

DAFTAR PUSTAKA

- Buslim, Nurhayati, and Rayi Pradono Iswara. 2019. "Pengembangan Algoritma Unsupervised Learning Technique Pada Big Data Analysis Di Media Sosial Sebagai Media Promosi Online Bagi Masyarakat." *Jurnal Teknik Informatika* 12(1):79–96. doi: 10.15408/jti.v12i1.11342.
- Heryati, Agustina, and Muhammad Izman Herdiansyah. 2020. "The Application of Data Mining by Using K-Means Clustering Method in Determining New Students' Admission Promotion Strategy." *International Journal of Engineering and Advanced Technology* 9(3):824–33. doi: 10.35940/ijeat.c5414.029320.
- Iqbal, Muhammad. 2019. "Klasterisasi Data Jamaah Umroh Pada Auliya Tour & Travel Menggunakan Metode K-Means Clustering." *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)* 5(2):97–104. doi: 10.33330/jurteks.v5i2.352.
- Fajar Nugroho, Khusus, Haji. 2019. "PROFILE." (2).
- MURTI, MIKAEL ADITYA WAHYU KRISNA. 2017. "Penerapan Metode K-Means Clustering Untuk Mengelompokkan Potensi Produksi Buah – Buah Di Provinsi Daerah Istimewa Yogyakarta." *Skripsi*.
- Saikin, Saikin, and Kusri Kusri. 2019. "Model Data Mining Untuk Karakteristik Data Traveller Pada Perusahaan Tour and Travel." *Jurnal Manajemen Informatika Dan Sistem Informasi* 2(2):61. doi: 10.36595/misi.v2i2.105.
- Seimahuira, Syarah. 2021. "Implementasi Datamining Dalam Menentukan Destinasi Unggulan Berdasarkan Online Reviews Tripadvisor Menggunakan Algoritma K-Means." *Technologia: Jurnal Ilmiah* 12(1):53. doi: 10.31602/tji.v12i1.4229.
- Syahputra, Siswan, Suci Ramadani, Akim Manoor, and Hara Pardede. 2020. "MENENTUKAN STRATEGI PROMOSI MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS Didapatkan Informasi Anggota Cluster 1 Terdiri Dari 164 Siswa Yang Berasal Dari Kecamatan Kuala Sebanyak 75 Siswa , Dengan Asal Sekolah Terbanyak Dari SMP Negeri 1 Salapian Sebanyak 21 S." 4(1):7–14.

DAFTAR RIWAYAT HIDUP



Muhammad Ichsan adalah nama penulis skripsi ini. Penulis lahir dari orang tua bernama Azmi Saputra dan Eka Sukaesih sebagai anak pertama. Alamat lengkap penulis Desa Sukadamai Sp. Surabaya Lueng Bata Banda Aceh. Penulis dilahirkan dan dibesarkan di Banda Aceh pada 4 September 1998. Penulis menempuh pendidikan dimulai dari TK Bungoeng Jumpa

pada tahun 2003 – 2004, melanjutkan ke MI 3 Banda Aceh, 2010 – 2016 kemudian melanjutkan ke Pondok Pesantren Darussalam Gontor Ponorogo. Pada tahun 2017 penulis melanjutkan pendidikan Strata Satu (S1) dan penulis mengambil program studi Informatika Fakultas Teknik Universitas Nurul Jadid Karanganyar Paiton Probolinggo.

Akhir kata penulis mengucapkan rasa syukur yang sebesar-besarnya atas terselesaikannya skripsi ini. Penulis dapat dihubungi melalui email ikhsansyahputra98@gmail.com

LAMPIRAN

Lampiran 1. Source Code

```
%matplotlib inline

import matplotlib.pyplot as plt

import pandas as pd

import seaborn as sns; sns.set() # for plot styling

import numpy as np

from sklearn.datasets import make_blobs

from sklearn.cluster import KMeans

import seaborn as sns

data = pd.read_csv('DATA_NEW.csv', index_col=0)

data.head(20)

data['ALAMAT'].describe()

data['PAKET'].describe()

data['ALAMAT'].value_counts()

x_ = data['ALAMAT']

y_ = data['PEKERJAAN']

label = data ['PAKET']

plt.scatter(x_,y_, c = label, cmap='plasma')
```

```

from sklearn.cluster import KMeans

from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()
x = scaler.fit_transform(x)
x[:5]

kmeans = KMeans(n_clusters=3, random_state=30)
kmeans.fit(x)

kmeans.labels_

#sum of square error kalau gak ada cluster
sse = []
index = range(1,10)
for i in index:
    kmeans = KMeans(n_clusters=i,random_state=30)
    kmeans.fit(x)
    sse_ = kmeans.inertia_
    sse.append(sse_)
    print(1, sse_)

plt.plot (index, sse)
plt.xlabel('n_cluster')
plt.ylabel('SSE')
plt.show()

```

```
#evaluasi rand score

from sklearn import metrics

rand = []

index = range(1,10)

for i in index:

    kmeans = KMeans(n_clusters=i,random_state=30)

    kmeans.fit(x)

    rand_ = metrics.adjusted_rand_score(y, kmeans.labels_)

    rand.append(rand_)

    print(i, rand_)

plt.plot (index, rand)

plt.xlabel('n_cluster')

plt.ylabel('rand score')

plt.show()

sns.pairplot(data,hue='clusters')

  


#generate dataset

data = make_blobs(n_samples=10,

    centers=4,

    n_features=2,

    cluster_std=1.6,

    random_state=50)

points = data[0]

kmeans = KMeans(n_clusters=4)

kmeans.fit(points)
```

```
plt.scatter(cluster[0][0],cluster[0][1],marker='*',s=500,  
color=('blue'))
```

```
plt.scatter(cluster[1][0],cluster[1][1],marker='*',s=500,  
color=('blue'))
```

```
plt.scatter(cluster[2][0],cluster[2][1],marker='*',s=500,  
color=('blue'))
```

```
plt.scatter(cluster[3][0],cluster[3][1],marker='*',s=500,  
color=('blue'))
```

```
plt.scatter(points[y_km == 0,0], points[y_km == 0,1], s=50,  
color =('cyan'))
```

```
plt.scatter(points[y_km == 1,0], points[y_km == 1,1], s=50,  
color =('red'))
```

```
plt.scatter(points[y_km == 2,0], points[y_km == 2,1], s=50,  
color =('gold'))
```

```
plt.show()
```