**DOKUMEN PROYEK**

**PEMBELAJARAN MESIN**

**” …….”**

**KELOMPOK 4**



**Disusun Oleh :**

|  |  |
| --- | --- |
| Samuel Tobing | 11S20001 |
| Michael Napitupulu | 11S20035 |
| Rolasta Sitorus | 11S20021 |
| Fany Lumbansiantar | 11S20000 |
| Lusye Pasaribu | 11S20000 |
| Aditya Pratama Gultom | 11S20000 |

**FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO**

**SARJANA INFORMATIKA**

**2023**

# Daftar Isi

[1 Daftar Isi 2](#_Toc137397869)

[2 BAB 1. Pendahuluan 3](#_Toc137397870)

[2.1 1.1 Latar Belakang 3](#_Toc137397871)

[2.2 1.2 Pembagian Tugas 4](#_Toc137397872)

[3 BAB 2. Deskripsi dan Cara Kerja Algoritma 5](#_Toc137397873)

[3.1 2.1 Deskripsi Algoritma 5](#_Toc137397874)

[3.2 2.2 Kelebihan dan Kekurangan 5](#_Toc137397875)

[3.3 2.3 Cara Kerja Algoritma 7](#_Toc137397876)

[4 BAB 3. Deskripsi Aplikasi 8](#_Toc137397877)

[4.1 3.1 Deskripsi singkat 8](#_Toc137397878)

[4.2 3.2 Diagram Flow 8](#_Toc137397879)

[BAB 4. Hasil Pengolahan Model 9](#_Toc137397880)

[BAB 5. Kesimpulan 9](#_Toc137397881)

# BAB 1. Pendahuluan

## 1.1 Latar Belakang

Pembelajaran Mesin merupakan metode analisis data yang mengautomasi pembuatan model analitik. Ini adalah cabang dari kecerdasan buatan yang berdasarkan ide bahwa sistem dapat belajar dari data, mengidentifikasi pola dan mengambil keputusan dengan sedikit intervensi dari manusia. Karena teknologi komputasi yang baru, pembelajaran mesin saat ini tidak seperti pembelajaran mesin di masa lampau. Teknologi ini terlahir dari pengenalan pola dan teori bahwa komputer dapat belajar tanpa perlu di program untuk menjalankan tugas-tugas khusus.

Beberapa metode pembelajaran mesin yang populer dan paling banyak di adopsi adalah *Unsupervised Learning* dan *Supervised Learning*, tetapi ada juga metode pembelajaran mesin lainnya. Dimana pada proyek kali ini, kami menggunakan metode Unsupervised learning adalah salah satu pendekatan dalam machine learning di mana model atau algoritma belajar dari data tanpa ada label atau informasi terprediksi sebelumnya. Dalam unsupervised learning, tujuannya adalah untuk menemukan pola, struktur, atau hubungan yang tersembunyi dalam data tanpa adanya petunjuk eksternal . Algoritma yang digunakan didalam proyek ini adalah Mean-Shift Clustering .Mean shift adalah sebuah algoritma yang digunakan dalam analisis citra dan pengenalan pola untuk melakukan segmentasi objek. Tujuan utama dari algoritma ini adalah untuk mengelompokkan titik-titik data dalam ruang fitur berdasarkan kesamaan karakteristiknya.

## 1.2 Pembagian Tugas

Berikut ini adalah pembagian tugas yang dilakukan dalam proses pengerjaan proyek Pembelajaran Mesin :

|  |  |
| --- | --- |
| **Anggota** | **Deskripsi tugas** |
| Samuel Tobing | Pembuatan Model |
| Michael Napitupulu | Pembuatan Model, |
| Rolasta Sitorus | Pembuatan Model, |
| Fany LumbanSiantar | Pembuatan Model, |
| Lusye Pasaribu | Pembuatan Model, |
| Aditya Gultom | Pembuatan Model, |

# BAB 2. Deskripsi dan Cara Kerja Algoritma

## 2.1 Deskripsi Algoritma

Algoritma yang digunakan didalam proyek ini adalah Algoritma Mean-Shift. Algoritma yang termasuk dalam Unserpevised learning yaitu salah satu pendekatan dalam machine learning di mana model atau algoritma belajar dari data tanpa ada label atau informasi terprediksi sebelumnya. Dalam unsupervised learning, tujuannya adalah untuk menemukan pola, struktur, atau hubungan yang tersembunyi dalam data tanpa adanya petunjuk eksternal. Algoritma Mean shift adalah sebuah algoritma yang digunakan dalam analisis citra dan pengenalan pola untuk melakukan segmentasi objek. Tujuan utama dari algoritma ini adalah untuk mengelompokkan titik-titik data dalam ruang fitur berdasarkan kesamaan karakteristiknya.

Prinsip dasar dari mean shift adalah melakukan pergeseran (shift) titik pusat atau centroid dari suatu kelompok titik data berdasarkan konsentrasi titik-titik yang ada di sekitarnya. Pergeseran tersebut dilakukan secara iteratif hingga mencapai kondisi konvergensi, yaitu ketika titik pusat tidak mengalami perubahan yang signifikan.

## 2.2 Kelebihan dan Kekurangan

Berikut ini kelebihan dan kekurangan Algortima Mean Shift.

**Kelebihan Mean Shift :**

1. Kemampuan mengelompokkan data dengan bentuk kelompok yang kompleks: Mean shift dapat mengelompokkan data dengan bentuk yang tidak teratur atau tidak berbentuk bola. Hal ini membuatnya lebih fleksibel dalam mengatasi data yang memiliki pola pengelompokan yang kompleks dan tidak konvensional.
2. Nonparametrik: Mean shift tidak mengharuskan adanya asumsi tertentu tentang distribusi data. Ini berarti algoritma ini dapat diterapkan pada berbagai jenis data tanpa harus memenuhi persyaratan distribusi tertentu.
3. Tidak memerlukan jumlah kelompok yang diketahui sebelumnya: Mean shift dapat menentukan jumlah kelompok secara otomatis berdasarkan data yang ada. Hal ini berguna ketika kita tidak memiliki informasi sebelumnya tentang jumlah kelompok yang ada dalam data.
4. Tidak terpengaruh oleh urutan data: Mean shift tidak tergantung pada urutan data, artinya pengurutan data yang berbeda tidak akan menghasilkan pengelompokan yang berbeda. Ini menjadikannya lebih stabil dan konsisten dalam pengelompokan data.

**Kekurangan Mean Shift:**

1. Sensitif terhadap parameter bandwidth: Algoritma mean shift memerlukan pemilihan parameter bandwidth yang tepat. Jika nilai bandwidth terlalu kecil, dapat menghasilkan kelompok yang terlalu banyak dan terlalu kecil. Sebaliknya, jika nilai bandwidth terlalu besar, kelompok yang dihasilkan bisa menjadi terlalu luas dan tidak memadai. Pemilihan bandwidth yang tepat dapat mempengaruhi kualitas pengelompokan.
2. Komputasi yang intensif: Proses pergeseran titik pusat pada setiap iterasi mean shift memerlukan komputasi yang intensif, terutama jika jumlah data besar. Ini dapat mengakibatkan waktu eksekusi yang lama untuk set data yang besar.
3. Kurang efektif pada data dengan dimensi tinggi: Mean shift dapat menghadapi tantangan dalam mengelompokkan data dengan dimensi tinggi. Ini dikarenakan fenomena "curse of dimensionality" di mana dengan peningkatan dimensi data, ruang fitur menjadi sangat jarang diisi, dan proses pergeseran titik pusat menjadi kurang efektif.
4. Sensitif terhadap inisialisasi awal: Inisialisasi titik pusat awal dalam mean shift dapat mempengaruhi hasil akhir. Jika titik pusat awal ditempatkan dengan cara yang tidak tepat, dapat menghasilkan pengelompokan yang tidak optimal.
5. Dalam penggunaan mean shift, penting untuk mempertimbangkan dan menyesuaikan parameter-parameter yang relevan serta memperhatikan karakteristik data yang akan dielompokkan untuk memperoleh hasil yang baik.

## 2.3 Cara Kerja Algoritma

Berikut adalah langkah-langkah utama dalam cara kerja algoritma Mean Shift:

1. Inisialisasi: Tentukan titik pusat awal untuk setiap kelompok atau titik data yang akan diklasterkan.
2. Penentuan tetangga: Untuk setiap titik pusat, tentukan tetangga-tetangga dari titik data yang berada dalam radius tertentu dari titik pusat. Radius ini didefinisikan sebelumnya sebagai parameter algoritma. Tetangga-tetangga ini akan digunakan untuk menghitung pusat massa (mean) pada langkah berikutnya.
3. Pergeseran titik pusat: Untuk setiap titik pusat, hitung pusat massa (mean) dari tetangga-tetangganya. Pusat massa dihitung dengan cara mengambil rata-rata dari semua tetangga yang termasuk dalam radius. Posisi titik pusat diperbarui dengan menggeser titik pusat awal ke pusat massa yang baru dihitung. Pergeseran ini dilakukan dengan meminimalkan jarak dari titik pusat awal ke pusat massa yang baru dihitung.
4. Konvergensi: Ulangi langkah 2 dan 3 hingga titik pusat tidak bergerak atau bergerak dalam pergeseran yang sangat kecil. Konvergensi tercapai ketika tidak ada perubahan signifikan dalam posisi titik pusat. Hal ini menandakan bahwa algoritma telah mencapai titik stabil dan titik-titik data telah terkelompokkan dengan baik.
5. Penentuan klaster: Setelah konvergensi tercapai, setiap titik data akan dikaitkan dengan titik pusat terdekatnya yang telah dikonvergensi. Titik-titik yang dikaitkan dengan titik pusat yang sama membentuk sebuah klaster.
6. Pada setiap iterasi, algoritma Mean Shift menggerakkan titik pusat ke arah dengan konsentrasi titik data yang lebih tinggi. Jadi, titik pusat secara iteratif menggeser posisinya menuju ke titik pusat yang lebih dominan dalam sekelilingnya. Hal ini memungkinkan algoritma untuk menemukan pusat-pusat klaster dan mengelompokkan titik-titik data berdasarkan karakteristik kesamaan.

Secara umum, algoritma Mean Shift bekerja berdasarkan prinsip pergeseran titik pusat (mean) yang dilakukan berulang kali hingga mencapai konvergensi. Proses ini memungkinkan algoritma untuk mengelompokkan data secara adaptif berdasarkan konsentrasi data dalam ruang fitur.

# BAB 3. Deskripsi Aplikasi

## 3.1 Deskripsi singkat

## 3.2 Diagram Flow

# 

# BAB 4. Hasil Pengolahan Model

# BAB 5. Kesimpulan