

**PENGAYAAN
PENGENALAN POLA**



Disusun oleh Kelompok 1 :

- | | |
|------------------------------|--------------------|
| 1. Farkhan | 20081010060 |
| 2. Azka Avicenna R | 20081010115 |
| 3. Refansya Rachmad A | 20081010006 |
| 4. Raviy Bayu Setiaji | 19081010010 |

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA
TIMUR
2022**

1. Regresi adalah suatu teknik analisis untuk mengidentifikasi hubungan di antara dua variabel atau lebih. Regresi termasuk ke dalam *supervised learning* yang digunakan untuk memprediksi nilai kontinu. Klasifikasi adalah sebuah teknik untuk mengklasifikasikan atau mengkategorikan beberapa item yang belum berlabel ke dalam sebuah set kelas diskrit. Klasifikasi mempelajari hubungan antara kumpulan variabel fitur dan variabel target yang bertipe kategori. Klasterisasi termasuk *unsupervised learning*. Klaster merupakan kumpulan data atau objek yang memiliki kemiripan satu sama lain di dalam kumpulan atau kelompok tersebut, dan beberapa objek di kelompok lain. Berbeda dengan klasifikasi, data yang dilatih pada klasterisasi tidak memiliki label. Dari penjelasan sebelumnya, maka secara singkat, berikut ini perbedaan antara regresi, klasifikasi, dan klasterisasi.

Regresi : *supervised learning*, prediksi data kontinu

Klasifikasi : *supervised learning*, prediksi data diskrit, memiliki label

Klasterisasi : *unsupervised learning*, tidak memiliki label

2. Pada paper kelompok kami, permasalahan yang diangkat ialah melakukan pendeteksian wajah menggunakan metode viola-jones. Metode viola-jones merupakan pendekatan baru yang dapat melakukan deteksi objek dengan tingkat akurasi yang tinggi serta kemampuan berjalan 15 frame per detik. Metode viola-jones menggunakan fitur Haar sebagai deskriptor kemudian menggabungkan *integral image* dan *AdaBoost* untuk mencari dan melakukan seleksi nilai fitur dan membentuk *cascade classifier*. *Classifier* tersebut yang akan digunakan untuk mendeteksi wajah pada gambar. Penggunaan *classifier* pada metode viola-jones menunjukkan bahwa metode viola-jones merupakan metode klasifikasi.
3. Algoritma K-Means adalah algoritma iteratif yang mencoba mempartisi kumpulan data ke dalam subkelompok (klaster) berbeda yang tidak tumpang tindih yang telah ditentukan sebelumnya di mana setiap titik data hanya dimiliki oleh satu kelompok. Ini mencoba membuat titik data intra-cluster semirip mungkin sambil juga menjaga klaster sejauh mungkin berbeda. Ini menetapkan titik data ke sebuah cluster sedemikian rupa sehingga jumlah jarak kuadrat antara titik data dan centroid cluster (rata-rata aritmatika dari semua titik data yang termasuk dalam klaster itu) adalah minimum. Semakin sedikit variasi yang kita miliki dalam klaster, semakin homogen (mirip) titik data berada dalam klaster yang sama.

Cara kerja algoritma kmeans adalah sebagai berikut:

- a) Tentukan jumlah kluster K.
- b) Inisialisasi centroid dengan terlebih dahulu mengacak dataset dan kemudian secara acak memilih K titik data untuk centroid tanpa penggantian.
- c) Terus iterasi sampai tidak ada perubahan pada centroid. yaitu penetapan titik data ke cluster tidak berubah.
- d) Hitung jumlah kuadrat jarak antara titik data dan semua centroid.
- e) Tetapkan setiap titik data ke kluster terdekat (centroid).
- f) Hitung centroid untuk kluster dengan mengambil rata-rata dari semua titik data yang dimiliki setiap kluster.

Algoritma klasterisasi (K-Means) memiliki tujuan untuk meminimalisasikan fungsi objective yang telah diset dalam proses klasterisasi. Tujuan tersebut dilakukan dengan cara meminimalkan variasi data yang ada di dalam cluster dan memaksimalkan variasi data yang ada di cluster lainnya. Contoh penerapan algoritma klasterisasi dapat ditemukan pada segmentasi customer bank atau segmentasi berita-berita online. Bisa juga digunakan untuk menentukan parameter jumlah data, kluster, dan atribut dalam penjurusan siswa.