

MENGELOMPOKKAN REMAJA BERDASARKAN NILAI IMT (INDEX MASA TUBUH) MENGUNAKAN ALGORITMA K – MEANS

Elsa Pratiwi¹

¹ Teknik Informatika, Institut Teknologi Pagar Alam, Jl.Masik siagim No 75 Simpang Mbacang Indonesia

¹ elsapратиwi2023@gmail.com

Kata kunci:

IMT, K – means, Data mining, Rafid minner, Google colab

Abstract

Menurut Menteri Kesehatan RI tahun 2010, batas usia remaja adalah antara 10 sampai 19 tahun dan belum kawin. Sementara menurut BKKBN batasan usia remaja adalah 10 – 21 tahun (Kemenkes, 2010).[1]. Tujuan penelitian adalah untuk menghasilkan pengelompokan nilai IMT pada posyandu remaja di updt Puskesmas Pengaringan dan Posyandu remaja wilayah updt Puskesmas Pengaringan. Berdasarkan observasi ke tempat penelitian updt Puskesmas rawat inap pengaringan dan posyandu remaja wilayah updt Puskesmas rawat inap Pengaringan. Dimana untuk menentukan IMT masih menghitung dengan excel dan data belum tersusun dengan rapi. Sehingga belum bisa mengukur pengelompokan status masa tubuh pada remaja secara tepat, untuk mengelompokkan IMT (Indeks Masa Tubuh) remaja masih harus mencari dan menghitung ulang data sesuai dengan kebutuhan, hal ini memakan waktu lama dan berisiko tidak akurat. Penelitian ini menggunakan algoritma K- means, Metode klasifikasi menggunakan data mining dan untuk pengujian menggunakan rafid minner dan google colab dengan metode pengembangan menggunakan Crips Dm, penelitian sebagai solusi untuk mempermudah Puskesmas terutama bagian posyandu remaja ketika kegiatan kelapangan, jika ada kasus obesitas kader posyandu remaja bisa memberika solusi supaya tidak menimbulkan penyakit akut dari 99 data remaja hasil pengujian dari rafid miner dan google menunjukkan akurasi 79 %.

Pendahuluan

A. Latar Belakang

Perkembangan zaman yang semakin maju pengaruh teknologi membuat semakin mudah mendapatkan informasi, apalagi informasi kesehatan yang semakin mudah diakses, usia rawan pada tumbuh kembang terjadi pada usia remaja.

Batasan usia remaja menurut WHO adalah 12 sampai 24 tahun (WHO, 2012). Menurut Menteri Kesehatan RI tahun 2010, batas usia remaja adalah antara 10 sampai 19 tahun dan belum kawin. Sementara menurut BKKBN batasan usia remaja adalah 10 – 21 tahun (Kemenkes, 2010).[1]. Pada masa ini upaya kesehatan remaja memiliki tujuan untuk mempersiapkan remaja menjadi dewasa yang sehat, cerdas, berkualitas, produktif dan serta berperan dalam menjaga, mempertahankan, dan meningkatkan kesehatannya. Untuk mengetahui semuanya kemungkinan yang baik maupun yang buruk tentang tumbuh kembang remaja dibutuhkan perhitungan yang baik sehingga bisa melihat kemungkinan apa yang akan mempengaruhi remaja, sehingga perhitungan yang tepat untuk digunakan adalah IMT (Indeks Masa Tubuh).

IMT (Indeks Masa Tubuh) atau BMI (Body Mass Index) yaitu angka penilaian standar berdasarkan proporsi perbandingan tinggi badan dan berat badan, sehingga dapat dikategorisasikan dalam golongan normal, kurang, berlebih, atau obesitas.[2]. Berdasarkan

observasi ke tempat penelitian updt Puskesmas rawat inap pengaringan dan posyandu remaja wilayah updt Puskesmas rawat inap Pengaringan. Dimana untuk menentukan IMT masih menghitung dengan excel dan data belum tersusun dengan rapi. Sehingga belum bisa mengukur pengelompokan status masa tubuh pada remaja secara tepat, untuk mengelompokkan IMT (Indeks Masa Tubuh) remaja masih harus mencari dan menghitung ulang data sesuai dengan kebutuhan, hal ini memakan waktu lama dan berisiko tidak akurat, sehingga dibutuhkan metode pengelompokan seperti K-means. K-Means adalah suatu metode cluster proses dalam pelatihan yang digunakan untuk mengelompokkan data menjadi beberapa cluster berdasarkan kemiripan antar data tersebut.[3].

Berdasarkan penelitian terdahulu menurut [1] dengan judul Aplikasi K means untuk mengelompokkan mahasiswa berdasarkan nilai body mass index (BMI) & ukuran kerangka hasil penelitiannya Berdasarkan hasil analisis Dari hasil penelitian terhadap 20 data sampel diperoleh 3 kelompok mahasiswa berdasarkan nilai BMI dan ukuran rangka, yaitu : BMI normal dan kerangka besar, BMI obesitas sedang dan kerangka sedang, BMI obesitas berat dan kerangka kecil.

Jadi, perhitungan dalam mengelompokkan remaja berdasarkan nilai IMT akan dikembangkan untuk keperluan pemantauan tumbuh kembang remaja dan mempermudah Puskesmas terutama bagian posyandu remaja ketika kegiatan kelapangan, jika ada kasus obesitas kader posyandu remaja bisa memberika solusi supaya tidak menimbulkan penyakit akut, untuk yang kurang berat badan kader posyandu dapat memberika gizi tambahan, karena efek kesehatan pada remaja sangat penting dan juga bisa meminimalisir terjadinya kesalahan dalam perekapan data. batasan masalah penelitian :

1. Algoritma digunakan K means
2. Data yang ada terdapat 1.095 seluruh remaja putra dan putri, terdiri dari 3 kelurahan yakni wilayah kelurahan jokoh, wilyah kelurahan candi jaya, dan wilayah kelurahan padang temu, namun untuk data digunakan yakni data remaja aktif posyandu yang berjumlah 99 remaja baik putra maupun putri tahun 2023
3. Data di analisis menggunakan Rapid miner
4. Pengujian menggunakan Confusion matrix dengan google colab

Berdasarkan latar belakang diatas tujuan adalah untuk menghasilkan pengelompokan nilai IMT pada posyandu remaja di updt Puskesmas Pengaringan dan Posyandu remaja wilayah updt Puskesmas Pengaringan

Berdasarkan latar balakang dan penelitian terdahulu diatas, maka peneliti mengajukan judul mengelompokkan remaja

Metode penelitian

A. OBSERVASI

Metode untuk melakukan observasi langsung ke tempat penelitian dan melakukan pengamatan terhadap permasalahan yang terjadi pada sistem yang sedang berjalan saat ini. [4] berikut instrument dari Observasi yang di dapat.

B. WAWANCARA

Wawancara adalah segi cara atau teknik pengumpulan data, dengan cara interview. [5] berikut instrument wawancara yang di ajukan Pertanyaan kepada pemangku program

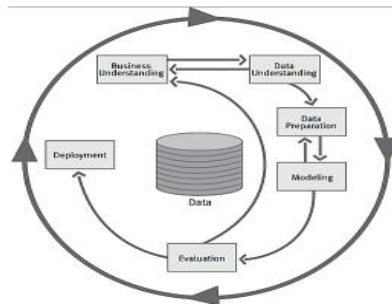
C. STUDI PUSTAKA

Dalam penulisan, peneliti menggunakan berbagai refrensi seperti buku dan jurnal yang berkaitan dengan penelitian, seperti pedoman penulisan pembuatan skripsi ,buku dan jurnal yang berkaitan dengan penelitian ini seperti buku data mining metode regresi linier sederhana sebagai landasan teori untuk menyelesaikan permasalahan yang dihadapi.

METODE PENGEMBANGAN SISTEM CRIPS – DM

Menurut Schröer et al., 2021):1 Cross Industry Standard Process for Data Mining (CRISP-DM) adalah standar model proses yang sering digunakan dalam data mining [6].

CRISP-DM terbagi menjadi 6 tahapan yaitu



gambar 1. alur Crisp-dm

1. BUSINESS UNDERSTANDING

Pada tahap ini bertujuan untuk memahami situasi dan permasalahan yang ada pada objek untuk menentukan tujuan. [6]. Permasalahan dalam objek untuk menentukan tujuan, objek penelitian ini adalah remaja rentang usia 10 – 24 tahun, belum adanya data pengecekan IMT (Indek Masa Tubuh) karena data masih dalam bentuk excel dan belum tersusun, sehingga belum bisa mengukur pengelompokkan status masa tubuh pada remaja, untuk mengelompokkan IMT (Indeks Masa Tubuh) remaja masih harus mencari dan menghitung ulang data sesuai dengan kebutuhan, hal ini memakan waktu lama dan berisiko tidak akurat, Dengan demikian diperlukan adanya pengolahan data yang nantinya dapat otomatis menghasilkan informasi pengelompokkan remaja berdasarkan nilai IMT.

2. DATA UNDERSTANDING

Pada tahap ini bertujuan untuk melakukan pengumpulan dan eksplorasi data untuk melihat kualitas data. Eksplorasi data dilakukan dengan pengecekan missing value, duplicate data dan outliers pada data yang akan digunakan. [6]. Hasil dari tahap ini adalah pengumpulan dan eksplorasi data. Data yang digunakan dalam penelitian sebanyak 99 data, dengan 8 atribut. Berikut data set yang digunakan.

Tabel 1. data set

NO	NAMA	ALAMAT	JENIS KELAMIN	UMUR	BB	TB	IMT	KETERANGAN
1	OCA	RIMBA CANDI	PEREMPUAN	12	40	140	20,4	NORMAL
2	FATIMA	RIMBA CANDI	PEREMPUAN	10	29	140	14,79	KURANG
....
99	DINDA DIVANA	PADANG TEMU	PEREMPUAN	15	50	155	20,83	NORMAL

3. DATA PREPARATION

Menentukan data transform, data training dan data testing [7]

1) DATA SELECTION

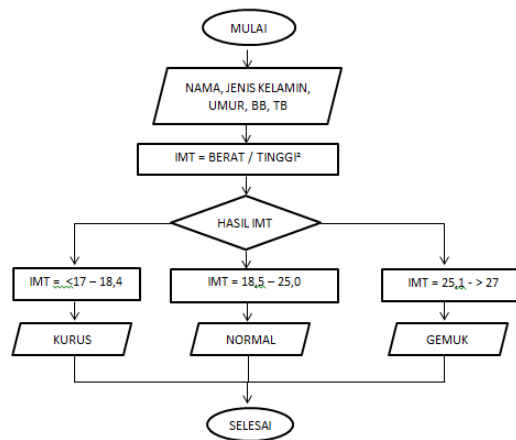
Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. [8]. Dari 99 data 8 atribut data akan di seleksi hingga menjadi 7 atribut saja yang akan digunakan. Yaitu NAMA, JENIS KELAMIN, UMUR, BB, TB, IMT, KETERANGAN

tabel 2. data selection

NO	NAMA	JENIS KELAMIN	UMUR	BB	TB	IMT	KETERANGAN
1	OCA	PEREMPUAN	12	40	140	20,4	NORMAL
2	FATIMA	PEREMPUAN	10	29	140	14,79	KURANG
....
99	DINDA DIVANA	PEREMPUAN	15	50	155	20,83	NORMAL

2) DATA TRANSFORMASI

Tahap selanjutnya yaitu tahap transformasi data. Transformasi data ini adalah mengubahnya Implementasi Rapid minner dari pengelompokkan remaja berdasarkan nilai IMT menggunakan K- means. [8]Flowchart perhitungan IMT pada remaja.



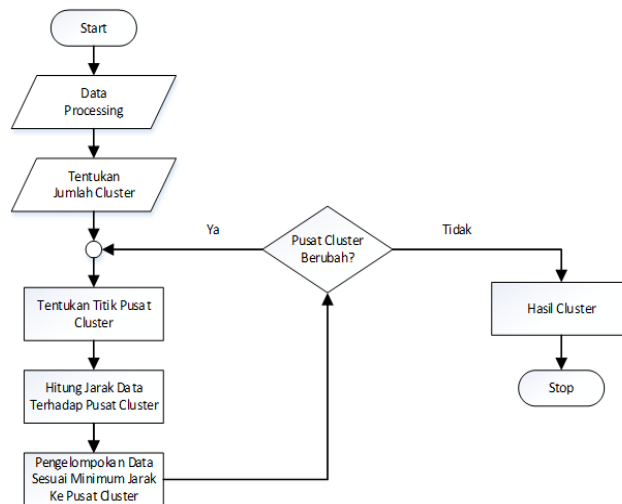
gambar 2. flow chart data tranformasi.

4. DATA CLEANING

Pembersihan data merupakan proses menghilangkan noise dan data yang tidak relevan. Pada umumnya data yang diperoleh, baik dari database memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Data- data yang tidak relevan itu juga lebih baik dibuang. [8]. Data yang di cleaning dari 99 data dengan 8 atribut yaitu , atribut ALAMAT. Sedang JENIS KELAMIN Dan KETERANGAN di tranformasikan kedalam Nominal to Numurical agar terbaca saat di kelompokkan

5. MODELING

Gunakan teknik matematika untuk mengidentifikasi pola dalam data kita. [7]

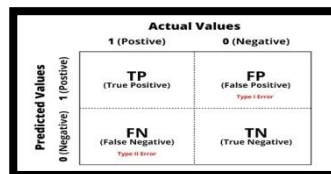


gambar 3. flow chart modeling

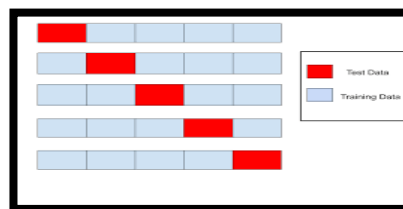
Tahap ini dilakukan pemodelan terhadap data yang ada. Pembuatan model dilakukan menggunakan algoritma K – Means dengan Rapid minner. Sehingga menghasilkan centroid, distance objek centroid, dan grouping based on minimum distance.

6. EVALUASI

Tinjau pola yang kita temukan dan nilai potensinya untuk penggunaan bisnis. [7]. Pada tahap ini, model sudah terbentuk dan diharapkan memiliki kualitas baik jika dilihat dari sudut pandang analisa data. Pada tahap ini akan dilakukan evaluasi terhadap keefektifan dan kualitas model sebelum digunakan dan menentukan apakah model dapat mencapai tujuan yang ditetapkan pada fase awal (Bussines Understanding). Pada pengukuran parameter algoritma K – means Pengujian yang digunakan peneliti yaitu menggunakan Confusion Matrix dan K-Fold Cross Validation yang akan digunakan pada software RapidMiner yang pada dasarnya hanya untuk memberikan informasi perbandingan hasil yang telah dilakukan oleh model yang digunakan dengan hasil klasifikasi sebenarnya dengan melihat akurasi yang tertinggi.



gambar 4 model evaluasi confision matrix



gambar 5 model evaluasi K fold cross validation

7. DEPLOYMENT

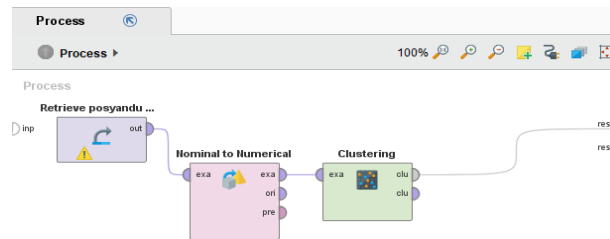
Manfaatkan penemuan kita dalam bisnis sehari-hari. [7]. Adalah tahap pelaporan pemodelan akhir dengan detail penyebarannya. Pada fase ini hasil dari model sudah keluar dan dapat digunakan untuk pengetahuan dan sudah diketahui algoritma mana yang terbaik akurasi atau sama saja untuk penelitian ini

Kesimpulan

Dari pengolahan data 99 data trener di peroleh hasil sebagai berikut:

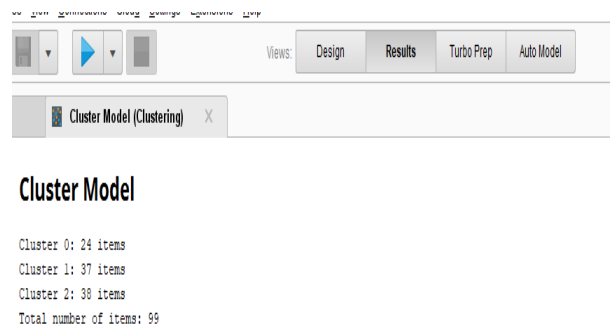
HASIL DARI RAFID MINNER

Pertama adalah menentukan hasil dari pengclusteran untuk cara – cara nya adalah Data yang di masukkan yaitu data posyandu yang telah di label dan menambahkan nominal to numerical, karena pada K – Means, karena Clustering hanya bisa membaca angka, jadi harus di ubah menjadi angka semua katagory sehingga memakai Nomerical to Numerical, baru bisa dikelompokkan makai clustering, berikut gambarnya:



gambar 6. Pengclusteran

Selanjutnya akan timbul hasil dari pengclusterannya seperti pada gambar di bawah ini:



Gambar 7. Hasil pengclusteran

Didapatkan hasilnya yakni: Hasil terdapat 3 cluster

Cluster 0 (kurus) terdapat 24 item / data

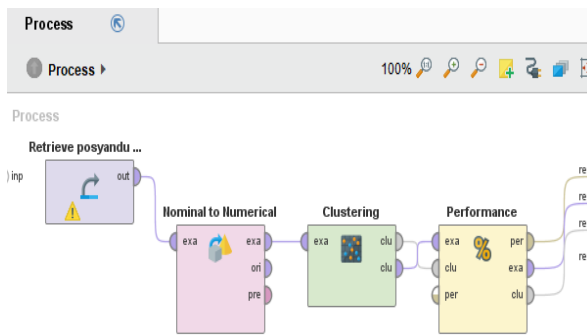
Cluster 1 (normal) terdapat 37 item / data

Cluster 2 (gemuk) terdapat 38 item / data

Total keseluruhan dari 99 item / data yang di kelola.

Yang kedua menentukan performance sehingga menghasilkan nilai centroid pada setiap cluster.

Seperti pada gambar di bawah ini:



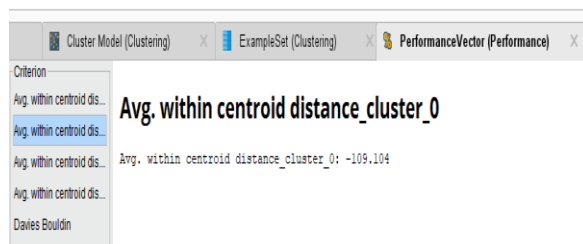
gambar 8. cluster performance

data yang dimasukkan adalah data yang telah di transformasi lalu di performance, setelah di performance. Setelah di performance di dapatkan hasil centroid, adapun hasil centriodnya sebagai berikut: Hasil centroid distance atau jarak dari performance yang telah di jalankan. Seperti pada gambar di bawah ini



gambar 9. centroid distance

Selanjutnya hasil centroid pada performance centroid cluster 0. seperti pada gambar di bawah ini:



gambar 10. centroid cluster 0

Selanjutnya hasil centroid pada performance centroid cluster 1. seperti pada gambar di bawah ini:



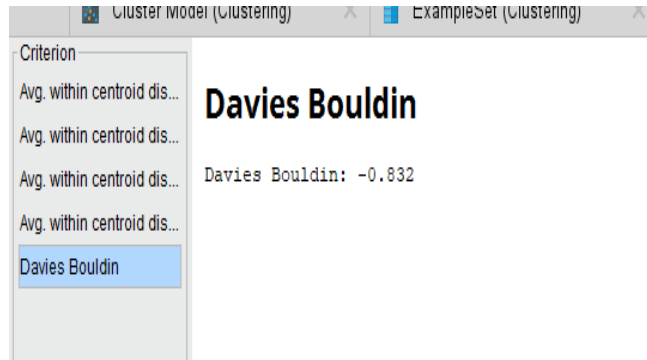
gambar 11. centroid 1

Selanjutnya hasil centroid pada performance centroid cluster 2. seperti pada gambar di bawah ini:



Gambar 12. Centroid cluster 2

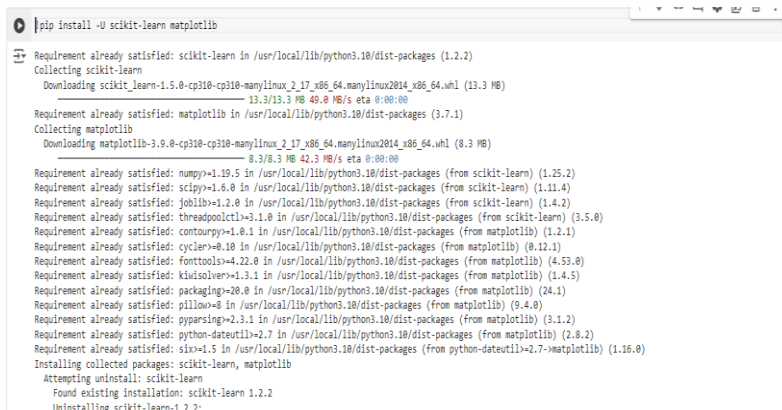
Selanjutnya hasil centroid pada performance davies Bouldin atau titik pusat terdekat . seperti pada gambar di bawah ini:



gambar 13. davies bouldin

HASIL DARI GOOGLE COLAB

Di dapat dengan pengodingan dengan bahasa pemrograman python. Adapun langkah – langkahnya sebagai berikut: Pertama melakukan penginstallan scikit-learn matplotlib, fungsi dari modul ini adalah untuk membantu melakukan processing data ataupun training data untuk kebutuhan data sciences. Cara menginstallnya seperti pada gambar di bawah ini:



gambar 14 install pip

kedua menginstall library Pandas dengan fungsi menganalisa data dan library pandas dengan fungsi memanipulasi data seperti proses pembacaan dokumen dengan format (xlsx,csv,txt dan sebagainya) berikut cara menginstallnya seperti gambar di bawah ini:


```
[ ] import pandas as pd
import numpy as np
```

gambar 15 import pandas dan numpy

Ketiga adalah cara untuk membaca data excel yang menggunakan algoritma K – Means siap diuji, dapat di lihat seperti gambar di bawah ini:

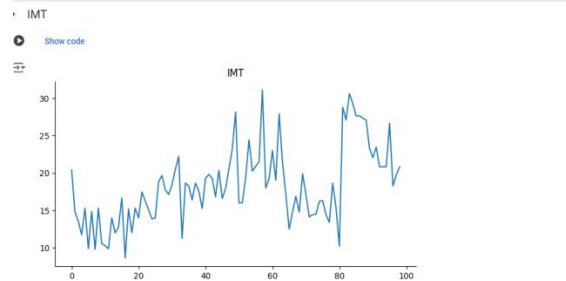
```
data = pd.read_excel('/content/posyandu.xlsx')
data
```

	NAMA	JENIS KELAMIN	URUR	BB	TB	IMT	KETERANGAN
0	Oca	P	12	40	140	20.40	Normal
1	Fatima	P	10	29	140	14.79	Kurang
2	Najua	P	12	27	141	13.66	Kurang
3	Affa	P	10	23	140	11.73	Kurang
4	Olivia	P	10	30	140	15.30	Kurang
...
94	Faisal Sidiq	L	15	50	155	20.83	Normal
95	M. riski	L	15	60	150	26.66	Gemuk
96	A. Noval	L	15	45	157	18.29	Normal
97	Adila	P	15	48	156	19.75	Normal
98	Dinda Divana	p	15	50	155	20.83	Normal

99 rows x 7 columns

gambar 16 data yg di upload

Dari data gambar diatas kita bisa langsung melihat Flot contohnya adalah Flot IMT. Di bawah ini adalah gambar dari Flot IMT



gambar 17 imt

Ke empat adalah cara untuk bisa membaca grafik dengan menggunakan library matplotlib inline yang dengan import matplotlib.pyplot as plt yang berfungsi sebagai pembaca grafik, allu mencari style grafik yang di inginkan contoh: seaborn whitegrid, lalu setelah itu membaca dari peng clusteran nya dengan predix yang ditentukan. Adapun untuk lebih jelasnya bisa di lihat pada gambar di bawah ini:

```
[ ] %matplotlib inline
import matplotlib.pyplot as plt
plt.style.use('seaborn-whitegrid')

~/python-input-1l-ca3kdh2l0cfv:3: MatplotlibDeprecationWarning: The seaborn styles shipped by Matplotlib are deprecated since 3.6, as they no longer correspond
plt.style.use('seaborn-whitegrid')

from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=3)
kmeans.fit(X)
y_kmeans = kmeans.predict(X)

/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:878: FutureWarning: The default value of 'n_init' will change from 10 to 'auto' in 1.4. Set the
warnings.warn()
```

gambar 18 penggambaran grafik

Setelah kode pengclusteran bisa di baca maka selanjutnya yaitu adalah mulai mengcluster remaja bersarkan IMT, untuk rumusnya dapat di lihat seperti di bawah ini:

```

from sklearn.metrics import pairwise_distances_argmin

def find_clusters(X, n_clusters, rseed=2):
    # 1. Randomly choose clusters
    rng = np.random.RandomState(rseed)
    i = rng.permutation(X.shape[0])[:n_clusters]
    centers = X[i]

    while True:
        # 2a. Assign labels based on closest center
        labels = pairwise_distances_argmin(X, centers)

        # 2b. Find new centers from means of points
        new_centers = np.array([X[labels == i].mean(0)
                                for i in range(n_clusters)])

        # 2c. Check for convergence
        if np.all(centers == new_centers):
            break
        centers = new_centers

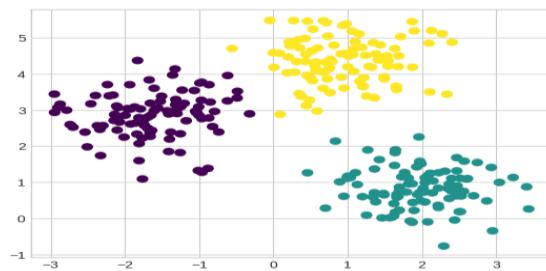
    return centers, labels

centers, labels = find_clusters(X, 3)
plt.scatter(X[:, 0], X[:, 1], c=labels,
            s=50, cmap='viridis');

```

gambar 19 rumus pengclusteran

Jika telah selesai perumusan maka akan di dapat letak cluster dan juga warna Cluster dari jumlah cluster yang ada



Gambar 20 cluster

Kelima adalah cara menentukan labels data remaja berdasarkan nilai, untuk rumusnya yang pertama dibaca dari data set grafik terdapat pada gambar di bawah ini:

```

from sklearn.cluster import KMeans # Import KMeans
from sklearn.datasets import make_moons
import matplotlib.pyplot as plt

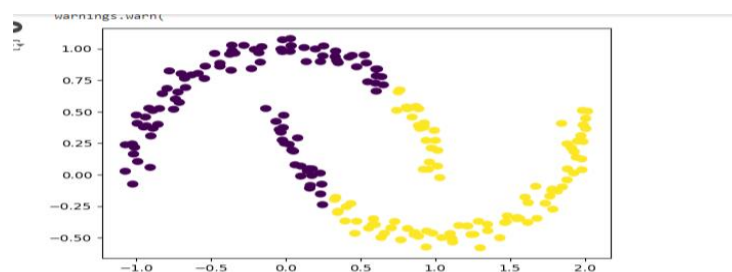
X, y = make_moons(200, noise=.05, random_state=0)

labels = KMeans(2, random_state=0).fit_predict(X)
plt.scatter(X[:, 0], X[:, 1], c=labels,
            s=50, cmap='viridis');

```

gambar 21 rumus make moon

adapun hasil dari pembacaan label dari data set grafik seperti pada gambar di bawah ini:



gambar 22. hasil make moon

cara menentukan labels data remaja berdasarkan nilai, untuk rumusnya yang kedua yaitu dengan membaca Spectral clustering dengan rumus seperti pada gambar di bawah ini

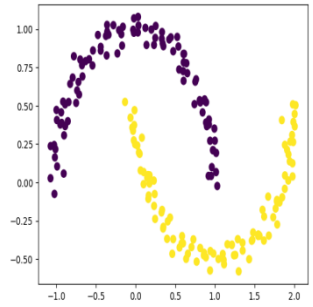
```

from sklearn.cluster import SpectralClustering
model = SpectralClustering(n_clusters=2, affinity='nearest_neighbors',
                           assign_labels='kmeans')
labels = model.fit_predict(X)
plt.scatter(X[:, 0], X[:, 1], c=labels,
            s=50, cmap='viridis');

```

gambar 23. rumus spectral clustering

adapun hasil dari pembacaan label dari data set spectral clustering seperti pada gambar di bawah ini



gambar 24 hasil spectral clustering

Ke enam adalah cara menentukan nilai digit pada remaja berdasarkan nilai IMT dengan rumus dan hasilnya, yang pertama digit data set dan cluster predix seperti pada gambar di bawah ini:

```

from sklearn.datasets import load_digits
digits = load_digits()
digits.data.shape

(1797, 64)

kmeans = KMeans(n_clusters=10, random_state=0)
clusters = kmeans.fit_predict(digits.data)
kmeans.cluster_centers_.shape

/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: FutureWarning: The default
warnings.warn(
(10, 64)

```

gambar 25. pendigitan

cara menentukan nilai digit pada remaja berdasarkan nilai IMT dengan rumus dan hasilnya, yang kedua data shape, cluster center shape seperti pada gambar di bawah ini:

```

from sklearn.datasets import load_digits
digits = load_digits()
digits.data.shape

(1797, 64)

kmeans = KMeans(n_clusters=10, random_state=0)
clusters = kmeans.fit_predict(digits.data)
kmeans.cluster_centers_.shape

/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: FutureWarning: The default
warnings.warn(
(10, 64)

from scipy.stats import mode
import numpy as np # Import the NumPy library and give it the alias 'np'

labels = np.zeros_like(clusters)
for i in range(10):
    mask = (clusters == i)
    labels[mask] = mode(digits.target[mask])[0]

```

gambar 26 pelebelan

HASIL EVALUASI

Terdapat convention matrix dari Algoritma yang digunakan dari pengclusteran remaja berdasarkan nilai IMT. Seperti pada gambar di bawah ini: Adapun rumus convention matrix sebagai berikut:

```
from sklearn.metrics import confusion_matrix
from sklearn.cluster import KMeans
from sklearn.datasets import load_iris
import matplotlib.pyplot as plt
import numpy as np

# Load the Iris dataset and fit a KMeans model
iris = load_iris()
iris_kmeans = KMeans(n_clusters=3).fit(iris.data) # Assuming 3 clusters for Iris

# Assuming 'y' holds the true labels for the Iris dataset
y = iris.target

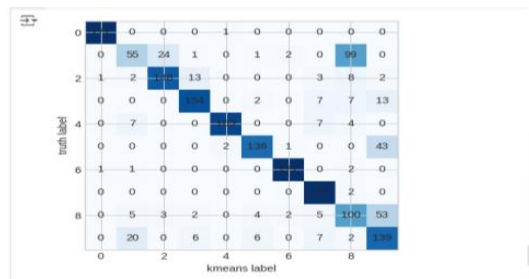
conf_matrix=confusion_matrix(y, iris_kmeans.labels_)

fig, ax = plt.subplots(figsize=(7.5, 7.5))
ax.matshow(conf_matrix, cmap=plt.cm.Blues, alpha=0.3)
for i in range(conf_matrix.shape[0]):
    for j in range(conf_matrix.shape[1]):
        ax.text(x=j, y=i, s=conf_matrix[i, j], va='center',
              ha='center', size='xx-large')

plt.xlabel('Predictions', fontsize=10)
plt.ylabel('Actuals', fontsize=10)
plt.title('Confusion Matrix', fontsize=10)
plt.show()
```

gambar 27 rumus convention matrix

Hasilnya sebagai berikut:



gambar 28 convention matrix

Berikut untuk akurasinya sebagai berikut:

```
from sklearn.metrics import accuracy_score
accuracy_score(digits.target, labels) * 100

79.35447968836951
```

gambar 29. akurasi

Berikut rumus diagram convention matrix sebagai berikut:

```
from sklearn.metrics import confusion_matrix
from sklearn.cluster import KMeans
from sklearn.datasets import load_iris
import matplotlib.pyplot as plt
import numpy as np

# Load the Iris dataset and fit a KMeans model
iris = load_iris()
iris_kmeans = KMeans(n_clusters=3).fit(iris.data) # Assuming 3 clusters for

# Assuming 'y' holds the true labels for the Iris dataset
y = iris.target

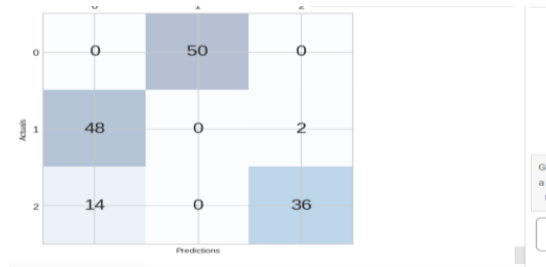
conf_matrix=confusion_matrix(y, iris_kmeans.labels_)

fig, ax = plt.subplots(figsize=(5, 5))
ax.matshow(conf_matrix, cmap=plt.cm.Blues, alpha=0.3)
for i in range(conf_matrix.shape[0]):
    for j in range(conf_matrix.shape[1]):
        ax.text(x=j, y=i, s=conf_matrix[i, j], va='center',
              ha='center', size='xx-large')

plt.xlabel('Predictions', fontsize=8)
plt.ylabel('Actuals', fontsize=8)
plt.title('Confusion Matrix', fontsize=8)
```

gambar 30 rumus tabel convention matrix

Untuk hasilnya sebai berikut;



gambar 31 rumus tabel convention matrix

DEVELOPMENT

Pada tahap ini akan dibuat laporan penelitian dan mempresentasikan dari model di atas yang telah melalui proses data mining. Hal-hal yang akan dipresentasikan mulai dari pengetahuan yang diperoleh selama proses tersebut berjalan sehingga dapat dipahami oleh pengguna. Tahapan pembuatan laporan dapat dilakukan setelah melalui proses *evaluation* dari penelitian terhadap model pengukuran ketepatan dan pola dalam.

Kesimpulan

Dapat disimpulkan bahwa berdasarkan proses keseluruhan yang telah dilakukan untuk mengelompokkan remaja berdasarkan IMT (index masa tubuh) menggunakan algoritma K – means , dari hasil perhitungan data trening sebanyak 99 menggunakan rafid minner dan Google colab tingkat akurasi mencapai 79%.

Ucapan terimakasih

Terimakasih saya ucapkan kepada kepala uptd rawat inap Puskesmas Pengaringan yang telah menerima sebagai tempat penelitian, terimakasih kepada pelaksana lapangan posyandu remaja beserta kader wilayah uptd rawat inap Puskesmas Pengaringan yang telah membantu berkoordinasi mengumpulkan data, terimakasih kepada dosen pengampuh mata kuliah yang telah senantiasa sabar membimbing dab terimakasih kepada teman – teman perjuangan yang terung memberian semangat untuk menyelesaikan tugas ini

Referensi

- [1] T. Rismawan, “APLIKASI K-MEANS UNTUK PENGELOMPOKKAN MAHASISWA,” vol. 2008, no. Snati, 2008.
- [2] M. U. Zamzami Hasibuan and P. A, “Sosialisasi Penerapan Indeks Massa Tubuh (IMT) di Suta Club,” *Cerdas Sifa Pendidik.*, vol. 10, no. 2, pp. 84–89, 2021, doi: 10.22437/csp.v10i2.15585.
- [3] P. M. K-means, M. Tholabah, R. A. Nugroho, and S. Mu, “Nusantara Computer and Design Review Pengelompokkan Status Gizi Anak dan,” pp. 7–14, 2024.
- [4] G. Agung *et al.*, “SEMBAKO BERBASIS ANDROID DENGAN STUDI,” vol. 3, no. 2, 2022.
- [5] Y. H. Nurhayati; Manja Purnasari, “Perancangan arsitektur sistem informasi enterprise pada puskesmas,” *Peranc. Arsit. Sist. ...*, vol.

- 7, no. 1, pp. 84–91, 2023, [Online]. Available:
<https://www.jurnal.kaputama.ac.id/index.php/JTIK/article/view/1245/793>
- [6] H. I. Islam, M. K. Mulyadien, and U. Enri, “Penerapan Algoritma C4. 5 dalam Klasifikasi Status Gizi Balita,” *J. Ilm. Wahana Pendidik.*, vol. 8, no. 10, pp. 116–125, 2022, [Online]. Available: <http://jurnal.peneliti.net/index.php/JIWP/article/view/1839>
- [7] D. M. For, “CRISP-DM : PENGANTAR,” 2014.
- [8] Y. Asriningtias and R. Mardhiyah, “Aplikasi Data Mining Untuk Menampilkan Informasi,” *Informatika*, vol. 8, no. 1, pp. 837–848, 2014.



Elsa pratiwi

Mahasiswa semester 6 yang kuliah
di jurusan Teknik Informatika di kampus
Institut Teknologi Pagar Alam.