

Soal nomor 2, 2502041956 - Nicholas Javier - LA05

Link Video : <https://www.youtube.com/watch?v=yF9EqOwTXt0>

1. [LO1, LO2, LO3] Di tahun 2023 terjadi kasus BSI yang mengalami serangan Ransomware, sehingga mengganggu fungsi Bank ini secara nasional. Pengguna twitter (X) banyak mengunggah tweet memberikan respon terhadap kasus ini. Penasaran dengan apa saja yang dibicarakan user mengenai kasus ini di twitter (X), anda diminta melakukan pemodelan terhadap data kumpulan tweet pengguna dengan detail sebagai berikut:

```
import pandas as pd
import re
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.cluster import KMeans
from sklearn.metrics import silhouette_score
import matplotlib.pyplot as plt
import nltk

df = pd.read_csv('data_2C.csv')
tweets = df['Tweet']
df.head(4)

{"summary":{"\n  \"name\": \"df\",\n  \"rows\": 6100,\n  \"fields\": [\n    {\n      \"column\": \"Unnamed: 0\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 1761,\n        \"min\": 0,\n        \"max\": 6099,\n        \"num_unique_values\": 6100,\n        \"samples\": [\n          2794,\n          2592,\n          5355\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"index\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 7103,\n        \"min\": 4,\n        \"max\": 24400,\n        \"num_unique_values\": 6100,\n        \"samples\": [\n          20818,\n          14133,\n          18997\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"URL\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 6100,\n        \"samples\": [\n          \"https://twitter.com/RezqieHardi/status/1656206329294491650\",\n          \"https://twitter.com/ZulheriIsra/status/1656619602082951168\",\n          \"https://twitter.com/beomracer/status/1656287367047811074\"\n        ],\n        \"semantic_type\": \"\"\n      }\n    }\n  ]\n}}
```

```

\"description\": \"\\n      }\n    },\n    {\n      \"column\":
\"Date\", \n      \"properties\": {\n        \"dtype\": \"object\", \n
\"num_unique_values\": 6061, \n        \"samples\": [\n
\"2023-05-11 08:50:41+00:00\", \n        \"2023-05-09
23:33:46+00:00\", \n        \"2023-05-10 12:52:20+00:00\"\\n
n      ], \n      \"semantic_type\": \"\\n\", \n
\"description\": \"\\n      }\n    },\n    {\n      \"column\":
\"Tweet\", \n      \"properties\": {\n        \"dtype\": \"string\", \n
\"num_unique_values\": 6089, \n        \"samples\": [\n          \"1,5
TB bank BSI Cicho\\n#\\u0e19\\u0e32\\u0e22\\u0e01\\u0e04\\u0e19\\
u0e17\\u0e35\\u0e4830\\n#\\u0e01\\u0e49\\u0e32\\u0e27\\u0e44\\u0e01\\
u0e25 Russia grup Mei lock bit rip hero\\nCharta Politika\", \n
\"Kok @bankbsi_id masih belum bisa transfer ke E-Wallet ya? .. sudah
hari ke 6 ini!\", \n          \"@bankbsi_id @rizkyamikaze
@intanpsupriatna @yazidm05 Saya gak bisa login\\ud83e\\udd26\\ud83e\\
udd26\\ud83e\\udd26\"\\n        ], \n      \"semantic_type\": \"\\n\", \n
\"description\": \"\\n      }\n    },\n    {\n      \"column\":
\"ID\", \n      \"properties\": {\n        \"dtype\": \"number\", \n
\"std\": 802471487627811, \n        \"min\": 1655862147572977665, \n
\"max\": 1658975122894774272, \n        \"num_unique_values\": 6100, \n
\"samples\": [\n          1656206329294491650, \n
1656619602082951168, \n          1656287367047811074\\n        ], \n
\"semantic_type\": \"\\n\", \n      \"description\": \"\\n      }\n    }
n  }, \n    {\n      \"column\": \"Username\", \n      \"properties\":
{\n        \"dtype\": \"string\", \n        \"num_unique_values\":
4544, \n        \"samples\": [\n          \"demaulan_\", \n
\"KaniaLara\", \n          \"BundaZahwa74\"\\n        ], \n
\"semantic_type\": \"\\n\", \n      \"description\": \"\\n      }\n    }
n  }, \n    {\n      \"column\": \"Replies\", \n      \"properties\":
{\n        \"dtype\": \"number\", \n        \"std\": 4, \n
\"min\": 0, \n        \"max\": 179, \n        \"num_unique_values\":
33, \n        \"samples\": [\n          16, \n          19, \n
20\\n        ], \n      \"semantic_type\": \"\\n\", \n
\"description\": \"\\n      }\n    },\n    {\n      \"column\":
\"Retweets\", \n      \"properties\": {\n        \"dtype\":
\"number\", \n        \"std\": 18, \n        \"min\": 0, \n
\"max\": 1177, \n        \"num_unique_values\": 38, \n
\"samples\": [\n          40, \n          20, \n          10\\n
n        ], \n      \"semantic_type\": \"\\n\", \n
\"description\": \"\\n      }\n    },\n    {\n      \"column\":
\"Likes\", \n      \"properties\": {\n        \"dtype\": \"number\", \n
\"std\": 83, \n        \"min\": 0, \n        \"max\": 5189, \n
\"num_unique_values\": 66, \n        \"samples\": [\n          188, \n
29, \n          0\\n        ], \n      \"semantic_type\": \"\\n\", \n
\"description\": \"\\n      }\n    },\n    {\n      \"column\":
\"Quotes\", \n      \"properties\": {\n        \"dtype\": \"number\", \n
\"std\": 2, \n        \"min\": 0, \n        \"max\": 134, \n
\"num_unique_values\": 17, \n        \"samples\": [\n          0, \n
1, \n          3\\n        ], \n      \"semantic_type\": \"\\n\", \n

```

```

{"description": "", "conversationId": 1474192728351150093, "number": 1656482931664367617, "std": 3711218339665346, "min": 1658972921505914881, "max": 1657533017081147393, "num_unique_values": 2455, "samples": [1656465377952681984], "semantic_type": "Text", "description": "Language", "properties": {"category": "Language", "num_unique_values": 12, "samples": [{"nl": "in", "fr": "in"}], "semantic_type": "Text", "description": "Links", "properties": {"category": "Links", "num_unique_values": 203, "samples": [{"https://jagobanking.onelink.me/9ryZ/f8ewsjyl", "https://brief.id/2023/05/beroperasi-akhir-pekan-bsi-terima-setoran-tunai-rp30-m/", "https://twitter.com/Paltiwest/status/1658330792001703937?t=Keg1LRVKzs_6dkrykIRmyg&s=19"}], "semantic_type": "Text", "description": "Media", "properties": {"category": "Media", "num_unique_values": 544, "samples": [{"https://pbs.twimg.com/media/FvwbFeJaMAE2dzN?format=jpg&name=orig", "https://pbs.twimg.com/media/Fv2v-epaQAUlUmI?format=jpg&name=orig", "https://pbs.twimg.com/media/FwUxXA-akAA5Mu6?format=jpg&name=orig"}], "semantic_type": "Text", "description": "Retweeted Tweet", "properties": {"category": "Retweeted Tweet", "num_unique_values": 0, "samples": [], "semantic_type": "Text", "description": "Bookmarks", "properties": {"category": "Bookmarks", "num_unique_values": 4, "samples": [17], "semantic_type": "Text", "description": ""}}], "type": "dataframe", "variable_name": "df"}

```

```

nltk.download('stopwords')
nltk.download('punkt')

```

```

def preprocess_tweet(tweet):
    tweet = re.sub(r"http\S+|www\S+|https\S+", '', tweet,
flags=re.MULTILINE)
    tweet = re.sub(r'\@w+|\#', '', tweet)
    tweet = re.sub(r'^A-Za-z\s', '', tweet)
    tweet = tweet.lower()

```

```
tokens = word_tokenize(tweet)
filtered_tokens = [word for word in tokens if word not in stopwords.words('english')]
return " ".join(filtered_tokens)

df['processed_tweet'] = df['Tweet'].apply(preprocess_tweet)

df.head(3)

[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Package punkt is already up-to-date!

{"summary": "{\n  \"name\": \"df\",\n  \"rows\": 6100,\n  \"fields\": [\n    {\n      \"column\": \"Unnamed: 0\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 1761,\n        \"min\": 0,\n        \"max\": 6099,\n        \"num_unique_values\": 6100,\n        \"samples\": [\n          2794,\n          2592,\n          5355\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"index\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 7103,\n        \"min\": 4,\n        \"max\": 24400,\n        \"num_unique_values\": 6100,\n        \"samples\": [\n          20818,\n          14133,\n          18997\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"URL\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 6100,\n        \"samples\": [\n          \"https://twitter.com/RezqieHardi/status/1656206329294491650\",\n          \"https://twitter.com/ZulheriIsra/status/1656619602082951168\",\n          \"https://twitter.com/beomracer/status/1656287367047811074\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"Date\",\n      \"properties\": {\n        \"dtype\": \"object\",\n        \"num_unique_values\": 6061,\n        \"samples\": [\n          \"2023-05-11 08:50:41+00:00\",\n          \"2023-05-09 23:33:46+00:00\",\n          \"2023-05-10 12:52:20+00:00\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"Tweet\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 6089,\n        \"samples\": [\n          \"TB bank BSI Cicho\\n#\\u0e19\\u0e32\\u0e22\\u0e01\\u0e04\\u0e19\\u0e17\\u0e35\\u0e4830\\n#\\u0e01\\u0e49\\u0e32\\u0e27\\u0e44\\u0e01\\u0e25 Russia grup Mei lock bit rip hero\\nCharta Politika\",\n          \"Kok @bankbsi_id masih belum bisa transfer ke E-Wallet ya? .. sudah hari ke 6 ini!\",\n          \"@bankbsi_id @rizkyamikaze @intanpsupriatna @yazidm05 Saya gak bisa login\\ud83e\\udd26\\ud83e\\udd26\\ud83e\\udd26\\n\", \n          \"\"]\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"ID\",\n      \"properties\": {\n        \"dtype\": \"number\",
```

```
\n      \"std\": 802471487627811,\n      \"min\": 1655862147572977665,\n      \"max\": 1658975122894774272,\n      \"num_unique_values\": 6100,\n      \"samples\": [\n        1656206329294491650,\n        1656619602082951168,\n        1656287367047811074\n      ],\n      \"semantic_type\": \"\",\n      \"description\": \"\",\n      \"column\": \"Username\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 4544,\n        \"samples\": [\n          \"KaniaLara\",\n          \"BundaZahwa74\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\",\n        \"column\": \"Replies\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 4,\n          \"min\": 0,\n          \"max\": 179,\n          \"num_unique_values\": 33,\n          \"samples\": [\n            16,\n            19,\n            20\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\",\n          \"column\": \"Retweets\",\n          \"properties\": {\n            \"dtype\": \"number\",\n            \"std\": 18,\n            \"min\": 0,\n            \"max\": 1177,\n            \"num_unique_values\": 38,\n            \"samples\": [\n              40,\n              20,\n              10\n            ],\n            \"semantic_type\": \"\",\n            \"description\": \"\",\n            \"column\": \"Likes\",\n            \"properties\": {\n              \"dtype\": \"number\",\n              \"std\": 83,\n              \"min\": 0,\n              \"max\": 5189,\n              \"num_unique_values\": 66,\n              \"samples\": [\n                188,\n                0\n              ],\n              \"semantic_type\": \"\",\n              \"description\": \"\",\n              \"column\": \"Quotes\",\n              \"properties\": {\n                \"dtype\": \"number\",\n                \"std\": 2,\n                \"min\": 0,\n                \"max\": 134,\n                \"num_unique_values\": 17,\n                \"samples\": [\n                  0,\n                  1,\n                  3\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\",\n                \"column\": \"conversationId\",\n                \"properties\": {\n                  \"dtype\": \"number\",\n                  \"std\": 3711218339665346,\n                  \"min\": 1474192728351150093,\n                  \"max\": 1658972921505914881,\n                  \"num_unique_values\": 2455,\n                  \"samples\": [\n                    1656482931664367617,\n                    1657533017081147393,\n                    1656465377952681984\n                  ],\n                  \"semantic_type\": \"\",\n                  \"description\": \"\",\n                  \"column\": \"Language\",\n                  \"properties\": {\n                    \"dtype\": \"category\",\n                    \"num_unique_values\": 12,\n                    \"samples\": [\n                      \"nl\",\n                      \"fr\"\n                    ],\n                    \"in\": \"\",\n                    \"semantic_type\": \"\",\n                    \"description\": \"\",\n                    \"column\": \"Links\",\n                    \"properties\": {\n                      \"dtype\": \"category\",\n                      \"num_unique_values\": 203,\n                      \"samples\": [\n                        \"https://jagobanking.onelink.me/9ryZ/f8ewsjyl\",\n                        \"https://brief.id/2023/05/beroperasi-akhir-pekan-bsi-terima-setoran-tunai-rp30-m/\"\n                      ]\n                    }\n                  }\n                }\n              }\n            }\n          }\n        }\n      }\n    }\n  }\n}
```

```

\"https://twitter.com/Paltiwest/status/1658330792001703937?
t=KeglLRVKzs_6dkrykIRmyg&s=19\"\\n        ],\\n
\"semantic_type\\\": \"\\\",\\n        \"description\\\": \"\\\"\\n        }\\
n    },\\n    {\\n        \"column\\\": \"Media\\\",\\n        \"properties\\\": {\\
n        \"dtype\\\": \"category\\\",\\n        \"num_unique_values\\\":
544,\\n        \"samples\\\": [\\n
\"https://pbs.twimg.com/media/FvwbFeJaMAE2dzN?format=jpg&name=orig\\\",\\
n        \"https://pbs.twimg.com/media/Fv2v-epaQAULUmI?
format=jpg&name=orig\\\",\\n
\"https://pbs.twimg.com/media/FwUxXA-akAA5Mu6?format=jpg&name=orig\\\"\\n
],\\n        \"semantic_type\\\": \"\\\",\\n        \"description\\\": \"\\\"\\n
}\\n    },\\n    {\\n        \"column\\\": \"Retweeted Tweet\\\",\\n
\"properties\\\": {\\n        \"dtype\\\": \"number\\\",\\n        \"std\\\":
null,\\n        \"min\\\": null,\\n        \"max\\\": null,\\n
\"num_unique_values\\\": 0,\\n        \"samples\\\": [],\\n
\"semantic_type\\\": \"\\\",\\n        \"description\\\": \"\\\"\\n        }\\
n    },\\n    {\\n        \"column\\\": \"Bookmarks\\\",\\n
\"properties\\\": {\\n        \"dtype\\\": \"number\\\",\\n        \"std\\\":
4,\\n        \"min\\\": 0,\\n        \"max\\\": 231,\\n
\"num_unique_values\\\": 17,\\n        \"samples\\\": [],\\n
\"semantic_type\\\": \"\\\",\\n        \"description\\\": \"\\\"\\n        }\\
n    },\\n    {\\n        \"column\\\": \"processed_tweet\\\",\\n
\"properties\\\": {\\n        \"dtype\\\": \"string\\\",\\n
\"num_unique_values\\\": 6010,\\n        \"samples\\\": [],\\n
\"semantic_type\\\": \"\\\",\\n        \"description\\\": \"\\\"\\n        }\\
n    }\\n ]\\n}\"\", \"type\": \"dataframe\", \"variable_name\": \"df\"}

```

```

vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(df['processed_tweet'])
X.shape

(6100, 9863)

```

a. Lakukan pengelompokan dengan menggunakan pendekatan clustering, silahkan ujicoba beberapa nilai cluster (k) yang mungkin anda peroleh dari data tersebut, berikan perbandingan silhouette score sebagai justifikasi jumlah cluster (k) yang akan anda pilih.

```

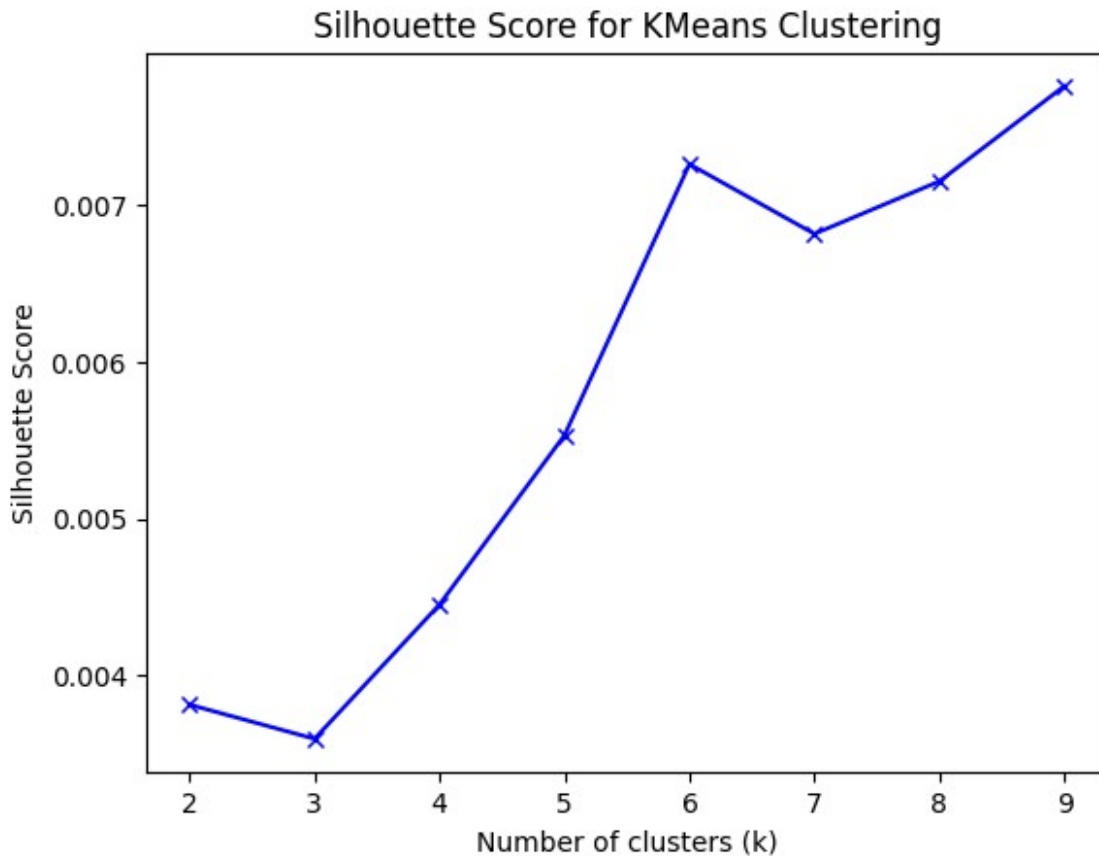
silhouette_scores = []

for k in range(2,10):
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(X)
    score = silhouette_score(X, kmeans.labels_)
    silhouette_scores.append(score)

plt.plot(K, silhouette_scores, 'bx-')
plt.xlabel('Number of clusters (k)')
plt.ylabel('Silhouette Score')

```

[illegible]



Best cluster: 9

Pada model KMeans Clustering, skor Silhouette menunjukkan seberapa mirip titik data dalam kluster yang sama dengan kluster lain, seperti yang ditunjukkan pada grafik berikut. Nilai Silhouette Score berkorelasi positif dengan kualitas clustering yang dihasilkan.

Dari grafik tersebut, Silhouette Score meningkat secara signifikan dari k=4 hingga k=6, dengan skor tertinggi pada k=6. Peningkatan ini menunjukkan bahwa model KMeans menghasilkan kluster yang lebih baik ketika k=6 dibandingkan dengan jumlah kluster lainnya. Namun, peningkatan tidak sejelas dari k=4 hingga k=6, dan k=8 dan k=9 menunjukkan peningkatan yang lebih kecil. Oleh karena itu, k=6 tampaknya menjadi juara untuk model KMeans.

Penggunaan k=6 sebagai jumlah kluster terbaik menunjukkan bahwa data yang dianalisis telah dikelompokkan menjadi enam kluster terpisah, yang memungkinkan pemahaman yang lebih baik tentang data dan segmentasi yang lebih jelas. Penggunaan k=6 akan memberikan pemisahan terbaik, yang membantu dalam interpretasi dan analisis lebih lanjut dari hasil kelompok.

b. Lakukan persona analysis untuk cluster yang anda hasilkan, jelaskan informasi yang anda peroleh mengenai cluster tersebut berdasarkan personanya.

```
kmeans = KMeans(n_clusters=best_k, random_state=42)
df['Cluster'] = kmeans.fit_predict(X)
```



```

for cluster_num in range(best_k):
    cluster_tweets = df[df['Cluster'] == cluster_num]
    ['processed_tweet']
    print(f"\nCluster {cluster_num}:")
    print(cluster_tweets.head(12))

```

Cluster 0:

```

31    ada yang transfer dari hari senin sampai sekar...
42                                     udah hari ke
46    mau transfer ke bank lain dari bsi udh aman bl...
50    halo admin saya pagi ini top sudah klik ok mal...
57    saya terakhir ke bsi customer service nya masi...
60    iyaa kak sama saya juga tranfer ke bri lewat m...
65    nah ini yg menarik di twit awal saya saya tuli...
69    dari hari jumat tanggal mei tf bsi mandiri den...
71    sdh hari saya ngga bisa pakai mobile banking b...
74    di aku msh ke menu saldo tiap ngeklik qris ata...
79    saya pernah rb dan bodohnya masih ga percaya k...
87    kirain internet saya yg bermasalah gara gabisa...
Name: processed_tweet, dtype: object

```

Cluster 1:

```

4                                     bca sy bisa nih kayaknya
5    normal apaan semalem mo tarik tunai masih nga ...
25                                     kok bisa malah suruh pinjeemm
36    nahnk gak bisaa wkwwk sabar ya top aja gk bisa
44    parah banget gak bisa dipake sampe sekarang ce...
64                                     udah bisa narik di atm min
80                                     atm bsi udah bisa kah
81    tolong donggg ini mau dipake transaksi dari ke...
82    admin permisi mau tanya info atm bsi area jogj...
94                                     tarik tunai bisa tidak
96    kapok banget cuma punya mbanking nya doang ga ...
97                                     pakai atm bca konvensional bisa pak
Name: processed_tweet, dtype: object

```

Cluster 2:

```

0    sistem sama orangnya harus di ganti gak ada si...
1                                     gak gitu
2    iya nasabah bank konven sprti bribnimandiri di...
6    wes bener bener dadi kampus obama malah dadi b...
7                                     bank jelek
8    mau sampe kapan klen error tiap jam ku pantau ...
9                                     lol ini respon paling bikin ngakak
10   iya ini kobisa ilo kena hack gabisa njajan di ...
11                                     hari bang kocak emng
12   bisaa tapi transfernnya antar bank bsi kakak
13   hai sobatrupiah sebelumnya kami turut bersimpa...

```

```
16      astaghfirulah begitu rupanya mbah es ya mbah u...
Name: processed_tweet, dtype: object
```

Cluster 3:

```
27      yg pulih mbanking nya kapan kak
37      masii belum bisa kak aku udah coba dari tadi
43      aku brusan transfer kak dh bisa lewat mbanking...
47      ituu dicoba terus aja kak aku awalnya juga gak...
98      kak transfer ke bank lain pake bsi udah bisa k...
122     belum semuanya bisa kak itu bertahap pemulihan...
127     hai kak info apakah sudah bisa atm bsi nya
128     coba ke cabang terdekat aja kak udah mulai pul...
155     iya kak tiap setengah jam sekali udah cek soal...
201     sudah berkala kak setiap jam blm bisa jg
233     coba dicek berkala aja kak kalau buat tfsama c...
250     sama kak aku jg kayak gitu nanti kalau saldony...
Name: processed_tweet, dtype: object
```

Cluster 4:

```
188      bisa dm aja kaa
276      admin cek dm pleaseee
284      dm minta tolong bantingin
624      cek dm min
726      min tolong baca dm saya
800      hallo tolong cek dm saya
844      cek dm min
936      cek dm min
955      miiin cek dm miiinn
976      halo admin tolong balas dm saya ya
1001     cek dm
1210     cek dm
Name: processed_tweet, dtype: object
```

Cluster 5:

```
17      masih belum bisa pake mbanking ya min
28      min belum bisa beli pulsa top ewallet dll dari...
61      ternyata hingga hari kelima ini bsi mobile mas...
75      hr ini sy belum bisa transfer di app atau di a...
93      masih
118     serius hari belum beneerrr
137     masih belum bisa transfer antar bank
148     ini apk mobile nya kok belum bisa juga ya di o...
157     kok sy belum bisa ya
169     apa apaan ini bsi sampai hari belum bisa digun...
200     sudah hari ke n bsi mobile msh belum bisa digu...
206     emas belum bisa diakses tolong segera ditindak...
Name: processed_tweet, dtype: object
```

Cluster 6:

```
3      gimana sih min ini kok gak bisa lagi bsi mobil...
```

```

14         biasa nya ga selama ini loh ini udh smpe hari
19     min bsi mobile masih gangguan sampe hari ini b...
24     ini gimnaa ya gimannaaa masih eror kah mbankin...
34     kapan bisa transaksi di marketplace semua masi...
49     adakah yang bisa memberikan penjelasan dari te...
56     min tolong benerin geh gaji misua di bsi lah k...
67         gangguan mbanking ini kapan selesainya
68         lagi maintenance udah hari
86     gw pindahkan juga uangnya ini error dari minggu...
90     assalamualaikum min mohon info mbanking lagi k...
101         udah aman hari ini

```

Name: processed_tweet, dtype: object

Cluster 7:

```

15     udah atm abis ilang mbanking gabisa udah hari ...
63     pantesan dari kemaren gue mau transfer gabisa ...
138         ko barusan gabisa lagi
139     min gimana nih masih error gabisa tf dan tarik...
140     sampe kapan mbanking gabisa dipake kasian anak...
156         lama bgt sich kack errornya ak gabisa tf nich
175     min ini kenapa masih gabisa top shopeepay ya
196     iya anjrit aku juga mo top udah beberapa hari ...
205     gabisa kak boro cek saldo blabla ini br klik l...
217     mbanking udah bisa blm tolong dong di atm jg g...
254     yang bener ajaa sih ini saya mau buka mbanking...
302     ini mo sampe kapan ya mbanking nya gabisa top ...

```

Name: processed_tweet, dtype: object

Cluster 8:

```

39     kalau topup ewallet msh pemulihan setau ku ka ...
73         sudah mulai kembali pulih ka atm dan kc
131     iya ka ini yang saya tahu proses pemulihan cob...
218     iya ka msh pemulihan terus tapi coba berkala a...
268     sabar ya ka aku juga nunggu mobilenya ini semo...
280     sekarang udh bisa belum daerah sana ka atm di ...
303     topup ewallet ya ka atau topup apa emoney bisa...
313         saran ku kalau urgent bgt bisa ke atm ka
324     aku bisa ka nyoba kemarin soalnya kalau mobile...
326     mabanking ku udh bisa ini ka coba aja di cek ka
348     ka saranku kalau urgent bisa cek di atm ka ata...
409     saran ku tf jgn via bi fast transfer online bi...

```

Name: processed_tweet, dtype: object

```

kmeans = KMeans(n_clusters=best_k, random_state=42)
df['cluster'] = kmeans.fit_predict(X)

```

```

from collections import Counter

```

```

def get_top_keywords(data, clusters, vectorizer, n_terms=10):
    df = pd.DataFrame(X.todense()).groupby(clusters).mean()

```

```

terms = vectorizer.get_feature_names_out()
top_keywords = {}
for i, row in df.iterrows():
    top_keywords[i] = [terms[t] for t in row.argsort()[-n_terms:]]
return top_keywords

top_keywords = get_top_keywords(df['processed_tweet'], df['cluster'],
vectorizer)

for cluster, keywords in top_keywords.items():
    print(f'Cluster {cluster}: {keywords}')

/usr/local/lib/python3.10/dist-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
    warnings.warn(

Cluster 0: ['bisa', 'saldo', 'transfer', 'masuk', 'di', 'lain',
'bank', 'bsi', 'saya', 'ke']
Cluster 1: ['cek', 'di', 'min', 'masih', 'atm', 'udah', 'gak', 'blm',
'ga', 'bisa']
Cluster 2: ['ya', 'dan', 'yang', 'syariah', 'yg', 'ada', 'ini', 'di',
'bank', 'bsi']
Cluster 3: ['kalo', 'hati', 'ke', 'belum', 'aku', 'udah', 'coba',
'aja', 'bisa', 'kak']
Cluster 4: ['baca', 'bales', 'ya', 'admin', 'dong', 'saya', 'tolong',
'min', 'cek', 'dm']
Cluster 5: ['sampai', 'bsi', 'ya', 'mobile', 'transfer', 'ini', 'min',
'masih', 'bisa', 'belum']
Cluster 6: ['masih', 'sampai', 'bsi', 'sampe', 'lagi', 'hari',
'error', 'min', 'ini', 'kapan']
Cluster 7: ['ke', 'atm', 'bsi', 'mbanking', 'hari', 'ya', 'ini',
'mau', 'min', 'gabisa']
Cluster 8: ['udh', 'kalau', 'urgent', 'terus', 'cek', 'atm',
'berkala', 'bisa', 'coba', 'ka']

```

Analisis Kluster Berdasarkan Personanya (Ringkas):

Cluster 0:

- **Persona:** Pengguna mengalami masalah transfer antar bank.
- **Informasi:** Kesulitan transfer dari BSI ke bank lain, frustrasi karena masalah belum terselesaikan.

Cluster 1:

- **Persona:** Pengguna mengalami masalah dengan ATM dan transaksi tunai.
- **Informasi:** Kesulitan menggunakan ATM BSI, bertanya kapan ATM berfungsi normal.

Cluster 2:

- **Persona:** Pengguna marah dan frustrasi dengan masalah sistem BSI.
- **Informasi:** Keluhan sistem sering error, ketidakpuasan dengan layanan BSI, keinginan pindah bank.

Cluster 3:

- **Persona:** Pengguna mencari informasi pemulihan mobile banking.
- **Informasi:** Bertanya kapan layanan pulih, disarankan cek berkala.

Cluster 4:

- **Persona:** Pengguna meminta bantuan melalui DM.
- **Informasi:** Meminta admin cek DM untuk solusi personal.

Cluster 5:

- **Persona:** Pengguna melaporkan masalah mobile banking berlanjut.
- **Informasi:** Belum bisa transaksi, transfer, top-up, frustrasi dengan durasi masalah.

Cluster 6:

- **Persona:** Pengguna ingin penjelasan dan solusi cepat.
- **Informasi:** Menanyakan penyelesaian masalah, berharap perbaikan segera.

Cluster 7:

- **Persona:** Pengguna mengalami masalah umum dengan mobile banking.
- **Informasi:** Tidak bisa transaksi, transfer, cek saldo, keluhan masalah lama.

Cluster 8:

- **Persona:** Pengguna memberikan saran dan info pemulihan layanan.
- **Informasi:** Berbagi info pemulihan, saran alternatif menggunakan ATM, cek berkala.

C. Anda tertarik untuk membandingkan hasil clustering dengan metode lain yaitu topic extraction, sehingga anda perlu melakukan pemodelan topic extraction. Jelaskan metode topic extraction yang anda lakukan

```
from sklearn.decomposition import LatentDirichletAllocation

lda = LatentDirichletAllocation(n_components=best_k, random_state=42)
lda.fit(X)

LatentDirichletAllocation(n_components=9, random_state=42)
```

Pustaka sklearn digunakan untuk menerapkan analisis topik dalam kode yang diberikan. Model Latent Dirichlet Allocation (LDA) adalah metode pemodelan topik yang menganggap setiap dokumen sebagai campuran dari beberapa topik, dengan `n_components=best_k` menunjukkan jumlah topik yang diperkirakan ideal (`best_k`) yang diinginkan untuk ditentukan dalam analisis. Selanjutnya, model dilatih dengan data masukan dari `X`. Untuk memastikan bahwa hasilnya dapat diulang dalam kondisi yang sama, kondisi random disetel ke 42. Tujuan dari proses ini

adalah untuk menemukan struktur topik yang tersembunyi dalam koleksi dokumen yang diberikan kepada Anda.

```
df['Topic'] = lda.transform(X).argmax(axis=1)

for topic_num in range(8):
    topic_tweets = df[df['Topic'] == topic_num]['processed_tweet']
    print(f"\nTopic {topic_num}:")
    print(topic_tweets.head(12))

    topic_words = lda.components_[topic_num]
    top_words_indices = topic_words.argsort()[-10:]
    top_words = [vectorizer.get_feature_names_out()[i] for i in
top_words_indices]
    print(f"Top words: {top_words}")
```

Topic 0:

```
6      wes bener bener dadi kampus obama malah dadi b...
20      ahjangan mengadingadi sampean minga bisadari t...
54      wani ne akun palsu koment di akun real
134      wanjer
161      gua gak bisa transfer weiiii kapan betulnya si...
184      ada pernyataan resmi dari mereka
204      klo hanya liat ui bsi oke banget sih dibanding...
216      apakah mungkin dibalik semua ini ada persaing...
257      hallloo dari bunga mayang lampung utara berapa...
313      saran ku kalau urgent bgt bisa ke atm ka
330      bukan masalah pindah bank atau ga nya ingat na...
348      ka saranku kalau urgent bisa cek di atm ka ata...
```

Name: processed_tweet, dtype: object

Top words: ['kalau', 'hubungi', 'ke', 'coba', 'urgent', 'center', 'terdekat', 'call', 'kc', 'ka']

Topic 1:

```
8      mau sampe kapan klen error tiap jam ku pantau ...
13      hai sobatrupiah sebelumnya kami turut bersimpa...
33      memalukan banget
38      hai candra sebelumnya kami turut bersimpatia at...
62      duitmu ilang ra dab
84      serangan siber yang membuat pelayanan pt bank ...
85      iyaa masak muncul permintaan tidak diproses mulu
119      andaipun data yg di ambil tidak penting tetapi...
121      sejak akhir april dari gambar terlihat saldo j...
130      saham bank besar kompak turun usai data bsi bo...
133      indonesia tengah dibuat geger dengan serangan ...
149      ojk harus tegur krn gangguan sistem miliknya s...
```

Name: processed_tweet, dtype: object

Top words: ['serangan', 'lockbit', 'dan', 'yang', 'nasabah', 'data', 'syariah', 'bsi', 'bank', 'indonesia']

Topic 2:

```
25          kok bisa malah suruh pinjeemm
35  narik pulsanya banyak cuma buat dengerin lagu ...
182 bsi bergerak cepat nih memperbaiki gangguan ma...
207          harus diantisipasi
272          iyes bcs se urgent itu mau transkasi
275  ketua forum keamanan cyber sendiri bilang kalo...
325          menarik untuk disimak
410          kok orang bnyk uitnya yaa
439  parah banget ya ternyata bank bsi ga bisa dipe...
483          hari ketiga ngandat
485          oalah
523          hilang nya nopo
```

Name: processed_tweet, dtype: object

Top words: ['perbaiki', 'yang', 'sm', 'bsi', 'yg', 'ada', 'bank', 'alhamdulillah', 'parah', 'penipu']

Topic 3:

```
40          kasus paralon apa sih
72          pan nanyanya yg net banking min
112         windowsnya ori apa ga aja gatau
147  yakin masukin bri saham bris dipegang oleh bri
152  harusnya dirutnya bikin konpers minta maaf dan...
156  lama bgt sich kack errornya ak gabisa tf nich
171  buktinya kuat di rekaman cctv terlihat jelas s...
176  bank aceh gt bsi lagian anggota dpra naon bang...
187  bank muammalat seperti mendapat durian runtuh ...
240          ingat dosa
251  tottebag muat banyak bank bsi bca syariah khai...
396          duo infamous bank bsi sama bsi
```

Name: processed_tweet, dtype: object

Top words: ['lu', 'ziva', 'selasa', 'lockbit', 'nabila', 'milikmu', 'bsi', 'ipk', 'khairunnisa', 'bank']

Topic 4:

```
1          gak gitu
2  iya nasabah bank konven sprti bribnimandiri di...
9          lol ini respon paling bikin ngakak
16  astaghfirullah begitu rupanya mbah es ya mbah u...
29          coba tolong di perhatikan lg
48  bayangin se pusing apa manajemen bank bsi seka...
51  data bocor ga hanya instansi ini beberapa inst...
55  bank bsi malah kena serangan hoax begini ruqya...
65  nah ini yg menarik di twit awal saya saya tuli...
76  masih banyak anak bangsa yg mampu dan pastinya...
77  dept kalian gk profesional lebih baik di cut s...
89  bahaya data nasabah juga bocor ini ada beberap...
```

Name: processed_tweet, dtype: object

Top words: ['syariah', 'ya', 'ada', 'nya', 'itu', 'ini', 'yg', 'di',


```
'bsi', 'bank']
```

Topic 5:

```
22                                     anjing lu
26    iya setuju maksudnya juga jangan taruh telur d...
57    saya terakhir ke bsi customer service nya masi...
104                                     okeh aku coba thank
165    iya kalo arsitektur data backupnya bagus reboo...
247    enih mahasiswa duitnya smua di bank mau beli s...
289    sampai kpn bosss iuran udah pada nunggak ni ap...
306    gue pemakai bni juga bri liat org pd komplain ...
346                                     wasyu dalam bentuk plaintext heu
366                                     jadi ormas ajalah
383    ini bsi sama kaya ditempatku kerja mesin absen...
387                                     mantab
```

Name: processed_tweet, dtype: object

Top words: ['yg', 'serem', 'ke', 'kena', 'wkwkwk', 'ya', 'nunggu', 'ngeri', 'bsi', 'iya']

Topic 6:

```
7                                     bank jelek
11                                     hari bang kocak emng
103    entah apa yang ada dibenak pikiran pembuat apl...
125    bahasa inggrisnya kayak bahasa terjemahan buka...
145                                     pertanyaan yang sama semoga aman
181    nyimak ga bisa bayangin duit segitu ilang
190    dalam wktu buan itu bolak balik ke kanto dan k...
232    pulih mbanking gimana hellow buta apa gimana
239    diliatin donk deskripsi transaksinya apa
262    okee kak thx u sarannya udh bisa aku cobaaa
273    serem amat bang harus di lanjutkan bang biar k...
282    weh rombongan yaa memanfaatkan momentum
```

Name: processed_tweet, dtype: object

Top words: ['bank', 'jugaaa', 'samaaa', 'bulan', 'tipu', 'jelek', 'itu', 'april', 'bye', 'bang']

Topic 7:

```
0    sistem sama orangnya harus di ganti gak ada si...
3    gimana sih min ini kok gak bisa lagi bsi mobil...
4                                     bca sy bisa nih kayaknya
5    normal apaan semalem mo tarik tunai masih nga ...
10   iya ini kobisa ilo kena hack gabisa njajan di ...
12   bisaa tapi transfernya antar bank bsi kakak
14   biasa nya ga selama ini loh ini udh smpe hari
15   udah atm abis ilang mbanking gabisa udah hari ...
17   masih belum bisa pake mbanking ya min
18   aku masih ada tmrw sama hana sih tp gaji dr bs...
19   min bsi mobile masih gangguan sampe hari ini b...
21   iya betul ni yg ngehack lagi prepare buat pilkada
```

Name: processed_tweet, dtype: object

```
Top words: ['ke', 'udah', 'ya', 'di', 'min', 'masih', 'belum', 'ini', 'bsi', 'bisa']
```

```
for topic_num in range(8):  
    topic_words = lda.components_[topic_num]  
    top_words_indices = topic_words.argsort()[-10:]  
    top_words = [vectorizer.get_feature_names_out()[i] for i in  
top_words_indices]  
    print(f"Top words: {top_words}")
```

```
Top words: ['kalau', 'hubungi', 'ke', 'coba', 'urgent', 'center',  
'terdekat', 'call', 'kc', 'ka']
```

```
Top words: ['serangan', 'lockbit', 'dan', 'yang', 'nasabah', 'data',  
'syariah', 'bsi', 'bank', 'indonesia']
```

```
Top words: ['perbaiki', 'yang', 'sm', 'bsi', 'yg', 'ada', 'bank',  
'alhamdulillah', 'parah', 'penipu']
```

```
Top words: ['lu', 'ziva', 'selasa', 'lockbit', 'nabila', 'milikmu',  
'bsi', 'ipk', 'khairunnisa', 'bank']
```

```
Top words: ['syariah', 'ya', 'ada', 'nya', 'itu', 'ini', 'yg', 'di',  
'bsi', 'bank']
```

```
Top words: ['yg', 'serem', 'ke', 'kena', 'wkwkwk', 'ya', 'nunggu',  
'nger', 'bsi', 'iya']
```

```
Top words: ['bank', 'jugaaa', 'samaaa', 'bulan', 'tipu', 'jelek',  
'itu', 'april', 'bye', 'bang']
```

```
Top words: ['ke', 'udah', 'ya', 'di', 'min', 'masih', 'belum', 'ini',  
'bsi', 'bisa']
```

Analisis Persona Berdasarkan Topic Extraction (Ringkas)

Topic 0:

- **Persona:** Pemberi saran dan pencari solusi.
- **Informasi:** Sarankan penggunaan ATM jika mobile banking bermasalah.

Topic 1:

- **Persona:** Khawatir serangan siber.
- **Informasi:** Diskusi tentang serangan siber dan kebocoran data nasabah.

Topic 2:

- **Persona:** Tidak puas dengan layanan.
- **Informasi:** Keluhan tentang layanan buruk dan kredibilitas bank.

Topic 3:

- **Persona:** Bingung dengan masalah teknis.
- **Informasi:** Meminta klarifikasi masalah sistem dan aplikasi.

Topic 4:

- **Persona:** Marah dengan manajemen.

- **Informasi:** Komentar negatif tentang manajemen dan kekhawatiran data nasabah.

Topic 5:

- **Persona:** Frustrasi dengan pelayanan pelanggan.
- **Informasi:** Keluhan tentang customer service dan sistem tidak efisien.

Topic 6:

- **Persona:** Menganggap bank buruk.
- **Informasi:** Komentar negatif tentang kualitas layanan BSI.

Topic 7:

- **Persona:** Mengeluh tentang mobile banking.
- **Informasi:** Keluhan tentang mobile banking yang bermasalah dan belum terselesaikan.

d. Lakukan persona analysis kembali untuk topik yang anda peroleh, bandingkan apa yang anda peroleh dari pendekatan clustering dan topic extraction. Paparkan hal-hal menarik yang berhasil anda ekstrak

```
kmeans = KMeans(n_clusters=best_k, random_state=42)
df['Cluster'] = kmeans.fit_predict(X)

for cluster_num in range(best_k):
    cluster_tweets = df[df['Cluster'] == cluster_num]
    ['processed_tweet']
    print(f"\nCluster {cluster_num}:")
    print(cluster_tweets.head(12))

df['Topic'] = lda.transform(X).argmax(axis=1)

for topic_num in range(8):
    topic_tweets = df[df['Topic'] == topic_num]['processed_tweet']
    print(f"\nTopic {topic_num}:")
    print(topic_tweets.head(12))
```

Cluster 0:

```
31 ada yang transfer dari hari senin sampai sekar...
42                                     udah hari ke
46 mau transfer ke bank lain dari bsi udh aman bl...
50 halo admin saya pagi ini top sudah klik ok mal...
57 saya terakhir ke bsi customer service nya masi...
60 iyaa kak sama saya juga tranfer ke bri lewat m...
65 nah ini yg menarik di twit awal saya saya tuli...
69 dari hari jumat tanggal mei tf bsi mandiri den...
71 sdh hari saya ngga bisa pakai mobile banking b...
74 di aku msh ke menu saldo tiap ngeklik qris ata...
79 saya pernah rb dan bodohnya masih ga percaya k...
87 kirain internet saya yg bermasalah gara gabisa...
```

Name: processed_tweet, dtype: object

Cluster 1:

```
4          bca sy bisa nih kayaknya
5    normal apaan semalem mo tarik tunai masih nga ...
25         kok bisa malah suruh pinjeemm
36    nahnh gak bisaa wkwkwk sabar ya top aja gk bisa
44    parah banget gak bisa dipake sampe sekarang ce...
64         udah bisa narik di atm min
80         atm bsi udah bisa kah
81    tolong donggg ini mau dipake transaksi dari ke...
82    admin permisi mau tanya info atm bsi area jogj...
94         tarik tunai bisa tidak
96    kapok banget cuma punya mbanking nya doang ga ...
97         pakai atm bca konvensional bisa pak
```

Name: processed_tweet, dtype: object

Cluster 2:

```
0    sistem sama orangnya harus di ganti gak ada si...
1          gak gitu
2    iya nasabah bank konven sprti bribnimandiri di...
6    wes bener bener dadi kampus obama malah dadi b...
7          bank jelek
8    mau sampe kapan klen error tiap jam ku pantau ...
9          lol ini respon paling bikin ngakak
10   iya ini kobisa ilo kena hack gabisa njajan di ...
11          hari bang kocak emng
12         bisaa tapi transfernnya antar bank bsi kakak
13   hai sobatrupiah sebelumnya kami turut bersimpa...
16   astaghfirulah begitu rupanya mbah es ya mbah u...
```

Name: processed_tweet, dtype: object

Cluster 3:

```
27          yg pulih mbanking nya kapan kak
37         masii belum bisa kak aku udah coba dari tadi
43   aku brusan transfer kak dh bisa lewat mbanking...
47   ituu dicoba terus aja kak aku awalnya juga gak...
98   kak transfer ke bank lain pake bsi udah bisa k...
122  belum semuanya bisa kak itu bertahap pemulihan...
127         hai kak info apakah sudah bisa atm bsi nya
128  coba ke cabang terdekat aja kak udah mulai pul...
155  iya kak tiap setengah jam sekali udah cek soal...
201         sudah berkala kak setiap jam blm bisa jg
233  coba dicek berkala aja kak kalau buat tfsama c...
250  sama kak aku jg kayak gitu nanti kalau saldony...
```

Name: processed_tweet, dtype: object

Cluster 4:

```
188          bisa dm aja kaa
276         admin cek dm pleaseeee
```

```
284          dm minta tolong bantingin
624          cek dm min
726          min tolong baca dm saya
800          hallo tolong cek dm saya
844          cek dm min
936          cek dm min
955          miiiii cek dm miiinn
976      halo admin tolong balas dm saya ya
1001         cek dm
1210         cek dm
```

Name: processed_tweet, dtype: object

Cluster 5:

```
17          masih belum bisa pake mbanking ya min
28      min belum bisa beli pulsa top ewallet dll dari...
61      ternyata hingga hari kelima ini bsi mobile mas...
75      hr ini sy belum bisa transfer di app atau di a...
93          masih
118         serius hari belum beneerrr
137         masih belum bisa transfer antar bank
148      ini apk mobile nya kok belum bisa juga ya di o...
157         kok sy belum bisa ya
169      apa apaan ini bsi sampai hari belum bisa digun...
200      sudah hari ke n bsi mobile msh belum bisa digu...
206      emas belum bisa diakses tolong segera ditindak...
```

Name: processed_tweet, dtype: object

Cluster 6:

```
3          gimana sih min ini kok gak bisa lagi bsi mobil...
14          biasa nya ga selama ini loh ini udh smpe hari
19      min bsi mobile masih gangguan sampe hari ini b...
24      ini gimnaa ya gimannaaa masih eror kah mbankin...
34      kapan bisa transaksi di marketplace semua masi...
49      adakah yang bisa memberikan penjelasan dari te...
56      min tolong benerin geh gaji misua di bsi lah k...
67          gangguan mbanking ini kapan selesainya
68          lagi maintenance udah hari
86      gw pindahin juga uangnya ini error dari minggu...
90      assalamualaikum min mohon info mbanking lagi k...
101         udah aman hari ini
```

Name: processed_tweet, dtype: object

Cluster 7:

```
15      udah atm abis ilang mbanking gabisa udah hari ...
63      pantesan dari kemaren gue mau transfer gabisa ...
138         ko barusan gabisa lagi
139      min gimana nih masih error gabisa tf dan tarik...
140      sampe kapan mbanking gabisa dipake kasian anak...
156         lama bgt sich kack errornya ak gabisa tf nich
175      min ini kenapa masih gabisa top shopeepay ya
```

```
196 iya anjrit aku juga mo top udah beberapa hari ...
205 gabisa kak boro cek saldo blabla ini br klik l...
217 mbanking udah bisa blm tolong dong di atm jg g...
254 yang bener ajaa sih ini saya mau buka mbanking...
302 ini mo sampe kapan ya mbanking nya gabisa top ...
Name: processed_tweet, dtype: object
```

Cluster 8:

```
39 kalau topup ewallet msh pemulihan setau ku ka ...
73 sudah mulai kembali pulih ka atm dan kc
131 iya ka ini yang saya tahu proses pemulihan cob...
218 iya ka msh pemulihan terus tapi coba berkala a...
268 sabar ya ka aku juga nunggu mobilenya ini semo...
280 sekarang udh bisa belum daerah sana ka atm di ...
303 topup ewallet ya ka atau topup apa emoney bisa...
313 saran ku kalau urgent bgt bisa ke atm ka
324 aku bisa ka nyoba kemarin soalnya kalau mobile...
326 mabanking ku udh bisa ini ka coba aja di cek ka
348 ka saranku kalau urgent bisa cek di atm ka ata...
409 saran ku tf jgn via bi fast transfer online bi...
Name: processed_tweet, dtype: object
```

Topic 0:

```
6 wes bener bener dadi kampus obama malah dadi b...
20 ahjangan mengadingadi sampean minga bisadari t...
54 wani ne akun palsu koment di akun real
134 wanjer
161 gua gak bisa transfer weiiii kapan betulnya si...
184 ada pernyataan resmi dari mereka
204 klo hanya liat ui bsi oke banget sih dibanding...
216 apakah mungkin dibalik semua ini ada persaing...
257 hallloo dari bunga mayang lampung utara berapa...
313 saran ku kalau urgent bgt bisa ke atm ka
330 bukan masalah pindah bank atau ga nya ingat na...
348 ka saranku kalau urgent bisa cek di atm ka ata...
Name: processed_tweet, dtype: object
```

Topic 1:

```
8 mau sampe kapan klen error tiap jam ku pantau ...
13 hai sobatrupiah sebelumnya kami turut bersimpa...
33 memalukan banget
38 hai candra sebelumnya kami turut bersimpati at...
62 duitmu ilang ra dab
84 serangan siber yang membuat pelayanan pt bank ...
85 iyaa masak muncul permintaan tidak diproses mulu
119 andaipun data yg di ambil tidak penting tetapi...
121 sejak akhir april dari gambar terlihat saldo j...
130 saham bank besar kompak turun usai data bsi bo...
133 indonesia tengah dibuat geger dengan serangan ...
149 ojk harus tegur krn gangguan sistem miliknya s...
```

Name: processed_tweet, dtype: object

Topic 2:

```
25          kok bisa malah suruh pinjeemm
35  narik pulsanya banyak cuma buat dengerin lagu ...
182 bsi bergerak cepat nih memperbaiki gangguan ma...
207          harus diantisipasi
272          iyes bcs se urgent itu mau transkasi
275  ketua forum keamanan cyber sendiri bilang kalo...
325          menarik untuk disimak
410          kok orang bnyk uitnya yaa
439  parah banget ya ternyata bank bsi ga bisa dipe...
483          hari ketiga ngandat
485          oalah
523          hilang nya nopo
```

Name: processed_tweet, dtype: object

Topic 3:

```
40          kasus paralon apa sih
72          pan nanyanya yg net banking min
112         windowsnya ori apa ga aja gatau
147  yakin masukin bri saham bris dipegang oleh bri
152  harusnya dirutnya bikin konpers minta maaf dan...
156  lama bgt sich kack errornya ak gabisa tf nich
171  buktinya kuat di rekaman cctv terlihat jelas s...
176  bank aceh gt bsi lagian anggota dpra naon bang...
187  bank muammalat seperti mendapat durian runtuh ...
240          ingat dosa
251  tottebag muat banyak bank bsi bca syariah khai...
396          duo infamous bank bsi sama bsi
```

Name: processed_tweet, dtype: object

Topic 4:

```
1          gak gitu
2  iya nasabah bank konven sprti bribnimandiri di...
9          lol ini respon paling bikin ngakak
16  astaghfirullah begitu rupanya mbah es ya mbah u...
29          coba tolong di perhatikan lg
48  bayangin se pusing apa manajemen bank bsi seka...
51  data bocor ga hanya instansi ini beberapa inst...
55  bank bsi malah kena serangan hoax begini ruqya...
65  nah ini yg menarik di twit awal saya saya tuli...
76  masih banyak anak bangsa yg mampu dan pastinya...
77  dept kalian gk profesional lebih baik di cut s...
89  bahaya data nasabah juga bocor ini ada beberap...
```

Name: processed_tweet, dtype: object

Topic 5:

```
22          anjing lu
26  iya setuju maksudnya juga jangan taruh telur d...
```



```

57      saya terakhir ke bsi customer service nya masi...
104                                     okeh aku coba thank
165      iya kalo arsitektur data backupnya bagus reboo...
247      enih mahasiswa duitnya smua di bank mau beli s...
289      sampai kpn bosss iuran udah pada nunggak ni ap...
306      gue pemakai bni juga bri liat org pd komplain ...
346                                     wasyu dalam bentuk plaintext heu
366                                     jadi ormas ajalah
383      ini bsi sama kaya ditempatku kerja mesin absen...
387                                     mantab

```

Name: processed_tweet, dtype: object

Topic 6:

```

7                                     bank jelek
11                                     hari bang kocak emng
103      entah apa yang ada dibenak pikiran pembuat apl...
125      bahasa inggrisnya kayak bahasa terjemahan buka...
145                                     pertanyaan yang sama semoga aman
181      nyimak ga bisa bayangin duit segitu ilang
190      dalam wktu buan itu bolak balik ke kanto dan k...
232      pulih mbanking gimana hellow buta apa gimana
239      diliatin donk deskripsi transaksinya apa
262      okee kak thx u sarannya udh bisa aku cobaaa
273      serem amat bang harus di lanjutkan bang biar k...
282      weh rombongan yaa memanfaatkan momentum

```

Name: processed_tweet, dtype: object

Topic 7:

```

0      sistem sama orangnya harus di ganti gak ada si...
3      gimana sih min ini kok gak bisa lagi bsi mobil...
4                                     bca sy bisa nih kayaknya
5      normal apaan semalem mo tarik tunai masih nga ...
10     iya ini kobisa ilo kena hack gabisa njajan di ...
12     bisaa tapi transfernya antar bank bsi kakak
14     biasa nya ga selama ini loh ini udh smpe hari
15     udah atm abis ilang mbanking gabisa udah hari ...
17     masih belum bisa pake mbanking ya min
18     aku masih ada tmrw sama hana sih tp gaji dr bs...
19     min bsi mobile masih gangguan sampe hari ini b...
21     iya betul ni yg ngehack lagi prepare buat pilkada

```

Name: processed_tweet, dtype: object

Kesamaan Persona: Masalah utama pengguna seperti transfer antar bank, masalah dengan mobile banking, dan kekhawatiran tentang serangan siber diidentifikasi melalui clustering dan topic extraction.

Kekhususan Topik: Clustering menyatukan tweet berdasarkan keluhan atau masalah yang sama.

Tema-tema tertentu seperti serangan siber, solusi, dan komentar manajemen biasanya menjadi fokus dari Topic Extraction.

Tambahan informasi tentang topik ekstraksi: Topik 1 membahas kekhawatiran tentang serangan siber dan kebocoran data, yang tidak begitu jelas terlihat dalam clustering. Topik 0 memberikan rekomendasi praktis seperti menggunakan ATM, yang tidak muncul sebagai kelompok terpisah dalam clustering.

Ekstraksi Informasi Berharga: Clustering memberikan gambaran umum tentang jenis keluhan utama, tetapi kurang spesifik dalam mengidentifikasi tema tertentu. Topic Extraction memungkinkan identifikasi topik-topik spesifik seperti rekomendasi tentang penggunaan ATM dan diskusi tentang serangan siber, memberikan wawasan yang lebih terfokus.

Kesimpulan: Clustering metode memberikan gambaran umum tentang kelompok keluhan utama pengguna, sedangkan topic extraction dengan LDA memberikan pemahaman yang lebih rinci dan spesifik tentang masalah yang dihadapi pengguna dalam keluhan mereka. Kombinasi kedua metode ini dapat membantu memberikan pemahaman yang lebih baik tentang masalah yang dihadapi pengguna dan cara mereka menyampaikan keluhan mereka.