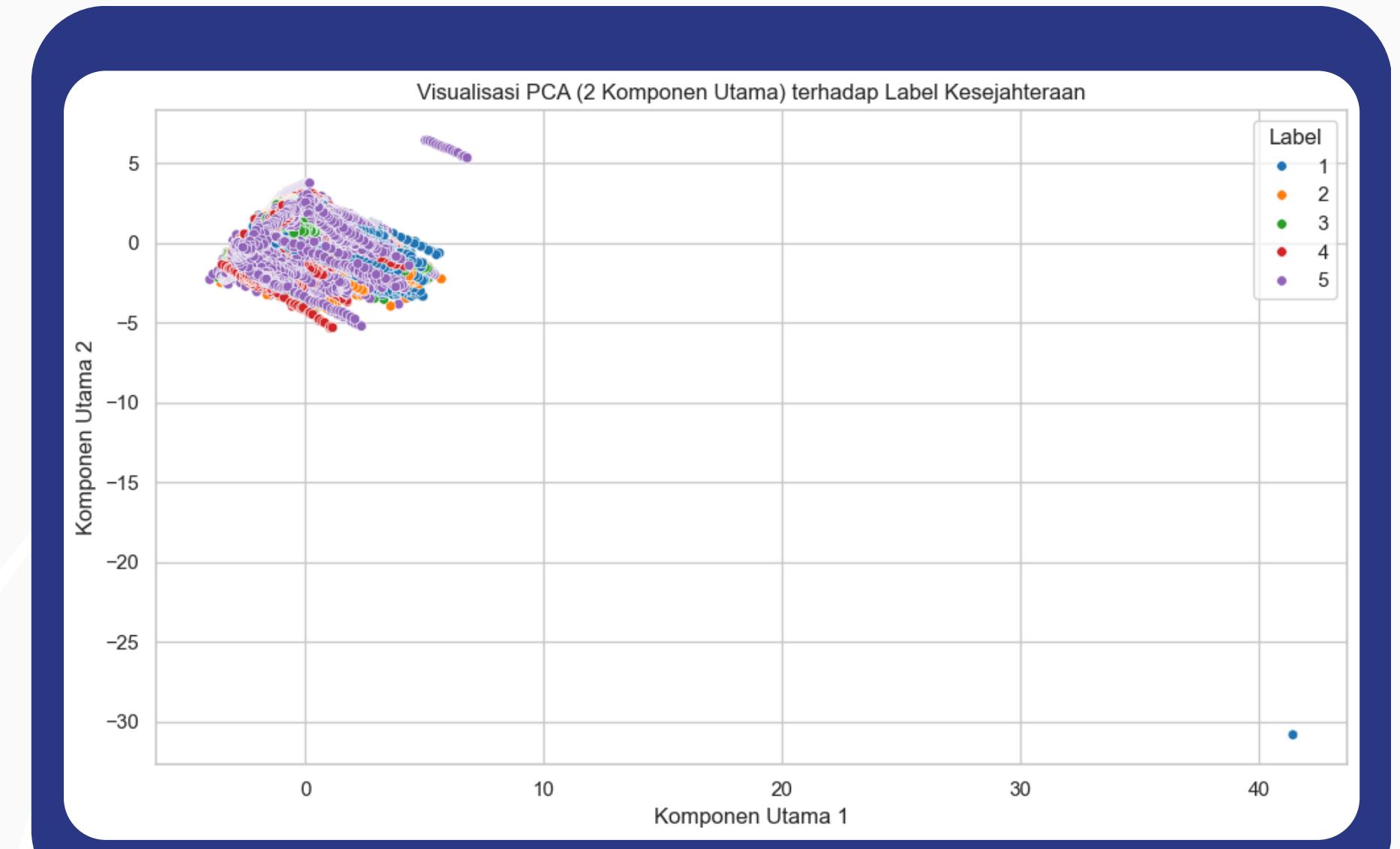
- Distribusi label menunjukkan mayoritas peserta merasa cukup hingga sangat bahagia (label 4–5)]
- Korelasi:
  - HR (+), EDA (+), CO (-), HRV (IBI) (-)
  - Faktor fisiologis lebih berpengaruh daripada faktor lingkungan



> Fauzan Azhima

# Visualisasi PCA

- PCA digunakan untuk reduksi dimensi ke 2 komponen utama
- Visualisasi scatter menunjukkan pemisahan jelas antara label 1 dan 5
- Label 2–4 tumpang tindih (klasifikasi lebih menantang)



> Fauzan Azhima

# Preprocessing

- Normalisasi Min-Max (dari peneliti)
- Resampling dan interpolasi ke 1Hz
- Tidak ada missing value berarti
- Label sudah berbentuk integer (1–5), tidak perlu encoding tambahan



> Fauzan Azhima

# Splitting Data

- 70% training: 29.705 sampel
- 30% testing: 12.731 sampel
- Stratifikasi digunakan agar distribusi label tetap seimbang



> Fauzan Azhima

# Model Machine Learning

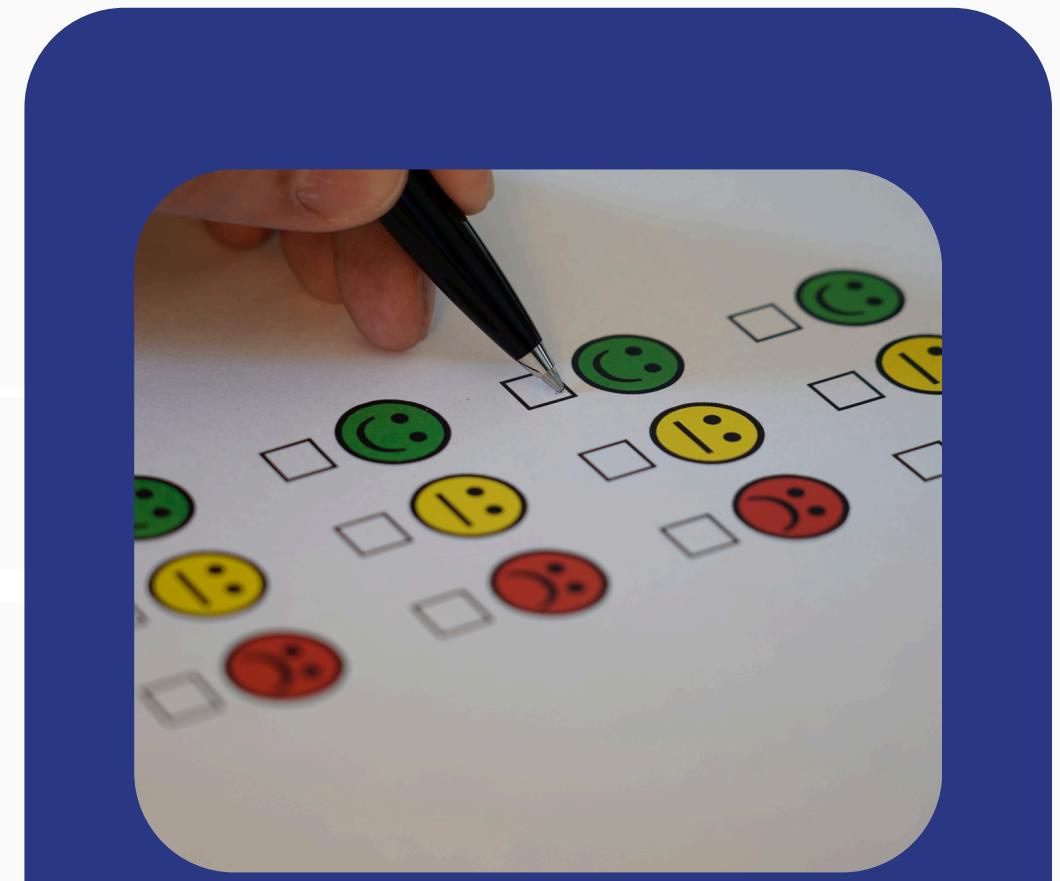
- Decision Tree Classifier: `max_depth = 8`
- Random Forest Classifier: 100 trees,  
`max_depth = 10`
- k-Nearest Neighbors (k-NN): `k = 5`,  
Euclidean distance



> Fauzan Azhima

# Evaluasi Model

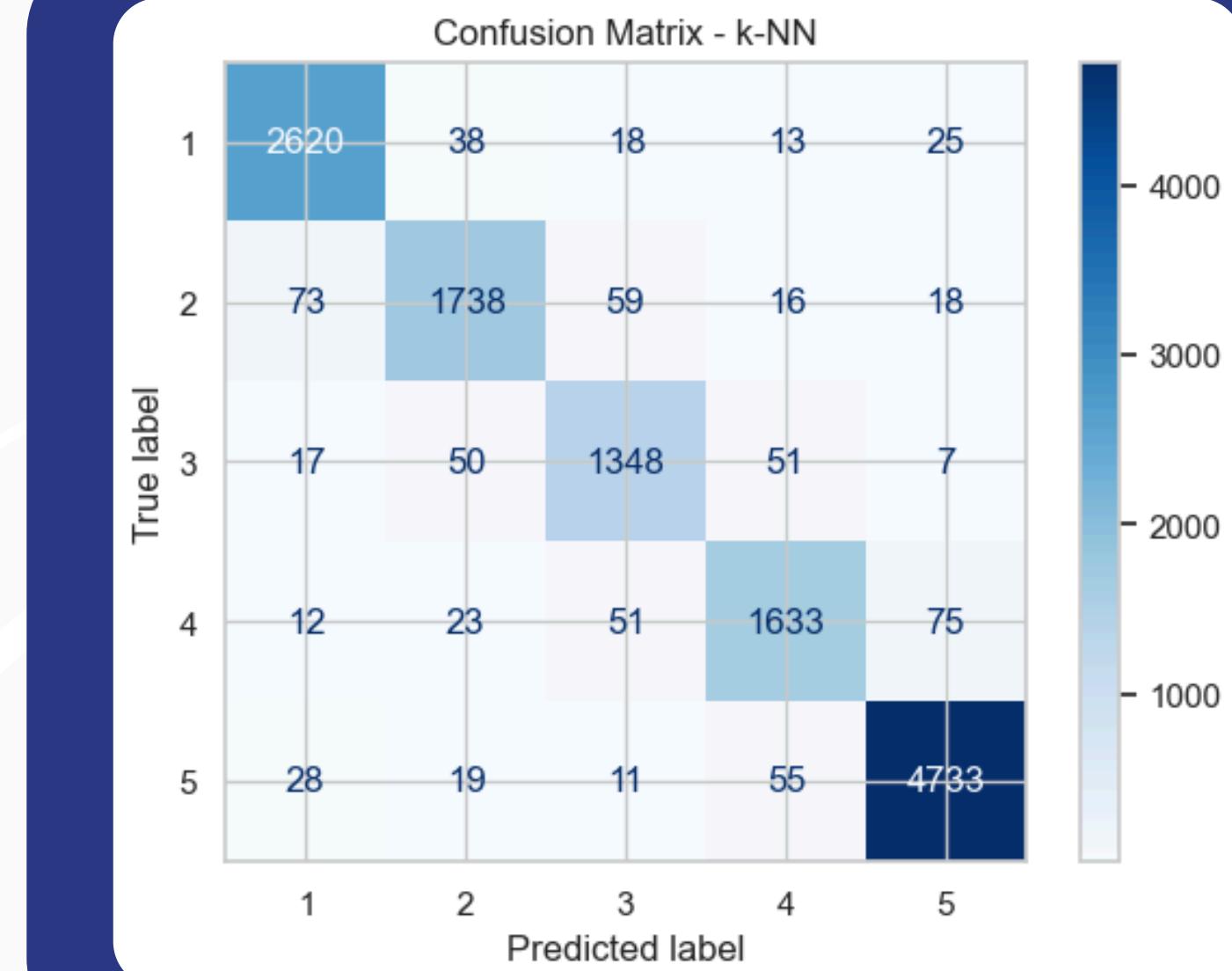
- Decision Tree: Akurasi 71.6%
- Random Forest: Akurasi 90.1%
- k-NN: Akurasi 94.8% → performa terbaik
- Model diuji menggunakan classification report (precision, recall, f1-score)



> Fauzan Azhima

# Confusion Matrix

- Confusion matrix menunjukkan:
  - Prediksi akurat pada label 1 dan 5
  - Kesalahan paling banyak antara label 2–4



# Feature Importance

- Feature importance dari Random Forest:
  - Paling berpengaruh: HR, CO, EDA, HRV
  - Insight: Fitur fisiologis lebih menentukan prediksi dibanding polusi udara



> Fauzan Azhima

# Kesimpulan

- Data DigitalExosome efektif untuk klasifikasi kesejahteraan mental
- k-NN menghasilkan akurasi tertinggi (94.8%)
- Insight:
  - Fitur fisiologis dominan → fokus pada wearable lebih efektif



# Terima Kasih

Ujian Tengah Semester Machine Learning  
Fauzan Azhima | 105222003

Analisis Kesejahteraan Mental  
Menggunakan Machine Learning  
pada Data DigitalExosome



Studi Kasus 3

