

BAB I PENDAHULUAN

1.1 Latar Belakang

Bumi terdiri dari berbagai benua yang di dalamnya terdapat suatu Negara dan kepulauan. Setiap Negara maupun pulau dihuni oleh suku yang berbeda-beda. Hal ini kemudian menyebabkan adanya keberagaman suku bangsa antar Negara maupun daerah. keberagaman ini bukan hanya mencakup suku bangsa saja, tetapi juga budaya dan ras. Ras merupakan pengelompokan manusia berdasarkan ciri fisik atau biologis yang diwariskan, serta tempat asal biologis keturunan (generasi). Ras seseorang dapat menentukan daerah asli tempat seseorang itu berasal, serta menentukan garis keturunannya. Ras bisa dibedakan dan diidentifikasi dengan spesifik, seperti bentuk mata, bentuk hidung, warna kulit, dan lain-lain. Ada beberapa macam ras yang tersebar di berbagai Negara dan daerah, seperti ras Mongoloid, Negroid dan Kaukasoid.

Seiring berjalannya waktu, populasi penduduk dunia semakin bertambah dan tersebar di berbagai Negara dan daerah. Persebaran ini dikarenakan terjadinya transmigrasi dari berbagai Negara maupun daerah ke Negara atau daerah tertentu. Transmigrasi ini disebabkan oleh beberapa faktor, diantaranya adalah terjadinya penjajahan dan perebutan kekuasaan pada zaman dahulu kala, kebutuhan ekonomi dan sosial, pengetahuan, pekerjaan, dan lain-lain. Sehingga mereka menetap, berkeluarga, beranak-pinak dan menjadi warga Negara di Negara maupun di daerah tersebut. Persebaran penduduk ini menjadikan ras tersebar di berbagai Negara dan daerah.

Pada saat ini, teknologi berkembang sangat pesat, terutama di bidang Teknologi Informasi. *Trending* pada bidang teknologi informasi saat ini adalah *machine learning* atau *deep learning*. Maka dari itu, kasus penelitian ini akan mengimplementasikan sistem *machine learning*. Banyak penelitian yang menggunakan metode klasifikasi untuk mengklasifikasikan objek tertentu. Metode klasifikasi memiliki kelemahan, yaitu membutuhkan data *training* (perlu dilatih terlebih dahulu) agar dapat melakukan klasifikasi. Sedangkan metode *clustering* tidak membutuhkan data training terlebih dahulu, karena metode

clustering merupakan *unsupervised learning*. Metode *clustering* merupakan metode pengelompokan data berdasarkan karakteristik yang ditemukan ke dalam beberapa *cluster* atau kelompok, sehingga data dalam satu *cluster* memiliki tingkat kesamaan yang maksimum dan data antar *cluster* memiliki tingkat kemiripan yang minimum. Maka dari itu, penelitian ini menggunakan metode *clustering* untuk mengelompokkan ras. Ada beberapa algoritma untuk metode *clustering*, salah satunya adalah algoritma *K-Means*. Algoritma *K-Means* merupakan salah satu algoritma *clustering* dengan metode partisi (*partition method*) yang berbasis titik pusat (*centroid*) dan memiliki kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien. Maka dari itu, penelitian ini lebih tepat menggunakan algoritma *k-means*.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, dirumuskan masalah yaitu bagaimana langkah mengelompokkan ras dengan menerapkan *machine learning* menggunakan teknik *clustering* dengan algoritma *k-means*.

1.3 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah untuk mengelompokkan ras menggunakan metode *clustering* dengan algoritma *k-means*. Adapun manfaatnya yaitu, menambah wawasan dengan mengetahui bagaimana ciri fisik dari setiap ras yang berbeda, membangun rasa toleransi terhadap perbedaan fisik seseorang, dan mengetahui keragaman ras dapat membantu kita menjalin hubungan sosial dengan baik.

1.4 Batasan Masalah

Agar pengerjaan tugas akhir ini menjadi lebih terarah dan mendapatkan hasil yang lebih spesifik, maka pembahasan dibatasi pada ruang lingkup sebagai berikut :

1. Metode yang digunakan adalah metode *clustering* dengan algoritma *k-means*
2. Mengelompokkan ras mongoloid, kaukasoid, dan negroid
3. Dataset yang digunakan adalah foto wajah manusia yang terdiri dari tiga macam ras secara acak

4. Dataset yang digunakan berjumlah 300 foto secara acak yang terdiri dari 100 foto dari masing-masing ras
5. Bahasa pemrograman yang digunakan adalah *Python*
6. *Input* berupa folder yang berisi foto wajah manusia
7. *Output training* berupa foto wajah manusia yang sudah dikelompokkan ke dalam folder berdasarkan kesamaan dari masing-masing ras
8. *Testing* berupa suatu foto wajah manusia

1.5 Metodologi Penelitian

Metodologi yang digunakan pada penelitian ini yaitu :

1. Pengumpulan data

Dataset yang digunakan pada pengelompokkan ras ini berupa folder yang berisi foto-foto dari beberapa macam ras, yaitu Kaukasoid, Mongoloid, dan Negroid yang berformat .jpg. Langkah pertama dalam pembuatan dataset ini adalah mengunduh dataset dari website (<https://www.kaggle.com/zuruoke/race-classification?select=Racist>).

Selanjutnya memilih 100 foto dari masing-masing ras dan menyimpannya ke dalam satu folder. Kemudian dilakukan tahap *preprocessing* dengan melakukan pemotongan pada setiap foto yang difokuskan pada area wajah. Hal ini dilakukan untuk menerapkan *feature extraction* pada objek area wajah agar hasil pengelompokkan lebih optimal. Jadi, jumlah dataset yang digunakan berjumlah 300 foto, yang terdiri dari 100 foto dari masing-masing ras.

2. Analisis

Proses ini dilakukan analisis sistem menggunakan metode *clustering* dengan Algoritma K-Means.

3. Perancangan

Pada tahap ini, perancangan sistem *clustering* ras menggunakan bahasa pemrograman *python* dengan *code editor Google Collaboratory*.

1.6 State of the Art

Dalam penyusunan tugas akhir ini, mengambil beberapa referensi materi dari penelitian sebelumnya, termasuk jurnal yang berkaitan dengan penelitian ini. Referensi dari beberapa jurnal beserta ringkasannya sebagai berikut ini :

Tabel 1. 1 State of the Art

No	Jurnal	Ringkasan	Perbedaan
1	<p>Judul : Penerapan Algoritma K-Means Untuk Clustering Data Obat-obatan pada RSUD Pekanbaru</p> <p>Oleh : Gustientiedina, M. Hasmil Adiya, dan Yenny Desnelita</p> <p>Tahun : 2019</p>	<p>Jurnal tersebut membahas tentang pengelompokkan data obat-obatan di RSUD Pekanbaru berdasarkan kebutuhan pemakaian obat-obatan tersebut. Dataset nya berupa nama obat-obatan dengan atribut pemakaian. Pengelompokkan ini menerapkan metode Clustering dengan Algoritma K-Means</p>	<p>Perbedaan jurnal tersebut dengan penelitian ini adalah dataset yang digunakan pada penelitian ini berupa foto atau gambar. Diterapkan dengan <i>Machine Learning</i></p>
2	<p>Judul : Analisis K-Means Clustering pada Data Sepeda Motor</p> <p>Oleh : Rozzi Kesuma Dinata, Safwandi, Novia Hasdyna, dan Nur Azizah</p> <p>Tahun : 2020</p>	<p>Jurnal tersebut membahas tentang pengelompokkan data sepeda motor menjadi tiga kelompok, yaitu murah, standar, dan mahal. Pengelompokkan ini menerapkan metode Clustering dengan Algoritma K-Means</p>	<p>Perbedaan jurnal tersebut dengan penelitian ini adalah dataset yang digunakan pada penelitian ini berupa foto atau gambar. Diterapkan dengan <i>Machine Learning</i></p>

3	<p>Judul : Penerapan Algoritma K-Means Untuk Pengelompokan Diagnosa Penyakit Mata Berdasarkan Rentang Usia</p> <p>Oleh : Fitria Kurnia, Ichsan Fahmi, Erwin Wahyudi, dan Godlief E.S Mige</p> <p>Tahun : 2019</p>	<p>Jurnal tersebut membahas tentang pengelompokan diagnosa penyakit mata berdasarkan rentang usia. Membandingkan hasil perhitungan antara menggunakan aplikasi sistem dengan menggunakan <i>tools Rapidminer</i></p>	<p>Perbedaan jurnal tersebut dengan penelitian ini adalah dataset yang digunakan pada penelitian ini berupa foto atau gambar. Kemudian tidak melakukan perbandingan hasil perhitungan dengan <i>tools Rapidminer</i></p>
4	<p>Judul : Penerapan K-Means Clustering dalam Pengelompokan Lipstik Matte Lip Cream Berdasarkan Warna RGB</p> <p>Oleh : Ananda Mulia Alhumaerah, Arief Bramanto Wicaksono Putra, dan Rihartanto</p> <p>Tahun : 2019</p>	<p>Jurnal tersebut membahas tentang pengelompokan Lipstik Mate Lip Cream berdasarkan warna RGB dengan menggunakan K-Means Clustering di <i>Matlab</i></p>	<p>Perbedaan jurnal tersebut dengan penelitian ini adalah bahasa yang digunakan pada penelitian ini yaitu bahasa <i>Python</i> dengan <i>code editor Google Collaboratory</i>. Kemudian tidak mengkonversikan nya ke <i>grayscale</i></p>
5	<p>Judul : Algoritma K-Means Untuk Segmentasi Kematangan Buah Jeruk Berdasarkan Kemiripan Warna</p> <p>Oleh : Mhd Furqan, Sriani, dan</p>	<p>Jurnal tersebut membahas tentang penerapan Algoritma K-Means dalam segmentasi citra buah jeruk</p>	<p>Perbedaan jurnal tersebut dengan penelitian ini adalah pada penelitian ini tidak melakukan segmentasi citra.</p>

	Atiqah Aulia Tahun : 2022	untuk menentukan kematangan berdasarakan warna.	
--	----------------------------------	---	--

1.7 Sistematika Penulisan

Secara garis besar, sistematika penulisan tugas akhir ini tersusun sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini membahas tentang latar belakang, rumusan masalah, tujuan dan manfaat, batasan masalah, *State of the art*, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini membahas tentang tinjauan pustaka yang berkaitan dengan tugas akhir, seperti *machine learning*, *clustering*, dan algoritma *k-means*.

BAB III METODE PENELITIAN

Pada bab ini membahas tentang analisis masalah kebutuhan sistem, pengumpulan data dan perancangan.

BAB IV IMPLEMENTASI

Pada bab ini membahas tentang pengujian dan hasil dari sistem untuk mengelompokkan ras menggunakan algoritma *k-means*.

BAB V PENUTUP

Pada bab ini berisi tentang kesimpulan dan saran dari hasil penelitian yang telah dilakukan.