

**PENERAPAN K-MEANS DAN ALGORITMA GENETIKA UNTUK  
MENYELESAIKAN MTSP**

(Studi Kasus Pada Perjalanan Menuju Seluruh SMA di Kabupaten Probolinggo)

**SKRIPSI**



**OLEH:**

**MUHAMMAD FAIZ NAILUN NI'AM**

NIM : 1842200034

UNIVERSITAS NURUL JADID

PAITON PROBOLINGGO

**FAKULTAS SOSIAL DAN HUMANIORA**

PROGRAM STUDI PENDIDIKAN MATEMATIKA

JUNI 2022

**PENERAPAN K-MEANS DAN ALGORITMA GENETIKA UNTUK  
MENYELESAIKAN MTSP**

(Studi Kasus Pada Perjalanan Menuju Seluruh SMA di Kabupaten Probolinggo)

**SKRIPSI**

DIAJUKAN KEPADA UNIVERSITAS NURUL JADID  
PAITON PROBOLINGGO UNTUK MENYELESAIKAN  
SALAH SATU PERSYARATAN DALAM MENYELESAIKAN  
PROGRAM SARJANA PENDIDIKAN MATEMATIKA

**OLEH :**

**MUHAMMAD FAIZ NAILUN NI'AM**

NIM : 1842200034

UNIVERSITAS NURUL JADID  
PAITON PROBOLINGGO  
**FAKULTAS SOSIAL DAN HUMANIORA**  
PROGRAM STUDI PENDIDIKAN MATEMATIKA

JUNI 2022

## **ABSTRAK**

#MTSP isi abstrak

## KATA PENGANTAR

Segenap puji dan syukur kami panjatkan kepada Tuhan Yang Maha Esa karena atas berkat, karunia, dan rahmat yang diberikan oleh-Nya, sehingga penulis berkesempatan untuk menyelesaikan penelitian skripsi ini yang berjudul "Penerapan *K-means* dan Algoritma Genetika untuk Menyelesaikan MTSP (Studi Kasus pada Perjalanan Menuju Seluruh SMA di Kabupaten Probolinggo)".

Dalam penelitian dan penyusunan laporan ini, telah mendapatkan banyak bantuan, dukungan, bimbingan, serta do'a yang sangat berharga sehingga penulisan skripsi ini dapat berjalan sesuai dengan waktu yang telah ditentukan. Oleh karena itu kami ingin mengucapkan terima kasih yang sebesar-besarnya kepada pihak yang telah memberikan dukungan dan membantu selama penyusunan skripsi ini.

1. Keluarga terutama orang tua yang telah memberikan semangat dan motivasi dalam menyelesaikan penelitian ini.
2. Bapak Nur Hamid, M.Si, Ph.D., dan Ibu Shofia Hidayah, M.Pd., selaku dosen pembimbing yang telah meluangkan waktu dan tenaganya serta memberikan bimbingan, wawasan, dan ilmu yang sangat berharga dalam menyelesaikan penelitian ini.
3. Ibu [[#nama bu olief]] selaku dosen penguji proposal yang telah memberikan masukan, nasehat, dan perbaikan pada proposal penelitian yang sebelumnya telah dibuat.
4. Seluruh teman-teman dari Prodi Matematika angkatan 2018 sebagai teman seperjuangan dalam menyelesaikan skripsi yang juga saling memberikan semangat dan dukungan dalam menyelesaikan skripsi.
5. Semua pihak yang secara langsung maupun tidak langsung terlibat dalam membantu penulis menyelesaikan skripsi ini.

Kami menyadari bahwa dalam penelitian ini masih terdapat kekurangan dan jauh dari kata sempurna. Oleh karena itu kami menerima dengan baik segala kritik dan masukan yang dapat membangun bagi kami. Kami berharap skripsi ini dapat memberi manfaat bagi banyak pihak.

Probolinggo, 26 Mei 2022

Penulis

## DAFTAR ISI

<b>I</b>	<b>PENDAHULUAN</b>	<b>1</b>
A.	Latar Belakang Masalah . . . . .	1
B.	Rumusan Masalah . . . . .	2
C.	Tujuan Penelitian dan Pengembangan . . . . .	2
D.	Manfaat Penelitian . . . . .	3
E.	Batasan Masalah Penelitian . . . . .	3
<b>II</b>	<b>KAJIAN PUSTAKA</b>	<b>5</b>
A.	Penelitian Relevan . . . . .	5
B.	Dasar Teori . . . . .	6
1.	<i>Multiple Traveling Salesman Problem</i> . . . . .	6
2.	Algoritma . . . . .	7
3.	Algoritma <i>k</i> -means . . . . .	7
4.	Centroid . . . . .	8
5.	Algoritma Genetika . . . . .	8
6.	Fitness . . . . .	9
7.	<i>Crossover</i> . . . . .	9
8.	Mutasi . . . . .	10
<b>III</b>	<b>METODE PENELITIAN</b>	<b>11</b>
A.	Model Penelitian dan Pengembangan . . . . .	11
B.	Prosedur Penelitian dan Pengembangan . . . . .	12
1.	Data Penelitian . . . . .	12
2.	Instrumen Pendukung . . . . .	12
3.	Langkah-langkah Dalam Tahap Pengolahan Data . . . . .	14

<b>IV HASIL</b>	<b>17</b>
A. Penyajian Data Uji Coba . . . . .	17
1. Pengambilan Data Lokasi . . . . .	17
2. Proses Pengklasteran Data . . . . .	19
3. Proses TSP menggunakan Algoritma Genetika . . . . .	20
4. Hasil dari proses <i>k</i> -means dan Algoritma Genetika . . . . .	21
5. Percobaan menggunakan banyak klaster yang berbeda . . . . .	22
<b>V PENUTUP</b>	<b>26</b>
A. Kesimpulan . . . . .	26
B. Saran . . . . .	26
<b>1 DATASET</b>	<b>30</b>
A. Nama dan Koordinat SMA di Kabupaten Probolinggo . . . . .	30

## DAFTAR GAMBAR

II.1	Contoh MTSP 6 klaster . . . . .	6
II.2	Contoh MTSP 5 klaster . . . . .	7
II.3	Proses crossover . . . . .	10
II.4	Proses mutasi . . . . .	10
III.1	Visualisasi data jupyter notebook . . . . .	13
III.2	Menandai lokasi pada Google Earth . . . . .	14
III.3	Mengekspor data ke bentuk spreadsheet . . . . .	14
IV.1	Visualisasi lokasi SMA di Kabupaten Probolinggo . . . . .	18
IV.2	Visualisasi Centroid . . . . .	19
IV.3	Visualisasi klaster sesuai warna . . . . .	20
IV.4	Visualisasi hasil pembagian klaster dan rute yang dapat dilalui . . . . .	22
IV.5	Hasil rute dengan pembagian klaster berbeda . . . . .	24



## DAFTAR TABEL

IV.1	Koordinat titik-titik centroid . . . . .	18
IV.2	Urutan perjalanan pada klaster A . . . . .	22
IV.3	Total jarak dengan jumlah pembagian klaster yang berbeda . . . . .	25
1.1	Daftar nama-nama SMA di Kabupaten Probolinggo . . . . .	33

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang Masalah**

Kabupaten Probolinggo adalah salah satu dari beberapa kabupaten yang sedang berkembang di provinsi Jawa Timur. Banyak sekolah tingkat menengah yang tersebar di Kabupaten Probolinggo. Selain itu di Kabupaten Probolinggo terdapat beberapa kampus salah satunya adalah Universitas Nurul Jadid yang terletak di Kecamatan Paiton. Pada tahun-tahun sebelumnya kampus ini sering sekali mengadakan acara-acara besar seperti lomba dan olimpiade. Dalam acara-acara tersebut seringkali melakukan pendistribusian barang seperti undangan acara, pamflet, dan lain-lain kepada beberapa sekolah di Kabupaten Probolinggo. Oleh karena itu diperlukanlah sebuah pencarian rute yang efisien untuk menuju ke sekolah-sekolah tersebut agar dapat menghemat waktu dan tenaga dalam perjalanan. Permasalahan pencarian rute tersebut dalam hal ini dapat disebut dengan *Traveling Salesman Problem* (TSP). Sedangkan gabungan dari beberapa permasalahan TSP disebut *Multiple Traveling Salesman Problem* (MTSP).

Selama bertahun-tahun, telah banyak penelitian tentang MTSP. Berbagai metode telah banyak digunakan untuk mencari solusi dari permasalahan MTSP salah satunya adalah Algoritma Genetika (AG) dan *K-means*. Untuk melakukan proses pencarian solusi MTSP diperlukanlah proses pengklasteran (pengelompokan) terlebih dahulu, ada banyak cara untuk menggunakan AG dalam pengklasteran, terbukti bahwa metode ini dapat mengklaster data lebih cepat daripada beberapa algoritma lain yang digunakan untuk pengklasteran [10]. Kemampuan pengklasteran dari AG ini dimanfaatkan untuk mencari pusat klaster yang sesuai sehingga kesamaan dari klaster yang dihasilkan dioptimalkan [13]. Ada juga yang menggunakan metode paralel untuk TSP untuk meningkatkan efisiensi seperti pada artikel [11].

Namun, menurut artikel Zhang efisiensi AG akan menurun dengan cepat jika digunakan pada skala kota besar [22], berbeda dengan algoritma  $k$ -means dapat mengkluster terlebih dahulu sebelum melakukan pencarian solusi dari permasalahan TSP dan menghindari persilangan tiap rute salesman (pengantar barang) [12]. Penggunaan Algoritma Genetika dan algoritma  $k$ -means, algoritma ini merupakan algoritma yang digunakan untuk membagi data MTSP menjadi beberapa kluster, metode ini efektif untuk menyelesaikan MTSP, selain itu juga dapat menghindari persilangan rute antar salesman seperti yang dibahas oleh Lu pada artikelnya [12]. Dari gabungan semua perspektif tersebut, dalam penelitian ini, digunakanlah  $k$ -means dan Algoritma Genetika untuk menyelesaikan kasus pembagian kluster dan pencarian rute terdekat menuju seluruh SMA di Kabupaten Probolinggo.

## **B. Rumusan Masalah**

Berdasarkan latar belakang, maka rumusan masalah yang akan dikaji dalam penelitian ini sebagai berikut.

1. Bagaimana cara mencari solusi *multiple traveling salesman problem* dengan  $k$ -means dan algoritma genetika?
2. Bagaimana pembagian kluster dan penentuan rute terdekat menuju seluruh SMA di Kabupaten Probolinggo?

## **C. Tujuan Penelitian dan Pengembangan**

Berdasarkan latar belakang dan rumusan masalah, tujuan dari penelitian ini yaitu untuk.

1. Mengetahui cara menemukan solusi *multiple traveling salesman problem* dengan  $k$ -means dan algoritma genetika.

2. Menemukan solusi pembagian klaster dan penentuan rute terdekat menuju seluruh SMA di Kabupaten Probolinggo.

#### **D. Manfaat Penelitian**

Manfaat dari penelitian ini yaitu:

1. Bagi Penulis, mengetahui cara menyelesaikan kasus permasalahan *Multiple Traveling Salesman Problem* yang telah dipelajari yaitu dengan menggunakan metode *K-Means Clustering* dan Algoritma Genetika serta penulis dapat mengembangkan ilmu pemrograman python pada komputer.
2. Bagi Program Studi Pendidikan Matematika, menambah ilmu mengenai metode optimasi dan pencarian rute terdekat yang dapat diterapkan serta dipelajari kembali oleh mahasiswa pendidikan matematika untuk tahun-tahun selanjutnya, serta mengetahui rute-rute terdekat untuk menuju ke seluruh SMA di Kabupaten Probolinggo.
3. Bagi Masyarakat, dapat menggunakan metode tersebut untuk menyelesaikan kasus *Multiple Traveling Salesman Problem*, seperti penyebaran pestisida, pengintaian musuh pada militer, pendistribusian barang, dan lain-lain.

#### **E. Batasan Masalah Penelitian**

Berdasarkan latar belakang penelitian dan tujuan penelitian, batasan masalah dalam penelitian ini yaitu:

1. MTSP pada skripsi ini menggunakan 1 titik asal dan setiap salesman akan berangkat dan kembali pada simpul kota yang sama.
2. MTSP pada skripsi ini menggunakan *k-means* untuk pengklasteran dan algoritma genetika untuk menentukan rute terdekatnya.

3. Titik tujuan merupakan seluruh SMA di Kabupaten Probolinggo baik negeri maupun swasta.
4. Setiap titik tujuan diasumsikan selalu terhubung dan berjalan lurus.
5. Tidak ada prioritas kota mana saja yang dilalui terlebih dahulu.

## BAB II

### KAJIAN PUSTAKA

#### A. Penelitian Relevan

Ada beberapa hasil penelitian sebelumnya yang memiliki keterkaitan dengan penelitian ini. Penelitian berjudul "*Applying K-means and Genetic Algorithm for Solving MTSP*" [12]. Penelitian tersebut membahas tentang persilangan jalur antar tiap salesman yang dapat dihindari dengan menggunakan algoritma genetika dan *k*-means. Dari penelitian tersebut dihasilkan bahwa dengan penggunaan algoritma genetika dan *k*-means untuk menyelesaikan MTSP dapat meminimalisir terjadinya tabrakan antar salesman.

Penelitian kedua berjudul "*Optimasi Multiple Travelling Salesman Problem (M-TSP) Pada Penentuan Rute Optimal Penjemputan Penumpang Travel Menggunakan Algoritme Genetika*" [17]. Penelitian tersebut membahas tentang permasalahan MTSP yaitu beberapa orang salesman yang akan berangkat dari kantor *travel* menuju ke alamat penjemputan masing-masing penumpang. Pada permasalahan tersebut menggunakan representasi permutasi, proses reproduksi *crossover* dengan *one cut point crossover*, proses mutasi dengan *exchange mutation*, dan proses seleksi dengan *elitism selection*.

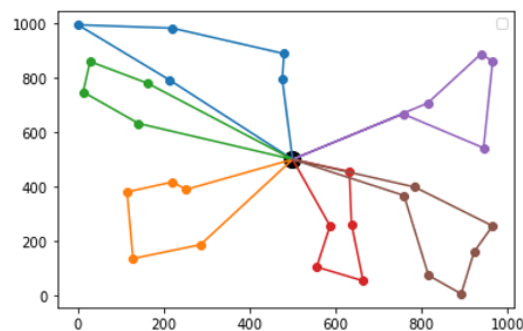
Mayuliana, N. K., Kencana, E. N., dan Harini, L. P. I. dalam artikelnya yang berjudul "Penyelesaian Multitraveling Salesman Problem dengan Algoritma Genetika" [15], mempelajari tentang kinerja algoritma genetika berdasarkan jarak minimum dan waktu pemrosesan yang diperlukan untuk 10 kali pengulangan untuk setiap kombinasi kota penjual. Artikel karangan Al-Khateeb, B., dan Yousif, M. berjudul "*SOLVING MULTIPLE TRAVELING SALESMAN PROBLEM BY MEERKAT SWARM OPTIMIZATION ALGORITHM*" [2] dalam artikel ini mengusulkan algoritma metaheuristik yang disebut algoritma *Meerkat Swarm Optimization* (MSO) untuk

memecahkan MTSP dan menjamin solusi berkualitas baik dalam waktu yang wajar untuk masalah kehidupan nyata.

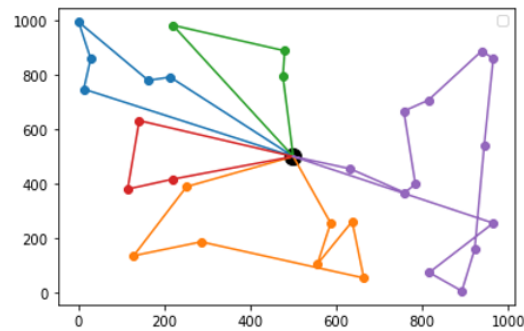
## B. Dasar Teori

### 1. *Multiple Traveling Salesman Problem*

*Travelling Salesman Problem* atau TSP adalah permasalahan pencarian rute paling efisien dalam sebuah perjalanan, sedangkan *Multiple Travelling Salesman Problem* (MTSP) adalah gabungan dari beberapa permasalahan TSP dengan titik kumpul dan titik kembali yang sama. Menurut Al-Omeir dan Ahmed, MTSP adalah salah satu kombinatorial optimasi masalah, yang dapat didefinisikan sebagai berikut: Ada  $m$  jumlah salesman yang harus melakukan perjalanan ke  $n$  sejumlah kota dimulai dengan depot dan berakhir di depot yang sama [3]. Selanjutnya para salesman harus melakukan perjalanan dari satu kota ke kota lain secara terus menerus tanpa mengulang kota mana saja yang telah dilintasi oleh para salesman dan mempertimbangkan jalur terpendek selama perjalanan tersebut. Metode MTSP sebenarnya banyak sekali, namun yang digunakan dalam penelitian ini adalah algoritma genetika dan algoritma  $k$ -means. Dalam hal ini data akan dibagi menjadi beberapa kluster terlebih dahulu sesuai dengan jumlah salesman dari perusahaan, seperti pada Gambar II.1 dan II.2



Gambar II.1: Contoh MTSP 6 kluster



Gambar II.2: Contoh MTSP 5 klaster

## 2. Algoritma

Maulana menyebutkan dalam artikelnya algoritma adalah kumpulan perintah untuk menyelesaikan suatu masalah dan diselesaikan dengan cara sistematis, terstruktur dan logis [14]. Algoritma digunakan untuk memecahkan permasalahan yang dialami oleh seorang pengguna program.

## 3. Algoritma *k*-means

*K*-Means adalah jenis metode klasifikasi tanpa pengawasan yang mempartisi item data menjadi satu atau lebih klaster [1]. *K*-Means mencoba untuk memodelkan suatu dataset ke dalam beberapa klaster sehingga item-item data dalam suatu klaster memiliki karakteristik yang sama dan memiliki karakteristik yang berbeda dengan klaster lainnya.

Menurut S Monalisa [16] tahapan mengklaster menggunakan algoritma *k*-means adalah sebagai berikut.

1. Menentukan banyak klaster
2. Memilih beberapa *centroid* secara acak sesuai banyak klaster



3. Hitung jarak titik ke centroid dengan rumus *Euclidean distance* seperti Persamaan (II.1).

$$d_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (\text{II.1})$$

4. Titik-titik yang tersebar masuk ke klaster yang sama dengan titik *centroid* yang paling dekat
5. Perbarui *centroid* dengan menghitung nilai rata-rata nilai pada masing-masing klaster
6. Lakukan iterasi sebanyak mungkin dengan kembali ke tahapan 3 sampai tidak ada perubahan klaster atau perubahan nilai *centroid*

#### 4. Centroid

Centroid adalah nilai yang dijadikan sebagai titik awal dimulainya sebuah pengelompokan (*clustering*) pada algoritma *k*-means [18]. Untuk melakukan pengelompokan data, dimulai dengan menghitung jarak dari setiap titik tujuan menuju ke setiap titik centroid sebagai awal pembentukan klaster, setelah itu titik-titik tujuan akan dikelompokkan menurut titik klaster terdekatnya. Dalam langkah tersebut akan dihitung titik-titik centroid yang baru menggunakan nilai rata-rata titik dari tiap klaster.

#### 5. Algoritma Genetika

Pada artikel Hermanto disebutkan bahwa algoritma genetika adalah algoritma yang digunakan untuk mencari solusi suatu permasalahan dengan cara yang lebih alami yang terinspirasi dari teori evolusi [7]. Dalam hal ini, algoritma genetika dapat juga digunakan untuk pencarian sebuah rute terpendek dalam sebuah kasus perjalanan.

Menurut Armanda RS [4] dalam artikelnya menyampaikan penyelesaian masalah menggunakan algoritma genetika memerlukan beberapa tahapan sebagai berikut:

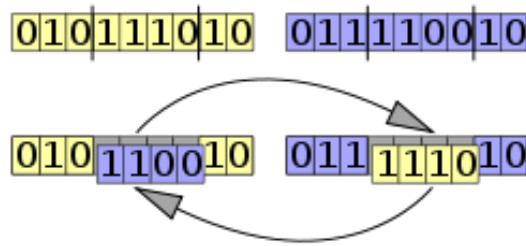
1. Menyiapkan populasi, dalam penelitian ini yang digunakan adalah data yang telah diklaster menggunakan algoritma *k*-means
2. Melakukan reproduksi dengan *crossover* dan mutasi pada pembentukan awal populasi
3. Seleksi dengan metode elitism
4. Menentukan nilai fitness agar mendapatkan solusi akhir yang optimal
5. Iterasi dilakukan untuk generasi berikutnya.

## **6. Fitness**

Fitness adalah suatu ukuran yang dijadikan acuan untuk mengetahui baik atau tidaknya suatu individu atau bisa disebut nilai dari fungsi tujuan [5]. Tujuan dari penggunaan algoritma genetika adalah untuk mengoptimalkan nilai fitness dengan cara mencari nilai fitness yang paling maksimal atau minimal. Seperti dalam penelitian ini yang tujuannya adalah mencari jarak yang paling minimal maka nilai fitness nya yang dicari adalah yang paling minimal juga.

## **7. Crossover**

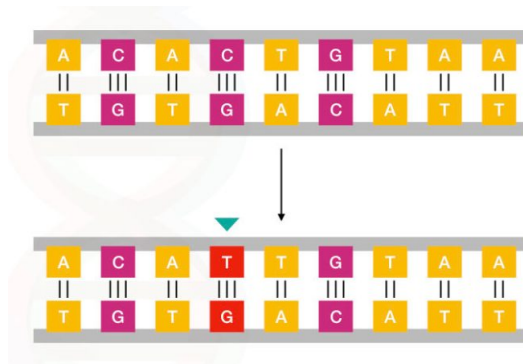
*Crossover* atau persilangan adalah operator dari algoritma genetika yang melibatkan dua induk untuk membentuk kromosom baru menurut artikel [6]. Dalam langkah ini dilakukan dengan cara menukar sebagian gen pada kromosom induk pertama dengan gen pada kromosom induk kedua seperti pada Gambar II.3. Proses *crossover* tersebut diterapkan pada setiap individu dengan probabilitas *crossover* ( $p_c$ ) yang telah ditentukan. Jika diterapkan *crossover* keturunan didapatkan dari kromosom-kromosom induk. Namun jika *crossover* tidak diterapkan satu induk dipilih secara random dengan  $p_c$  yang sama dan diduplikasi menjadi anak.



Gambar II.3: Proses crossover

## 8. Mutasi

Mutasi atau mutation adalah operator yang digunakan untuk mengubah gen-gen yang terdapat dalam kromosom. Model dalam proses ini sebagaimana yang terjadi dalam kehidupan alam [19] seperti pada Gambar II.4. Dalam proses mutasi akan dibangkitkan sebuah bilangan random sebagai Probabilitas mutasi ( $p_m$ ) yang sangat kecil. Mutasi diterapkan dengan tujuan untuk memperoleh nilai fitness yang lebih baik dari sebelumnya, dan lama-kelamaan akan menjadi solusi optimum yang diinginkan.



Gambar II.4: Proses mutasi

## **BAB III**

### **METODE PENELITIAN**

#### **A. Model Penelitian dan Pengembangan**

*Research and Development (R&D)* atau penelitian dan pengembangan adalah suatu metode penelitian yang digunakan untuk menghasilkan produk tertentu, dan menguji keefektifan produk [20]. Berdasarkan pendapat tersebut, metode *Research and Development (R&D)* atau penelitian dan pengembangan dalam bidang pendidikan merupakan penelitian yang bertujuan untuk menghasilkan atau mengembangkan dan memvalidasi suatu produk pendidikan secara efektif. Model penelitian dan pengembangan dalam skripsi ini melalui tahapan sebagai berikut.

1. Tahap pengumpulan data, kegiatan yang dilakukan pada tahap pertama adalah peneliti mengumpulkan data. Pada tahap ini peneliti juga mencari informasi data, yaitu membaca artikel penelitian sebelumnya yang berkaitan dan juga menyiapkan alat bantu atau aplikasi yang akan digunakan untuk membantu pengolahan data. Dari tahap ini data akan dikumpulkan untuk kemudian melanjutkan ke tahapan selanjutnya.
2. Tahap pengolahan data, pada tahap ini penulis mulai mengolah data yang telah dikumpulkan sebelumnya untuk di olah dan dari tahap ini akan dilakukan ujicoba untuk mengetahui keefektifan suatu produk.
3. Tahap analisis, setelah mendapatkan hasil uji coba peneliti mulai menganalisis hasil, menjabarkan, serta mengevaluasinya.
4. Tahap implementasi, pada tahap terakhir ini penelitian yang telah dievaluasi dapat digunakan dan diterapkan pada tempat penelitian.

## **B. Prosedur Penelitian dan Pengembangan**

### **1. Data Penelitian**

Berdasarkan studi kasus dalam skripsi ini, data yang akan digunakan dalam penelitian ini adalah data koordinat dari seluruh SMA di Kabupaten Probolinggo. Data nama-nama sekolah dikumpulkan dari <https://referensi.data.kemdikbud.go.id/> [8], dan data koordinat dikumpulkan melalui aplikasi Google Earth yang dapat diunduh langsung ke dalam bentuk spreadsheet, dapat dilihat pada Lampiran A.. Waktu yang diperlukan peneliti untuk mengumpulkan data dari web tersebut kurang lebih sekitar satu bulan.

### **2. Instrumen Pendukung**

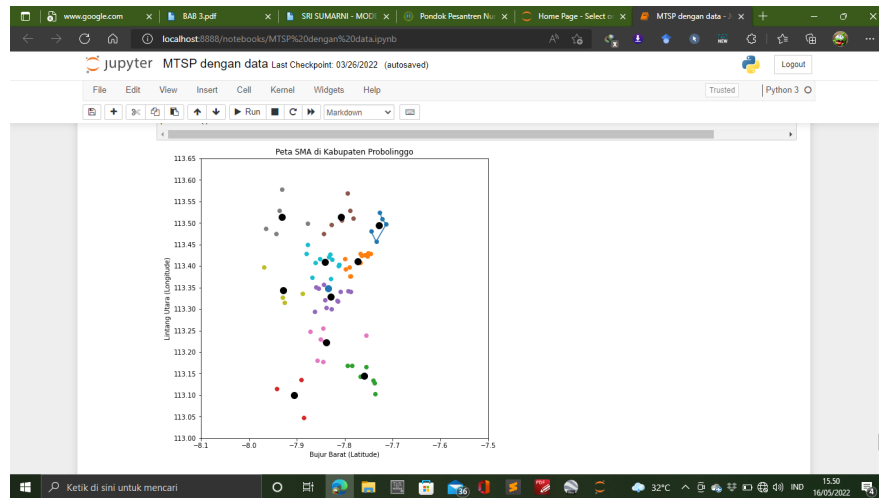
#### **1. Python**

Dalam penelitian ini akan digunakan bahasa pemrograman python untuk mempermudah pengerjaan. Bahasa python adalah bahasa pemrograman baru di masa sekarang, karena dalam bahasa ini lebih simple dan singkat dalam membuat program [21]. Bahasa pemrograman ini merupakan bahasa pemrograman yang paling mudah dipelajari dari pada bahasa pemrograman yang lain. Serta dalam bahasa pemrograman ini dapat menjalankan beberapa rumus matematika di dalamnya. Selain itu bahasa Python telah digunakan secara luas, dan masuk dalam 3 besar bahasa pemrograman yang digunakan dalam beberapa tahun belakangan.

#### **2. Jupyter Notebook**

Jupyter Notebook adalah aplikasi web gratis yang digunakan untuk membuat dan membagikan dokumen komputasi, hasil hitungan, visualisasi, dan teks [9]. Notebook ini juga mendukung 3 bahasa pemrograman salah satunya adalah

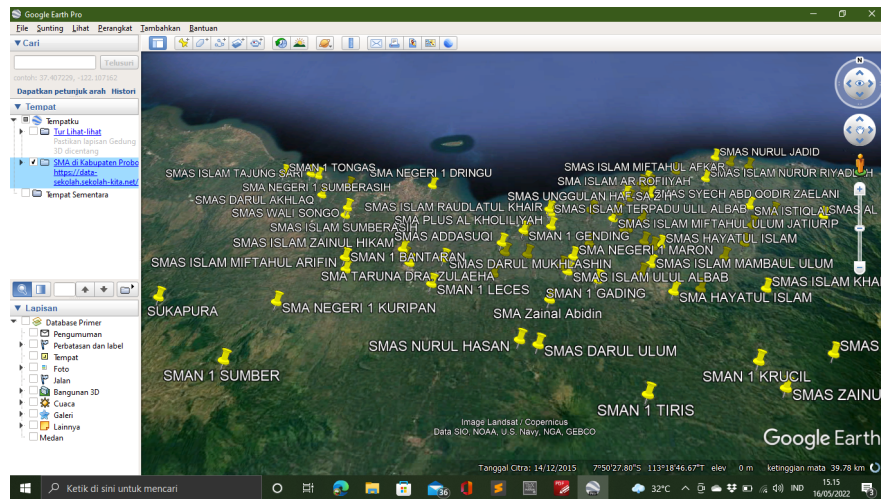
bahasa pemrograman python. Banyak kelebihan yang disajikan dari aplikasi ini salah satunya adalah mendokumentasikan kode, dan menjalankan kode dalam setiap sel, dan visualisasi data seperti pada Gambar III.1.



Gambar III.1: Visualisasi data jupyter notebook

### 3. Google Earth

Google Earth digunakan dalam penelitian ini untuk mengumpulkan koordinat lokasi seluruh SMA di Kabupaten Probolinggo. Aplikasi ini dapat menandai beberapa lokasi secara langsung seperti pada Gambar III.2 dan mengekspor kedalam bentuk spreadsheet seperti pada Gambar III.3. Data-data lokasi yang telah didownload ke dalam bentuk spreadsheet akan diproses menggunakan jupyter notebook.



Gambar III.2: Menandai lokasi pada Google Earth

Nama Sekolah	Latitude (sumbu Y)	Longitude (sumbu X)	Klaster	Urutan
SMAN 1 SUMBER	-7,94	113,11	D	
SMAN 1 SUKAPURA	-7,89	113,05	D	
SMAN 1 BANTARAN	-7,84	113,18	G	
SMAS ISLAM MIFTAHUL ARIFIN	-7,86	113,18	G	
SMA TARUNA DRA ZULAEHA	-7,85	113,23	G	
SMAN 1 LECES	-7,87	113,25	G	
SMAS ISLAM ZAINUL HIKAM	-7,84	113,22	G	
SMAS IT KYAI SEKAR AL AMRI	-7,84	113,22	G	
SMAS DARUL MUKHLASHIN	-7,85	113,26	G	
SMAS ISLAM MIFTAHUL ULUM GUNUNG GENI	-7,86	113,29	E	
SMA NAZHATUT THOLUBIN	-7,84	113,30	E	
SMAS ADDASUQI	-7,83	113,30	E	
SMAS ISLAM RADEN FATAH	-7,84	113,32	E	
SMAN 1 TIRIS	-7,97	113,40	I	
SMAS DARUL ULUM	-7,93	113,33	I	
SMAS NURUL HASAN	-7,93	113,31	I	
SMA HAYATUL ISLAM	-7,88	113,43	J	
SMAN 1 KRUCIL	-7,94	113,48	H	
SMAS ISLAM AR ROHMAH	-7,93	113,58	H	11
SMAS ISLAM NURUL HUDA	-7,94	113,53	H	10
SMAS ZAINUL HASAN 2 KRUCIL	-7,96	113,49	H	
SMAA IBEKADH - KAHART KODIM	-7,93	113,43	I	

Gambar III.3: Mengekspor data ke bentuk spreadsheet

### 3. Langkah-langkah Dalam Tahap Pengolahan Data

1. Menyiapkan data yang telah dikumpulkan sebelumnya.
2. Selanjutnya menentukan jumlah klaster yaitu sebanyak  $n$  klaster. Data yang telah dikumpulkan pada tahap ini akan dibagi menjadi beberapa klaster, metode yang digunakan algoritma  $k$ -means.

3. Langkah-langkah yang digunakan dalam metode *k*-means adalah sebagai berikut

- (a) Tentukan jumlah klaster
- (b) Pilih titik-titik centroid sesuai banyak klaster
- (c) Hitung jarak tiap titik tujuan dengan masing-masing *centroid* dengan rumus *Euclidean distance* seperti pada Persamaan (III.1).

$$d_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (\text{III.1})$$

- (d) Setelah seluruh titik masuk ke dalam klaster, hitung *centroid* baru dengan cara menghitung rata-rata titik pada klaster tersebut. Lakukan hal yang sama pada klaster yang lain.
  - (e) Ulangi langkah 3c sampai tidak ada perubahan pada anggota klaster.
4. Selanjutnya melakukan proses TSP pada setiap klaster yang telah dibagi, langkah-langkahnya adalah sebagai berikut.

- (a) Bangkitkan populasi awal yang berisi sejumlah kromosom
- (b) Hitung nilai *fitness* (jarak) tiap kromosom
- (c) Tetapkan probabilitas *crossover* ( $p_c$ ) dan bangkitkan bilangan random (0,0000 sampai 1,0000) pada tiap kromosom, jika bilangan random kurang dari  $p_c$  maka dilakukan *crossover*. Jika *fitness* kromosom hasil *crossover* lebih baik dari kromosom awal, maka kromosom awal digantikan dengan kromosom hasil *crossover*.
- (d) Tetapkan probabilitas mutasi ( $p_m$ ) dan bangkitkan bilangan random (0,0000 sampai 1,0000) pada setiap kromosom. Jika bilangan random kurang dari  $p_m$  maka akan dilakukan mutasi. Kromosom awal digantikan dengan kromosom hasil mutasi jika kromosom hasil mutasi memiliki *fitness* yang lebih baik dari kromosom awal.



- (e) Jika hasil kurang optimal, iterasi dilakukan dengan cara kembali ke tahapan 4b untuk generasi berikutnya sampai hasil yang dilakukan optimal atau mendekati optimal.
5. Ketika proses diatas selesai dilakukan maka dihasilkanlah pembagian klaster dan rute terdekat tiap klaster menuju seluruh SMA di Kabupaten Probolinggo
  6. Mengevaluasi data yang dihasilkan

## **BAB IV**

### **HASIL**

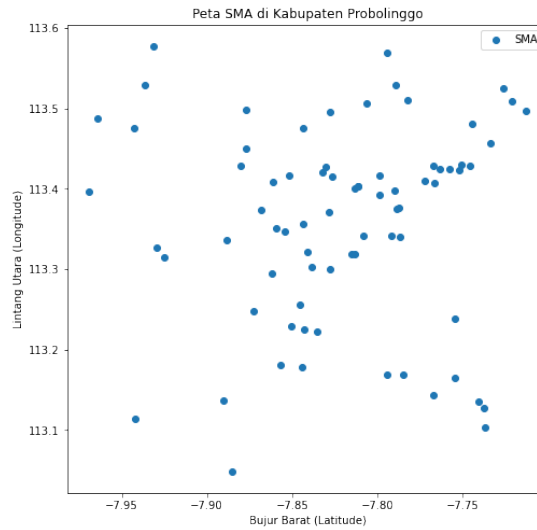
33

#### **A. Penyajian Data Uji Coba**

Pada penelitian ini dilakukan uji coba menggunakan data lokasi seluruh SMA di Kabupaten Probolinggo, dan dijalankan menggunakan python. Berikut adalah sajian data hasil uji coba.

##### **1. Pengambilan Data Lokasi**

Data yang digunakan adalah data koordinat lokasi yang diekspor melalui Google Earth. Pengujian Pengambilan data lokasi bertujuan untuk menunjukkan bahwa sistem mampu membaca input yang dimasukkan. Dapat dilihat pada Lampiran A. seluruh nama-nama SMA di Kabupaten Probolinggo beserta koordinat lokasinya. Visualisasi data dari koordinat-koordinat SMA di kabupaten probolinggo dapat dilihat pada Gambar IV.1.

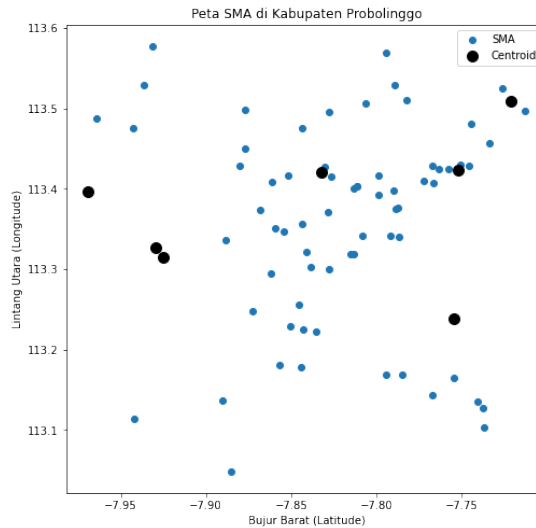


Gambar IV.1: Visualisasi lokasi SMA di Kabupaten Probolinggo

Setelah mendapatkan lokasi yang akan diproses, selanjutnya adalah menentukan beberapa titik centroid secara random, dalam penelitian ini akan diambil 8 centroid secara random seperti pada Tabel IV.1 dan Gambar IV.2.

Nama Centroid	Latitude (Sumbu X)	Longitude (Sumbu Y)
A	-7,81	113,14
B	-7,77	113,51
C	-7,77	113,40
D	-7,88	113,35
E	-7,93	113,51
F	-7,83	113,27
G	-7,84	113,42

Tabel IV.1: Koordinat titik-titik centroid



Gambar IV.2: Visualisasi Centroid

## 2. Proses Pengklasteran Data

Pada tahap ini metode yang digunakan adalah metode  $K$ -means untuk mengklaster data. Langkah-langkah nya adalah sebagai berikut.

1. Menentukan jumlah klaster, dalam hal ini yang digunakan adalah 8 klaster.
2. Memilih titik-titik centroid sebanyak jumlah klaster seperti pada Tabel IV.1 dan IV.2
3. Hitung jarak tiap titik sekolah yang ada dengan masing-masing *centroid*. Penghitungan jarak menggunakan *Euclidean distance* pada persamaan (IV.1).
4. Kelompokkan data ke dalam klaster yang memiliki jarak paling minimum.
5. Setelah seluruh titik sekolah masuk ke dalam klaster-klaster, hitung *centroid* yang baru dengan cara menghitung rata-rata titik sekolah yang ada di dalam klaster tersebut. Lakukan hal yang sama pada klaster yang lain.

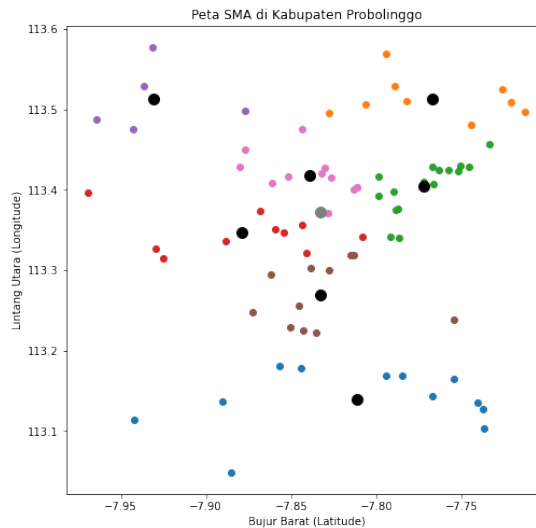
6. Jika terdapat perubahan klaster, maka ulangi langkah 3 hingga tidak ada perubahan anggota pada tiap klaster. Jika *centroid* yang baru tidak berubah dari sebelumnya, maka proses berhenti, karena *centroid* yang tidak berubah menyebabkan anggota klaster juga tidak berubah.

$$[(x,y), (a,b)] = \sqrt{(x-a)^2 + (y-b)^2} \quad (IV.1)$$

7. Setelah semua data terklasifikasi, selanjutnya adalah memperbarui titik-titik centroid dengan cara menghitung rata-rata tiap anggota klaster.

### 3. Proses TSP menggunakan Algoritma Genetika

Setelah data terklaster seperti pada Gambar IV.3 selanjutnya adalah mencari rute terdekatnya menggunakan Algoritma Genetika



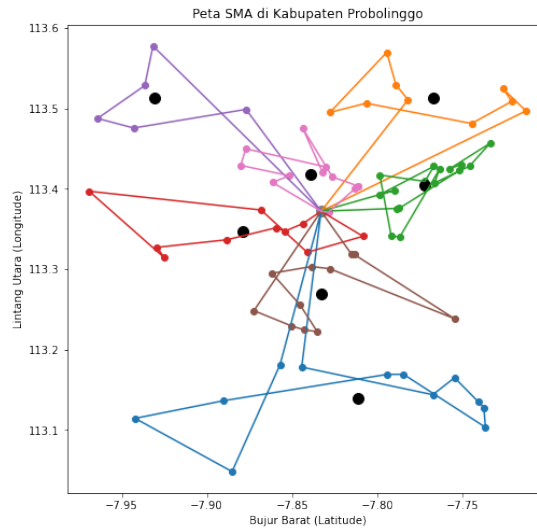
Gambar IV.3: Visualisasi klaster sesuai warna

1. Bangkitkan beberapa populasi awal berisi sejumlah kromosom yang di dalamnya terdapat urutan perjalanan menuju titik sekolah.

2. Hitung nilai *fitness* (total jarak) dari tiap kromosom.
3. Menetapkan probabilitas *crossover* ( $p_c$ ) dalam hal ini yang digunakan adalah  $p_c = 0,95$ . Bangkitkan bilangan random (0,0000 sampai 1,0000) pada setiap kromosom, kromosom dengan bilangan random kurang dari  $p_c$  maka akan dilakukan *crossover*. Jika kromosom hasil *crossover* memiliki *fitness* yang lebih baik dari kromosom awal, maka kromosom awal digantikan oleh kromosom hasil *crossover*.
4. Menetapkan probabilitas mutasi ( $p_m$ ), dalam hal ini digunakan  $p_m = 0,1$ . Bangkitkan bilangan random (0,0000 sampai 1,0000) pada setiap kromosom, kromosom yang memiliki bilangan random kurang dari  $p_m$  maka akan dilakukan mutasi. Jika kromosom hasil mutasi memiliki *fitness* yang lebih baik dari kromosom awal, maka kromosom awal digantikan oleh kromosom hasil mutasi.
5. Jika hasil kurang optimal, iterasi dilakukan dengan cara kembali ke tahapan (2) untuk generasi berikutnya sampai hasil yang dilakukan optimal atau mendekati optimal.

#### **4. Hasil dari proses *k*-means dan Algoritma Genetika**

Dari serangkaian proses di atas, menghasilkan beberapa rute optimal yang dapat dilalui seperti pada Gambar IV.4. Rute-rute yang dihasilkan telah diurutkan sebelumnya dan telah diekspor ke bentuk spreadsheet sehingga mempermudah pengguna dalam membaca data seperti pada Tabel IV.2 merupakan urutan perjalanan pada klaster A.



Gambar IV.4: Visualisasi hasil pembagian kluster dan rute yang dapat dilalui

Nama Sekolah	Latitude (Sumbu X)	Longitude (Sumbu Y)	Kluster	Urutan
SMAS ISLAM MIFTAHUL ARIFIN	-7,86	113,18	A	0
SMAN 1 SUKAPURA	-7,89	113,05	A	1
SMAN 1 SUMBER	-7,94	113,11	A	2
SMA NEGERI 1 KURIPAN	-7,89	113,14	A	3
SMAS ISLAM SUMBERASIH	-7,79	113,17	A	4
SMAS WALI SONGO	-7,78	113,17	A	5
SMAN 1 TONGAS	-7,74	113,10	A	6
SMAS ISLAM TAJUNG SARI	-7,74	113,13	A	7
SMA NEGERI 1 SUMBERASIH	-7,74	113,13	A	8
SMAS ASSUBHAN	-7,75	113,16	A	9
SMAS DARUL AKHLAQ	-7,77	113,14	A	10
SMAN 1 BANTARAN	-7,84	113,18	A	11

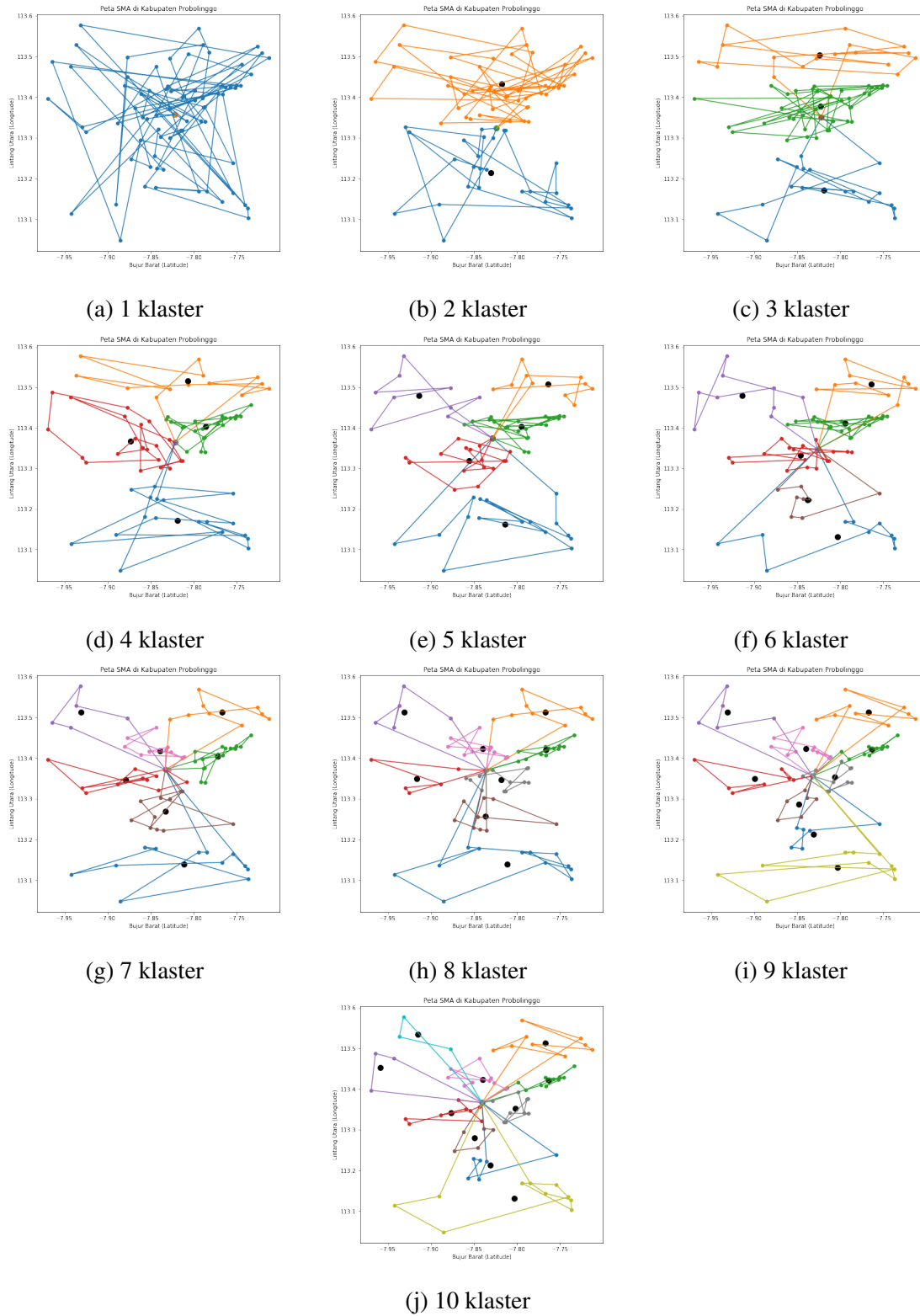
Tabel IV.2: Urutan perjalanan pada kluster A

## 5. Percobaan menggunakan banyak kluster yang berbeda

Pada penelitian ini tidak hanya menggunakan pembagian 8 kluster saja, namun sebelumnya telah diuji coba menggunakan banyak kluster yang berbeda seperti pada

Gambar IV.5. Dalam gambar tersebut telah diuji coba menggunakan pembagian 1 klaster hingga 10 klaster, namun setelah dihitung jarak yang minimal yang dapat dilalui berada pada pembagian 8 klaster seperti Tabel IV.3 membuktikan bahwa pembagian menjadi 8 klaster dapat menghasilkan jarak yang paling minimum.





Gambar IV.5: Hasil rute dengan pembagian kluster berbeda

Banyak Klaster	Total Jarak
1	11,02306125
2	7,454129387
3	6,444054025
4	5,452244797
5	5,152989696
6	4,899631043
7	4,857117163
8	4,763601887
9	5,08663193
10	5,023840617

Tabel IV.3: Total jarak dengan jumlah pembagian klaster yang berbeda

## **BAB V**

### **PENUTUP**

**A. Kesimpulan**

**B. Saran**

## DAFTAR PUSTAKA

- [1] Agusta, Y. (2007). K-means–penerapan, permasalahan dan metode terkait. *Jurnal Sistem dan informatika*, 3(1):47–60.
- [2] Al-Khateeb, B. and Yousif, M. (2019). Solving multiple traveling salesman problem by meerkat swarm optimization algorithm. *Journal of Southwest Jiaotong University*, 54(3).
- [3] Al-Omeir, M. A. and Ahmed, Z. H. (2019). Comparative study of crossover operators for the mtsp. In *2019 International Conference on Computer and Information Sciences (ICCIS)*, pages 1–6. IEEE.
- [4] Armanda, R. S. and Mahmudy, W. F. (2016). Penerapan algoritma genetika untuk penentuan batasan fungsi kenggotaan fuzzy tsukamoto pada kasus peramalan permintaan barang. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 3(3):169–173.
- [5] Basuki, A. (2003). Strategi menggunakan algoritma genetika. *Politeknik Elektronika Negeri Surabaya, Surabaya*.
- [6] Hardi, S. M., Zarlis, M., and Budiarti, E. (2014). Analisis mapping pada partially mapped crossover dalam algoritma genetika pada travelling salesman problem. *TECHSI-Jurnal Teknik Informatika*, 6(1).
- [7] Hermawanto, D. (2003). Algoritma genetika dan contoh aplikasinya. *Retrieved*, 10(25):2013.
- [8] Kemendikbud (2022). Data refrensi pendidikan. <https://referensi.data.kemdikbud.go.id/index11.php?kode=052000&level=2>. Accessed: 2022-04-19.

- [9] Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., and Willing, C. (2016). Jupyter notebooks – a publishing format for reproducible computational workflows. In Loizides, F. and Schmidt, B., editors, *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pages 87 – 90. IOS Press.
- [10] Krishna, K. and Murty, M. N. (1999). Genetic k-means algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(3):433–439.
- [11] Li, L., Zhang, K., Yang, S., and He, J. (2016). Parallel hybrid genetic algorithm for maximum clique problem on opencl. *Journal of Computational and Theoretical Nanoscience*, 13(6):3595–3600.
- [12] Lu, Z., Zhang, K., He, J., and Niu, Y. (2016). Applying k-means clustering and genetic algorithm for solving mtsp. pages 278–284.
- [13] Maii, U. and Bandyopadhyay, S. (2000). Genetic algorithm-based clustering technique. j. *Pattern Recogn*, 33:1455–1465.
- [14] Maulana, G. G. et al. (2017). Pembelajaran dasar algoritma dan pemrograman menggunakan el-goritma berbasis web. *J. Tek. Mesin*, 6(2):8.
- [15] Mayuliana, N. K., Kencana, E. N., and Harini, L. P. I. (2015). Penyelesaian multi traveling salesman problem dengan algoritma genetika. *E-Jurnal Mat*, 6(1):1–6.
- [16] Monalisa, S. (2018). Klasterisasi customer lifetime value dengan model lrfm menggunakan algoritma k-means. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 5(2):247–252.

- [17] Raditya, P. M. R. and Dewi, C. (2017). Optimasi multiple travelling salesman problem (m-tsp) pada penentuan rute optimal penjemputan penumpang travel menggunakan algoritme genetika. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548:964X.
- [18] Retno, S. (2019). Peningkatan akurasi algoritma k-means dengan clustering purity sebagai titik pusat cluster awal (centroid).
- [19] Rovie-Ryan, J., Abdullah, M., Sitam, F., Tan, S., Zainuddin, Z., Basir, M., Abidin, Z., Keliang, C., Denel, A., Joeneh, E., et al. (2014). Genetic diversity of macaca fascicularis (cercopithecidae) from penang, malaysia as inferred from mitochondrial control region segment. *Journal of Indonesian Natural History*, 2(1):14–25.
- [20] Sugiyono, D. (2016). Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan r&d.
- [21] Syahrudin, A. N. and Kurniawan, T. (2018). Input dan output pada bahasa pemrograman python. *Jurnal Dasar Pemrograman Python Stmik, January*, pages 1–7.
- [22] Zhang, K., Yang, S., Qiu, M., et al. (2014). Parallel genetic algorithm with opencl for traveling salesman problem. In *Bio-Inspired Computing-Theories and Applications*, pages 585–590. Springer.

## LAMPIRAN 1

### DATASET

#### A. Nama dan Koordinat SMA di Kabupaten Probolinggo

No.	Nama Sekolah	Latitude (Sumbu X)	Longitude (Sumbu Y)
1	SMA DARUL HIKMAH	-7,76	113,43
2	SMA DARUT TAQWA	-7,79	113,53
3	SMA HAYATUL ISLAM	-7,88	113,43
4	SMA IRSYADUL MUBTADIIN	-7,83	113,42
5	SMA ISLAM AR ROFIIYAH	-7,77	113,41
6	SMA ISLAM MIFTAHUL ULUM GUNUNG GENI	-7,86	113,29
7	SMA ISLAM SYARIF HIDAYATULLAH	-7,81	113,51
8	SMA ISTIQLAL	-7,78	113,51
9	SMA NAZHATUT THOLIBIN	-7,84	113,30
10	SMA NEGERI 1 DRINGU	-7,75	113,24
11	SMA NEGERI 1 KURIPAN	-7,89	113,14
12	SMA NEGERI 1 MARON	-7,84	113,36
13	SMA NEGERI 1 SUMBERASIH	-7,74	113,13
14	SMA NEGERI 2 KRAKSAAN	-7,73	113,46
15	SMA PLUS AL KHOLILIYAH	-7,81	113,34
16	SMA SIROJUL ARIFIN	-7,80	113,42
17	SMA TARUNA DRA. ZULAEHA	-7,85	113,23
18	SMA TERPADU DARUT TAUHID	-7,81	113,40
19	SMA UNGGULAN BADRIDDUJA	-7,75	113,42
20	SMA Zainal Abidin	-7,89	113,34
21	SMAN 1 BANTARAN	-7,84	113,18
22	SMAN 1 BESUK	-7,83	113,50

No.	Nama Sekolah	Latitude (Sumbu X)	Longitude (Sumbu Y)
23	SMAN 1 GADING	-7,87	113,37
24	SMAN 1 GENDING	-7,81	113,32
25	SMAN 1 KRAKSAAN	-7,76	113,42
26	SMAN 1 KRUCIL	-7,94	113,48
27	SMAN 1 LECES	-7,87	113,25
28	SMAN 1 PAITON	-7,72	113,51
29	SMAN 1 SUKAPURA	-7,89	113,05
30	SMAN 1 SUMBER	-7,94	113,11
31	SMAN 1 TIRIS	-7,97	113,40
32	SMAN 1 TONGAS	-7,74	113,10
33	SMAS ADDASUQI	-7,83	113,30
34	SMAS AL HASYIMI	-7,79	113,57
35	SMAS AL KHAIRIYAH	-7,75	113,43
36	SMAS ASSUBHAN	-7,75	113,16
37	SMAS DARUL AKHLAQ	-7,77	113,14
38	SMAS DARUL MUKHLASHIN	-7,85	113,26
39	SMAS DARUL ULUM	-7,93	113,33
40	SMAS HAYATUL ISLAM	-7,83	113,43
41	SMAS IHYAUL IMAN	-7,88	113,45
42	SMAS ISLAM AR ROHMAH	-7,93	113,58
43	SMAS ISLAM IRTIQOIYAH	-7,79	113,40
44	SMAS ISLAM KHAIRIYAH	-7,88	113,50
45	SMAS ISLAM MAMBAUL ULUM	-7,85	113,42
46	SMAS ISLAM MIFTAHUL AFKAR	-7,75	113,43
47	SMAS ISLAM MIFTAHUL ARIFIN	-7,86	113,18
48	SMAS ISLAM MIFTAHUL ULUM JATIURIP	-7,80	113,39
49	SMAS ISLAM MIFTAHUL ULUM OPO OPO	-7,83	113,41



No.	Nama Sekolah	Latitude (Sumbu X)	Longitude (Sumbu Y)
50	SMAS ISLAM NURUL HUDA	-7,94	113,53
51	SMAS ISLAM NURUR RIYADLAH	-7,74	113,48
52	SMAS ISLAM RADEN FATAH	-7,84	113,32
53	SMAS ISLAM RAUDLATUL KHAIR	-7,79	113,34
54	SMAS ISLAM SIROJUL UMMAH	-7,84	113,48
55	SMAS ISLAM SUMBERASIH	-7,79	113,17
56	SMAS ISLAM TAJUNG SARI	-7,74	113,13
57	SMAS ISLAM TERPADU ULIL ALBAB	-7,79	113,34
58	SMAS ISLAM ULUL ALBAB	-7,86	113,35
59	SMAS ISLAM ZAINUL HIKAM	-7,84	113,22
60	SMAS IT KYAI SEKAR AL AMRI	-7,84	113,22
61	SMAS MIFTAHUL HASANAIN	-7,81	113,40
62	SMAS MUHAMMAD SHODIQ	-7,83	113,37
63	SMAS MUHAMMADIYAH 3 PROBOLINGGO	-7,82	113,32
64	SMAS NURUL HASAN	-7,93	113,31
65	SMAS NURUL IMAN	-7,86	113,41
66	SMAS NURUL JADID	-7,71	113,50
67	SMAS SA ADAH NIZHAMUL ISLAM	-7,85	113,35
68	SMAS SYECH ABD QODIR ZAELANI	-7,77	113,43
69	SMAS TAMAN MADYA	-7,77	113,41
70	SMAS TUNAS LUHUR	-7,73	113,52
71	SMAS UNGGULAN HAF-SA Z H	-7,79	113,38
72	SMAS WALI SONGO	-7,78	113,17
73	SMAS ZAINUL HASAN 1	-7,79	113,38
74	SMAS ZAINUL HASAN 2 KRUCIL	-7,96	113,49
75	SMASI NURUL HIDAYAH	-7,81	113,40

Tabel 1.1: Daftar nama-nama SMA di Kabupaten Probolinggo

## **RIWAYAT HIDUP**