

# Pengembangan Aplikasi Web Untuk Mengklasifikasikan Penyebaran Informasi Akun Twitter Palang Merah Indonesia

Yefta Christian<sup>\*1</sup>, Amir<sup>2</sup>

<sup>1,2</sup>Universitas Internasional Batam

E-mail: <sup>\*1</sup>yefta@uib.ac.id, <sup>2</sup>amiruibbtm@gmail.com

## Abstrak

Kelangkaan darah dapat menjadi salah satu faktor utama hidup dan mati bagi pasien yang membutuhkan darah. Oleh sebab itu organisasi yang bekerja dalam pengelolaan darah Palang Merah Indonesia berupaya menyebarkan informasi untuk mengajak orang mendonorkan darah. Penyebaran informasi mengenai donor darah yang sering dilakukan yaitu melalui media sosial. Twitter adalah salah satu media sosial yang digunakan dalam penyebaran informasi tersebut. Metode yang digunakan untuk menyelesaikan sistem berupa metode waterfall dengan melakukan perencanaan konsep, pemodelan, pengerjaan sistem, pengujian hingga implementasi sistem. Informasi-informasi yang di sebarakan memiliki jumlah yang sangat banyak. Setiap informasi dapat diklasifikasikan menjadi beberapa kelompok data. Ada banyak cara untuk melakukan klasifikasi data informasi, salah satu jenis klasifikasi yang digunakan dalam pembuatan sistem yaitu *unsupervised classification* dimana klasifikasi dilakukan tanpa pengawasan atau sering disebut *clustering*. Metode *clustering* yang digunakan yaitu metode *k-means*. Teknik ini mempunyai kemampuan dalam klasifikasi data yang cukup besar jumlahnya relatif cepat dan juga efisien. Dalam *clustering* metode *k-means* diperlukan jumlah cluster optimal agar hasil dari *clustering* lebih baik. Untuk menentukan jumlah cluster yang optimal, Digunakan metode *Elbow* untuk mendapatkan nilai cluster terbaik. Hasil dari pengelompokan data disajikan dalam bentuk visualisasi, sehingga informasi menjadi lebih mudah dipahami dengan cepat karena didukung grafik dan tabel serta warna-warna yang responsif. Hasil output penelitian ini berupa sistem analisis penyebaran informasi berbasis web yang dapat mengklasifikasi penyebaran informasi yang didapatkan dari akun twitter Palang Merah Indonesia dan di visualisasikan kedalam bentuk diagram-diagram.

**Kata Kunci**— *Clustering, K-Means, Elbow, Visualisasi*

## 1. PENDAHULUAN

Salah satu cairan penting dalam tubuh kita yaitu darah. Cairan ini terdiri dari jaringan yang bertugas untuk mengedarkan sari makanan, mengangkut oksigen, serta mengedarkan hormon ke seluruh bagian tubuh [1]. Darah manusia diproduksi dari tubuh manusia dan dapat didonorkan kepada manusia lain dengan jumlah yang bukan tidak terbatas. Saat ini, belum ada produk lain yang dapat menjadi pengganti darah walaupun melalui proses kimia untuk menghasilkan darah. Dikarenakan kelangkaan darah sehingga darah dianggap sebagai komoditas yang sangat berharga. Rantai pasok darah merupakan sistem yang kompleks dan melibatkan berbagai parameter berbeda yang saling berhubungan. Beberapa parameter yang berhubungan seperti rumah sakit, bank darah, dan pendonor. Rantai pasok darah memiliki aspek-aspek sensitif lainnya seperti tingkat kebutuhan yang tinggi, sehingga merupakan salah satu faktor utama hidup dan mati bagi pasien. Organisasi atau Instansi yang menjadi bagian penting dalam pasokan darah yaitu bank darah dan rumah sakit, sehingga instansi tersebut berupaya untuk meminimalkan

kekurangan darah maupun darah yang kadarluarsa. Tantangan paling besarnya adalah ketidakpastian pada permintaan darah serta terbatasnya masa simpan darah [2].

Di Indonesia terdapat suatu organisasi yang bekerja dalam pengelolaan darah yaitu Palang Merah Indonesia. Ketersediaan darah sangat perlu di perhatikan untuk menolong pasien yang membutuhkan darah [3]. Organisasi ini juga berupaya dengan berbagai cara untuk meningkatkan ketersediaan darah. Salah satu cara untuk meningkatkan ketersediaan darah yaitu dengan meningkatkan penyebaran informasi akan kebutuhan darah. Kurangnya informasi dan kepedulian masyarakat akan kebutuhan darah dapat menghambat ketersediaan darah untuk orang yang membutuhkan. Pada zaman sekarang ini cara yang sering digunakan untuk meningkatkan penyebaran informasi yaitu melalui media sosial. Cara kerja media sosial juga sangat mudah dan cepat. Seseorang dapat menyebarkan informasi hanya menggunakan ponsel dan internet. Selain itu media sosial juga merupakan salah satu media yang mempermudah komunikasi [4]. Di era ini juga terdapat banyak platform-platform media sosial yang dapat digunakan. Berbagai media sosial memiliki pasar masing-masing, seperti twitter kebanyakan orang menggunakannya untuk membagikan informasi terkini. Bukan hanya berupa tulisan saja informasi yang dapat dibagikan melainkan video juga dapat dibagikan. Hal ini membuat banyaknya pengguna menggunakan media sosial platform ini. Jumlah pengguna juga dapat mempengaruhi tingkat penyebaran suatu informasi pada masyarakat. Indonesia juga menempati peringkat 5 pengguna twitter terbesar didunia pada tahun 2019 [5].

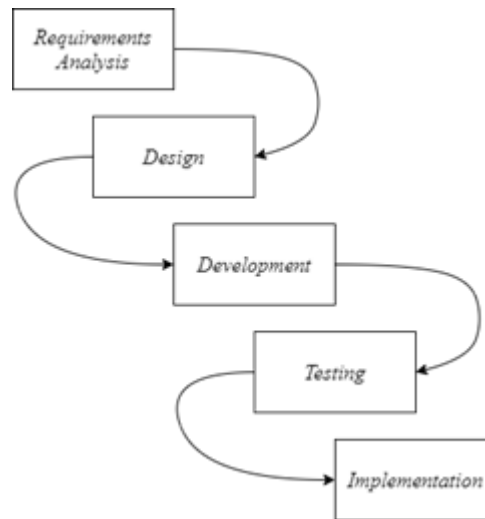
Informasi-informasi yang di sebarakan juga memiliki jumlah yang sangat banyak. Setiap informasi dapat klasifikasikan menjadi beberapa kelompok data. Klasifikasi merupakan suatu proses pencarian dari kumpulan model atau fitur data yang dapat digunakan untuk membedakan label kelas data. Tujuan tersebut agar model yang terbentuk dapat digunakan untuk membedakan kelompok data atau untuk memprediksi data dari suatu objek dengan tepat. Klasifikasi menjadi salah satu tugas dari pembelajaran mesin (machine learning). Ada banyak cara untuk melakukan klasifikasi data informasi, salah satu jenis klasifikasi yaitu unsupervised classification dimana klasifikasi dilakukan tanpa pengawasan atau sering disebut clustering. Dalam clustering metode yang sering digunakan yaitu metode k-means. Teknik ini mempunyai kemampuan dalam klasifikasi data yang cukup besar jumlahnya relatif cepat dan juga efisien. Dalam clustering metode k-means diperlukan jumlah cluster optimal agar hasil dari clustering lebih baik. Untuk menentukan jumlah cluster yang optimal, dapat menggunakan metode Elbow untuk mendapatkan nilai cluster terbaik [6]. Informasi yang telah dipisahkan kedalam kelas data, dapat digunakan untuk mengetahui bagaimana penyebaran suatu informasi pada kelas masing-masing. Informasi yang banyak cenderung membuat orang sulit untuk menangkap isi dari informasi tersebut. Oleh karena itu diperlukan cara untuk menyajikan hasil dari informasi itu menjadi suatu bentuk yang mudah di mengerti. Salah satu caranya yaitu dengan mengvisualisasikan hasil dari informasi yang telah di bagi menjadi kelas-kelas tersebut. Dengan menggunakan visualisasi data, informasi yang ada menjadi lebih mudah dipahami dengan cepat karena didukung grafik dan tabel serta warna-warna yang responsif [7].

Berdasarkan dari latar belakang yang telah dibahas, maka penulis berupaya untuk menerapkan aplikasi guna menyelesaikan masalah tersebut. Rancangan ini penulis rangkum dalam skripsi yang berjudul “Pengembangan Aplikasi Web untuk Mengklasifikasikan Penyebaran Informasi dari Akun Twitter Palang Merah Indonesia”. Dengan perancangan ini, penulis berharap dapat membantu Palang Merah Indonesia untuk mengetahui penyebaran informasi yang telah diklasifikasi pada akun Palang Merah Indonesia di media sosial twitter.

## 2. METODE PENELITIAN

Metode yang akan digunakan dalam penelitian ini oleh penulis adalah metodologi Waterfall. Waterfall merupakan salah satu metode dalam pengembangan software yang

pengerjaannya harus dilakukan secara berurutan. Tahapan dimulai dari tahap perencanaan konsep (requirements analysis), kemudian pemodelan (design), setelah itu develop atau pengerjaan (development), baru melakukan pengujian (testing) dan pada akhirnya implementasi (implementation).



Gambar 1. Metode Penelitian

### 2.1. Requirements Analysis

Pada tahap awal ini penulis akan melakukan diskusi bersama dosen pembimbing yang bertujuan untuk memahami kebutuhan dan batasan sistem yang akan di buat. Aplikasi web yang dirancang diharapkan dapat memberikan informasi yang lengkap.

### 2.2. Design

Pada tahap desain, akan dilakukan penerjemahan dari syarat dan kebutuhan ke suatu perancangan desain perangkat lunak yang dapat diperkirakan sebelum dibuatnya proses pengkodean (coding). Proses ini akan berfokus pada proses alur kerja sistem menggunakan flowchart, struktur data dengan menggunakan ER Diagram, arsitektur perangkat lunak dengan menggunakan Usecase Diagram, dan representasi interface dengan membuat mockup tampilan.

### 2.3. Development

Pada tahap ini akan dilakukan proses menerjemahkan perancangan desain ke bentuk yang dapat di mengerti oleh mesin, dengan menggunakan kode-kode bahasa pemrograman. Kode program yang dihasilkan dari modul-modul kecil dan kemudian akan digabungkan menjadi satu. Proses development juga dibagi menjadi beberapa bagian yaitu:

1. Sync Tweet Script
2. Analyze Tweet Script
3. Web Interface System

### 2.4. Testing

Di tahap ini akan dilakukan pengujian sehingga dapat mengetahui apakah sistem yang dibuat telah sesuai dengan desain. Pengujian dilakukan dengan black-box testing untuk memastikan sistem sesuai dengan desain dan fungsi yang dirancang. Selain itu pengujian juga

untuk mengetahui fungsi pada sistem sudah berjalan dengan baik atau tidak. Pengujian dilakukan dengan mengetest satu persatu fitur/fungsi pada sistem.

### 2.5. Implementation

Ini merupakan tahapan terakhir dalam metode model waterfall. Setelah melalui tahap testing jika sistem telah sesuai dengan desain dan fungsi yang di rancang maka akan dilakukan tahap implementasi sistem untuk digunakan.

## 3. HASIL DAN PEMBAHASAN

### 3.1. Requirements Analysis

Pada tahap requirement analysis yang telah dilakukan, sistem yang dibuat membutuhkan beberapa fitur untuk mendukung sistem. Fitur – fitur tersebut mencakup:

#### 1. Login

Sebagai fitur untuk validasi pengguna sistem agar dapat masuk kedalam sistem.

#### 2. Dashboard

Sebagai fitur halaman utama sistem untuk melihat hasil dari analisis tweet.

#### 3. Tweets

Sebagai fitur untuk melihat semua tweet yang telah disimpan di sistem.

#### 4. Sync Tweets

Sebagai fitur untuk sinkronisasi data tweet Palang Merah Indonesia di twitter ke dalam sistem.

#### 5. Analysis Tweets

Sebagai fitur untuk melakukan proses algoritma untuk menghasilkan hasil analisis tweet.

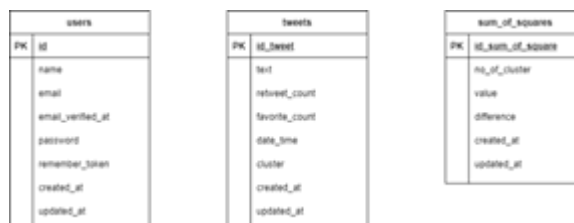
#### 6. Logout

Sebagai fitur untuk keluar dari sistem.

### 3.2. Design

Dalam perancangan sistem ini menggunakan beberapa diagram – diagram pendukung dalam mendesain pemodelan sistem. Berikut diagram – diagram yang digunakan yaitu:

#### 1. ER Diagram

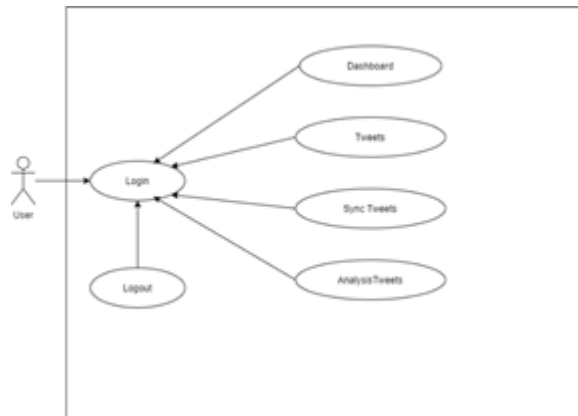


Gambar 2. ER Diagram Database Sistem

Dalam sistem ini terdapat 3 tabel dalam database. Tabel users merupakan tabel yang digunakan untuk menyimpan data user. Table tweets merupakan tabel yang digunakan untuk menyimpan seluruh data informasi tweet beserta cluster tweet. Sedangkan tabel sum\_of\_squares merupakan tabel yang digunakan untuk menyimpan hasil data sum of square error menggunakan metode elbow.

2. Use Case Diagram

Use case diagram merupakan pemodelan untuk menggambarkan kelakuan dari sistem yang dibuat dan mendeskripsikan sebuah interaksi antara satu atau lebih actor dengan sistem yang dibuat, serta digunakan untuk mengetahui fungsi apa saja yang ada didalam sebuah sistem dan siapa saja yang berhak menggunakan fungsi-fungsi tersebut.



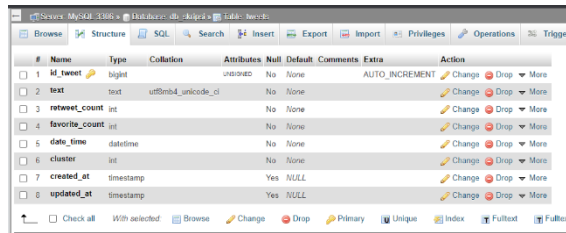
Gambar 3. Use Case Diagram

Pada gambar 3, digambarkan mengenai use case diagram dari sistem yang dibuat dimana terdapat 1 aktor yaitu user. Aktor user memiliki hak untuk login, mengakses dashboard, menampilkan informasi tweets, melakukan sinkronisasi data tweet dari akun twitter palang merah Indonesia ke sistem, melakukan proses analisis data tweets dengan metode elbow dan k-means, beserta fungsi logout.

Pada tahap desain yang telah dilakukan, menghasilkan sebuah database pendukung sistem yang dirancang. Database ini digunakan untuk menyimpan seluruh data informasi sistem.

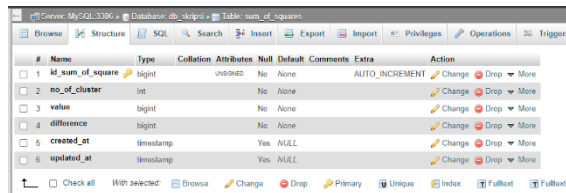
#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Actions
1	id	integer(11)	utf8mb4_general_ci		No	None		AUTO_INCREMENT	Change Stop Move
2	name	varchar(191)	utf8mb4_general_ci		No	None			Change Stop Move
3	email	varchar(191)	utf8mb4_general_ci		No	None			Change Stop Move
4	email_verified_at	timestamp			Yes	NULL			Change Stop Move
5	password	varchar(191)	utf8mb4_general_ci		No	None			Change Stop Move
6	remember_token	varchar(191)	utf8mb4_general_ci		Yes	NULL			Change Stop Move
7	remember_me	boolean			Yes	NULL			Change Stop Move
8	updated_at	timestamp			Yes	NULL			Change Stop Move

Gambar 4. Database tabel users



#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
1	id_tweet	bigint	utf8mb4	UNSIGNED	No	None		AUTO_INCREMENT	Change Drop More
2	text	text	utf8mb4_unicode_ci		No	None			Change Drop More
3	retweet_count	int			No	None			Change Drop More
4	favorite_count	int			No	None			Change Drop More
5	date_time	datetime			No	None			Change Drop More
6	cluster	int			No	None			Change Drop More
7	created_at	timestamp			Yes	NULL			Change Drop More
8	updated_at	timestamp			Yes	NULL			Change Drop More

Gambar 5. Database tabel tweets



#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra	Action
1	id_sum_of_square	bigint	utf8mb4	UNSIGNED	No	None		AUTO_INCREMENT	Change Drop More
2	no_of_cluster	int			No	None			Change Drop More
3	value	bigint			No	None			Change Drop More
4	difference	bigint			No	None			Change Drop More
5	created_at	timestamp			Yes	NULL			Change Drop More
6	updated_at	timestamp			Yes	NULL			Change Drop More

Gambar 6. Database tabel sum\_of\_squares

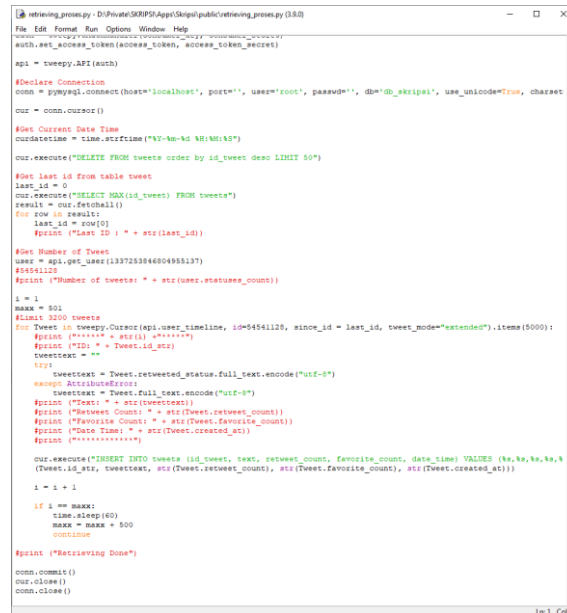
### 3.3. Development

Pada tahap development dilakukan pengerjaan script pengambilan data tweet, script melakukan analisis tweet dan aplikasi web.

#### 3.3.1. Sync Tweet Script

Sync Tweet Script merupakan script yang dibuat untuk mengambil data postingan dari akun twitter PMI. Script ini menggunakan bahasa pemrograman python dan juga library tweepy. Proses yang terdapat pada script ini yaitu:

1. Membuat koneksi ke api twitter PMI.
2. Membuat koneksi ke database sistem.
3. Menghapus 50 data terakhir dari database, agar 50 data terakhir dapat diupdate dari data terbaru.
4. Mengambil data terbaru dari twitter.
5. Menyimpan data kedalam database.



```

retrieving_process.py - D:\Private\SRIPS\Appl\Script\public\retrieving_process.py (3.8)
File Edit Format Run Options Window Help
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth)

#Declare Connection
conn = pymysql.connect(host='localhost', port=3306, user='root', passwd='', db='db_krispal', use_unicode=True, charset='utf8')
cur = conn.cursor()

#Get Current Date Time
curdatetime = time.strftime("%Y-%m-%d %H:%M:%S")
cur.execute("DELETE FROM tweets order by id_tweet desc LIMIT 50")

#Get last id from table tweet
last_id = 0
cur.execute("SELECT MAX(id_tweet) FROM tweets")
result = cur.fetchall()
for row in result:
    last_id = row[0]
#print ("Last ID : " + str(last_id))

#Get Number of Tweet
user = api.get_user(13972334680495117)
#54541128
#print ("Number of tweets: " + str(user.statuses_count))

i = 1
maxk = 501
#LIMIT 3200 tweets
for Tweet in tweepy.Cursor(api.user_timeline, id=54541128, since_id = last_id, tweet_mode="extended").items(5000):
    #Print ("*****" + str(i) + "*****")
    #Print ("ID: " + Tweet.id_str)
    tweettext = ""
    try:
        tweettext = Tweet.retweeted_status.full_text.encode("utf-8")
    except AttributeError:
        tweettext = Tweet.full_text.encode("utf-8")
    #Print ("Text: " + str(tweettext))
    #Print ("Retweet Count: " + str(Tweet.retweet_count))
    #Print ("Favorite Count: " + str(Tweet.favorite_count))
    #Print ("Date Time: " + str(Tweet.created_at))
    #Print ("*****")

    cur.execute("INSERT INTO tweets (id_tweet, text, retweet_count, favorite_count, date_time) VALUES (%s,%s,%s,%s,%s)"
    (Tweet.id_str, tweettext, str(Tweet.retweet_count), str(Tweet.favorite_count), str(Tweet.created_at)))

    i = i + 1

if i == maxk:
    time.sleep(10)
    maxk = maxk + 500
    continue

#print ("Retrieving Done")
conn.commit()
cur.close()
conn.close()

```

Gambar 7. Script yang digunakan untuk pengambilan data tweet dari twitter

### 3.3.2. Analyze Tweet Script

Analyze Tweet Script merupakan script yang dibuat untuk melakukan proses clustering data tweet. Script ini menggunakan bahasa pemrograman python dan juga library bantuan untuk proses clustering. Proses yang terdapat pada script ini yaitu:

1. Membuat koneksi ke database sistem.
2. Mengambil data tweet yang diperlukan dan memasukan kedalam sebuah variable objek.
3. Melakukan proses k-means sebanyak 10 kali dari jumlah 1 cluster – 10 cluster.
4. Melakukan proses elbows dari hasil cluster yang dihasilkan.
5. Menyimpan hasil dari proses elbows ke database.
6. Mengambil angka cluster optimal dari proses elbow.
7. Melakukan proses k-means sesuai dengan angka cluster optimal.
8. Melakukan update data tweet sesuai dengan cluster yang dihasilkan.

```

kmeans_elbow_groes.py - D:\Private\SKRIP2\Apl\Skrip\public\kmeans_elbow_groes.py (1.0)
File Edit Format Run Options Window Help
# library & datasets
import time
import pymysql
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from sklearn.cluster import KMeans
from scipy.spatial.distance import cdist

# Delete Connection
conn = pymysql.connect(host='localhost', port='', user='root', password='', db='db_skripsi', use_unicode=True, charset
cur = conn.cursor()

cur.execute("SELECT * FROM tweets WHERE date_time >= (LAST_DAY(NOW()) + INTERVAL 1 DAY - INTERVAL 3 MONTH) AND date

# Data Clean to K-Means Process
dataclean = np.array(['id_tweet', 'text', 'retweet_count', 'favorite_count'])
cur.execute("SELECT id_tweet, text, retweet_count, favorite_count FROM tweets WHERE date_time >= (LAST_DAY(NOW()) +
result = cur.fetchall()
for row in result:
    articlean = np.array([row[0],row[1],row[2],row[3]])
    dataclean = np.append(dataclean, articlean, axis=0)

# Set Clean to Pandas Dataframe
pdclean = pd.DataFrame(data=dataclean[1:,1:], index=dataclean[1:,0], columns=dataclean[0,1:])
pdclean.head()
#print(pdclean)
#print("\nFinish Get Data from DB")

# Choose Data Variable
# Set (Jump row no, after column 1 to 3 (column retweet_count and favorite_count))
pdclean_2 = pdclean.iloc[:, 1:3]
pdclean_2.head()
#print(pdclean_2)
#print("\nFinish Set used Data to Array")

# Elbow Method to get optimal Cluster
inertias = []
mapping = {}

K = range(1,11)
for k in K:
    # Do clustering kmeans
    kmeanModel = KMeans(n_clusters=k)
    kmeanModel.fit(pdclean_2)

    # sum of squared
    inertia.append(kmeanModel.inertia_)
    mapping[k] = kmeanModel.inertia_

    # Print cluster center point
    #print("Cluster Center")
    #print(kmeanModel.cluster_centers_)

    # Show Cluster data
    pdclean["cluster"] = kmeanModel.labels_

# Elbow Method to get optimal Cluster
inertias = []
mapping = {}

K = range(1,11)
for k in K:
    # Do clustering kmeans
    kmeanModel = KMeans(n_clusters=k)
    kmeanModel.fit(pdclean_2)

    # sum of squared
    inertia.append(kmeanModel.inertia_)
    mapping[k] = kmeanModel.inertia_

    # Print cluster center point
    #print("Cluster Center")
    #print(kmeanModel.cluster_centers_)

    # Show Cluster data
    pdclean["cluster"] = kmeanModel.labels_
    #print(kmeanModel.labels_)
    #print(kmeanModel.labels_)

#print("\nFinish Proceed sum of squared")

# Update data in database with optimal number cluster
oc = 1
curmax = 0
cursel = 0
#print("\nData sum of squared")

# Clear sum of square in DB
cur.execute("TRUNCATE TABLE sum_of_squares")

for key, val in mapping.items():
    if cursel > 1:
        selisih = cursel-val
    else:
        selisih = 0
    #print(f'{key} : {val}')
    #print('s : ', selisih)

    # Insert sum of square to DB
    cur.execute("INSERT INTO sum_of_squares (no_of_cluster, value, difference) VALUES (%s,%s,%s)",(str(key), str(val)
    cursel = val
    if selisih > curmax:
        oc = key
        curmax = selisih

#print("optimal cluster ", oc, " curmax ", curmax, " cursel ", cursel)
#print("\nFinish Check Optimal number of cluster ", oc)

# Do clustering kmeans with optimal number cluster
kmeanModel = KMeans(n_clusters=oc)
    
```

Gambar 8. Script yang digunakan untuk melakukan proses clustering part 1

```

kmeans_elbow_groes.py - D:\Private\SKRIP2\Apl\Skrip\public\kmeans_elbow_groes.py (1.0)
File Edit Format Run Options Window Help
# Elbow Method to get optimal Cluster
inertias = []
mapping = {}

K = range(1,11)
for k in K:
    # Do clustering kmeans
    kmeanModel = KMeans(n_clusters=k)
    kmeanModel.fit(pdclean_2)

    # sum of squared
    inertia.append(kmeanModel.inertia_)
    mapping[k] = kmeanModel.inertia_

    # Print cluster center point
    #print("Cluster Center")
    #print(kmeanModel.cluster_centers_)

    # Show Cluster data
    pdclean["cluster"] = kmeanModel.labels_
    #print(kmeanModel.labels_)
    #print(kmeanModel.labels_)

#print("\nFinish Proceed sum of squared")

# Update data in database with optimal number cluster
oc = 1
curmax = 0
cursel = 0
#print("\nData sum of squared")

# Clear sum of square in DB
cur.execute("TRUNCATE TABLE sum_of_squares")

for key, val in mapping.items():
    if cursel > 1:
        selisih = cursel-val
    else:
        selisih = 0
    #print(f'{key} : {val}')
    #print('s : ', selisih)

    # Insert sum of square to DB
    cur.execute("INSERT INTO sum_of_squares (no_of_cluster, value, difference) VALUES (%s,%s,%s)",(str(key), str(val)
    cursel = val
    if selisih > curmax:
        oc = key
        curmax = selisih

#print("optimal cluster ", oc, " curmax ", curmax, " cursel ", cursel)
#print("\nFinish Check Optimal number of cluster ", oc)

# Do clustering kmeans with optimal number cluster
kmeanModel = KMeans(n_clusters=oc)
    
```

Gambar 9. Script yang digunakan untuk melakukan proses clustering part 2

```

kmeans_elbow_proses.py - D:\Private\SCRIPTS\Apps\Skripsi\public\kmeans_elbow_proses.py (3.8.0)
File Edit Format Run Options Window Help
-----
for key, val in mapping.items():
    if cursor > 1:
        selisih = cursor - val
    else:
        selisih = 0
    print(f'{key} : {val}')
    print(f' : {selisih}')
# Insert sum of square to DB
cur.execute("INSERT INTO sum_of_squares (no_of_cluster, value, difference) VALUES (%s,%s,%s)", (str(key), str(val)
cursor = val
if selisih > curmax:
    oc = key
    curmax = selisih
print("Optimal cluster ", oc, " curmax ", curmax, " cursor ", cursor)
print("\nFinish Check Optimal number of cluster ", oc)

# Do clustering kmeans with optimal number cluster
kmeanModel = KMeans(n_clusters=oc)
kmeanModel.fit(pdclean_w)
# Print cluster center point with optimal number cluster
print("Cluster Center")
print(kmeanModel.cluster_centers_)
# Show Cluster data with optimal number cluster
pdclean["cluster"] = kmeanModel.labels_
print(kmeanModel.labels_)
print("\nDate Final")
print(pdclean)
# Update result to database
for index, row in pdclean.iterrows():
    print(row[0], row[2], row[3], index)
    cur.execute("UPDATE tweets SET cluster=%s WHERE id_tweet=%s", (row[3]+1, index))
# Showing Clustered Data
C = range(1, oc+1)
for c in C:
    cur.execute("SELECT id_tweet FROM tweets WHERE cluster=%s", (c))
    result = cur.fetchall()
    print("\nCluster ", c)
    k = 1
    #for row in result:
        print(k, " ", row[0])
        k = k + 1
    print("Total ", len(result), " Result")
# Commit in DB
conn.commit()
cur.close()
conn.close()

```

Gambar 10. Script yang digunakan untuk melakukan proses clustering part 3

### 3.3.3. Web Interface System

Pada bagian ini dilakukan pembuatan sistem dalam aplikasi web. Aplikasi web yang dibuat untuk menghubungkan data dalam database beserta user interface sehingga menjadi satu sistem. Aplikasi web ini menggunakan bahasa pemrograman php beserta framework pendukung seperti Laravel dan juga bootstrap. Beberapa interface yang dibuat seperti:

#### 1. Login

Interface ini dilakukan coding tampilan login berdasarkan mockup pada tahap design. Setelah tampilan selesai, dilakukan coding untuk proses login user dengan pengecekan data user pada database. Jika data user yang dimasukkan sesuai data pada database, maka sistem akan memberikan user akses ke dalam sistem.

#### 2. Layout dan Menu

Interface ini dilakukan coding tampilan layout beserta menu untuk seluruh halaman pada sistem.

#### 3. Dashboard

Interface ini dilakukan coding tampilan utama user berdasarkan mockup pada tahap design. Didalam dashboard berisi beberapa chart untuk hasil dari analisis tweet, kemudian juga terdapat tabel hasil analisis dan beberapa informasi yang dihasilkan dari analisis. Selain tampilan juga dilakukan coding untuk menarik data dari database beserta hasil dari analisis data sehingga dapat di tampilkan ke tampilan dashboard.

#### 4. Tweets

Interface ini dilakukan coding tampilan data tweet berdasarkan mockup pada tahap design. Didalam tampilan ini berisi data-data hasil dari pengambilan data twitter PMI dan juga nomor halaman untuk mempermudah user mengakses data sehingga

tampilan tidak kepanjangan. Proses pada tampilan ini dengan mengambil data tweet dari database yang sebelumnya sudah melakukan pengambilan dari twitter PMI.

#### 5. Sync Tweets

Interface ini berupa tombol menu Sync Tweets. Pada tombol ini akan dilakukan link untuk menjalankan script python sync tweet. Saat tombol di klik maka proses pengambilan data tweet dari akun twitter PMI akan di proses dan di simpan ke database.

#### 6. Analysis Tweets

Interface ini berupa tombol menu Analysis Tweets. Pada tombol ini akan dilakukan link untuk menjalankan script python analyze tweet. Saat tombol di klik maka proses algoritma akan dijalankan dan hasil analisis data tweet akan di simpan ke database.

#### 7. Logout

Interface ini berupa tombol menu logout. Pada tombol ini akan dibuat coding untuk mengeluarkan user dari sistem, sehingga sistem harus di login kembali jika ingin digunakan kembali.

### 3.4. Testing

Pada tahap testing, dilakukan proses pengujian terhadap script dan aplikasi web yang telah dibuat. Pengujian ini dilakukan agar dapat mengetahui apakah sistem yang telah dibuat sudah memenuhi kebutuhan yang diperlukan pada sistem. Selain itu pengujian juga untuk mengetahui fungsi pada sistem sudah berjalan dengan baik atau tidak. Pengujian dilakukan dengan mengetest satu persatu fitur/fungsi pada sistem.

Dari sistem yang telah dirancang, penulis melakukan pengujian pada fungsi sistem yang menghasilkan status dari pengujian sebagai berikut:

Tabel 1. Hasil pengujian pada fungsi-fungsi sistem

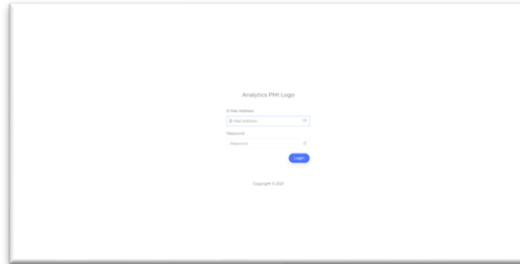
<i>Input</i>	<i>Expected</i>	<i>Output</i>	<i>Status</i>
Akses sistem awal	Sistem menampilkan halaman login	Halaman login ditampilkan	Berhasil
Email dan password benar	User berhasil masuk ke tampilan utama sistem	Halaman utama sistem ditampilkan	Berhasil
Email dan password salah	User gagal masuk ke tampilan utama sistem dan kembali ke halaman login	Halaman utama sistem tidak ditampilkan dan kembali ke tampilan login	Berhasil
Klik menu dashboard	Sistem menampilkan halaman dashboard	Halaman dashboard ditampilkan	Berhasil
Klik menu tweets	Sistem menampilkan halaman tweets	Halaman tweets ditampilkan	Berhasil

Klik menu sync tweets	Sistem melakukan proses sinkronisasi data dari twitter dengan database	Sistem melakukan proses sinkronisasi	Berhasil
Klik menu analysis tweets	Sistem melakukan proses analisis data dari tweet yang tersimpan dalam database	Sistem melakukan proses analisis tweet	Berhasil
Klik menu logout	User berhasil keluar dari sistem dan sistem menampilkan tampilan login	User keluar dari sistem dan halaman login ditampilkan	Berhasil

### 3.5. Implementation

Implementation merupakan tahapan terakhir dari seluruh proses yang dilakukan. Setelah melalui tahap testing jika sistem telah sesuai dengan desain dan fungsi yang di rancang, maka dilakukan implementasi sistem untuk digunakan. Hasil dari implementasi sebagai berikut:

#### 1. Login



Gambar 11. Tampilan halaman login hasil implementasi

Hasil dari implementasi halaman login, fitur dapat digunakan untuk user melakukan login ke dalam sistem. Pengecekan data login user berhasil dan sesuai dengan database.

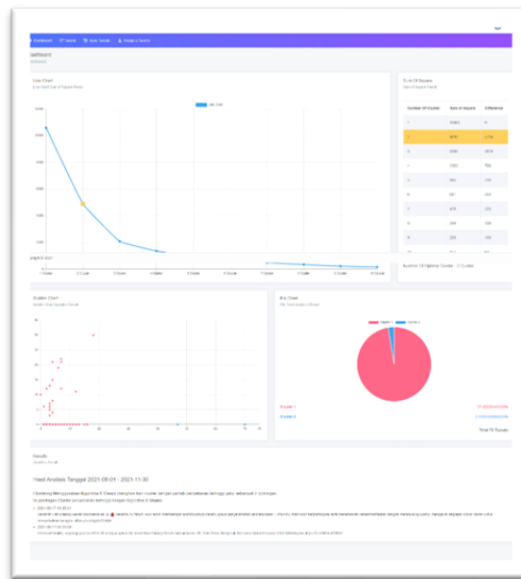
#### 2. Menu



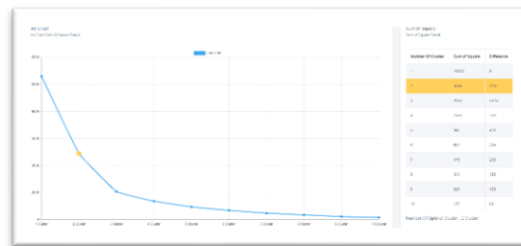
Gambar 12. Tampilan menu hasil implementasi

Hasil dari implementasi menu sistem telah sesuai dengan kebutuhan fitur dari perancangan dan design. Menu-menu tersebut dapat digunakan user untuk ke halaman yang dituju dan data yang di tampilkan sesuai dengan data di database.

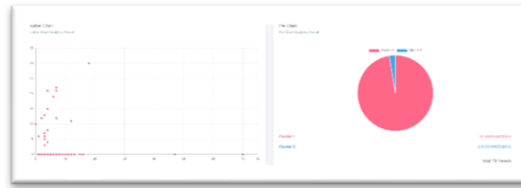
#### 3. Dashboard



Gambar 13. Tampilan halaman dashboard hasil implementasi keseluruhan Hasil dari implementasi dashboard berhasil menampilkan diagram-diagram beserta tabel dan informasi yang diperlukan. Pada dashboard terdapat line chart atau diagram garis dan juga tabel untuk mendukung informasi pada hasil metode elbows. Dashboard juga memiliki scatter chart atau diagram penyebaran dan pie chart atau diagram kue untuk menampilkan hasil analisis data dari metode k-means. Hasil dari seluruh analisis juga di tampilkan berupa tulisan pada bagian terakhir dashboard.



Gambar 14. Tampilan halaman dashboard hasil implementasi metode elbow Sebelum melakukan proses clustering data tweet, diperlukan proses analisis cluster untuk pencarian jumlah cluster optimal pada data tweet yang ada. Pada penelitian ini menggunakan metode elbow dalam proses analisis cluster. Data yang digunakan adalah data tweet PMI yang dimulai dari tanggal 01-Agustus-2021 sampai 30-November-2021. Data tiga bulan ini akan dilakukan proses metode elbow dimana clustering mulai dari jumlah 1 cluster hingga 10 cluster. Dari proses metode elbow akan dicari nilai jarak objek terdekat. Pada gambar 4.3.2 terdapat line chart untuk menggambarkan hasil elbow masing-masing jumlah cluster dan juga terdapat tabel sum of square error untuk menampilkan perbandingan jarak antar jumlah cluster. Kita dapat melihat hasil dari proses metode elbow dimana jumlah cluster yang paling optimal adalah jumlah cluster sebanyak 2.



Gambar 15. Tampilan halaman dashboard hasil implementasi metode k-means  
Melanjutkan dari proses sebelumnya, kita akan melakukan proses metode k-means berdasarkan jumlah cluster optimal yang didapatkan dari proses metode elbow. Setelah melakukan proses metode k-means kita mendapatkan hasil untuk cluster pertama sebanyak 77 tweet sedangkan cluster kedua sebanyak 2 tweet. Postingan pada cluster kedua yang merupakan cluster penyebaran informasi tertinggi yaitu:

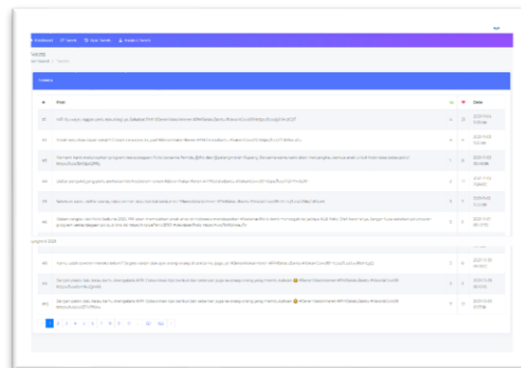
- 2021-09-17 13:45:21

Selamat Hari Palang Merah Indonesia ke-76 🇮🇩 Selama 76 tahun, PMI terus memberikan kontribusinya dalam upaya penyelamatan kemanusiaan. Untuk itu, mari turut berpartisipasi aktif menebarkan kebermanfaatannya dengan mendukung & mengikuti kegiatan donor darah untuk menyehatkan bangsa.  
<https://t.co/rgzA7Cb1jF>

- 2021-09-17 06:33:59

#TemanPemilih, segenap jajaran KPU RI mengucapkan Selamat Hari Palang Merah Indonesia ke-76. Mari Terus Bergerak Bersama Demi Sesama..!  
#KPU Melayani <https://t.co/N9asiC98hP>

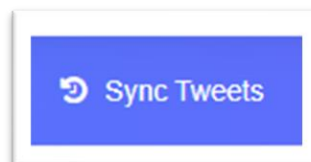
#### 4. Tweets



Gambar 16. Tampilan halaman tweets hasil implementasi

Hasil dari implementasi halaman tweets menampilkan seluruh data tweet yang telah di simpan didatabase. Halaman ini juga sudah memiliki paging atau tombol halaman data sehingga mempermudah user melihat data tweet.

#### 5. Sync Tweets

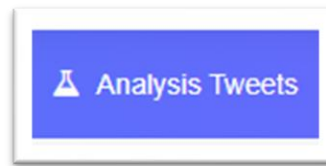


Gambar 17. Tampilan tombol menu sync tweets hasil implementasi

Hasil dari implementasi tombol menu sync tweets, data tweet dapat di ambil dari akun twitter PMI dan disimpan ke database. Proses sync tweets tergantung pada

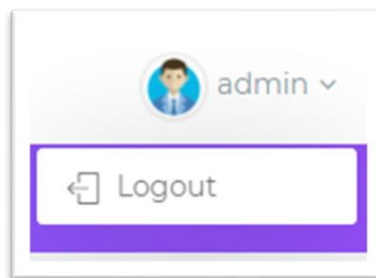
banyaknya data dari akun twitter PMI, jika banyak maka proses akan memakan waktu lebih lama.

6. Analysis Tweets



Gambar 18. Tampilan tombol menu analisis tweets hasil implementasi Hasil dari implementasi tombol menu analisis tweets, data tweet dapat dilakukan proses metode elbows dan k-means dan disimpan ke database. Proses analisis tweets tergantung pada banyaknya data tweet yang disimpan di database, jika banyak maka proses akan memakan waktu lebih lama.

7. Logout



Gambar 19. Tampilan tombol logout hasil implementasi Hasil dari implementasi tombol menu logout, user dapat keluar dari sistem setelah menekan tombol logout. Setelah logout sistem akan menuju user ke halaman login.

#### 4. KESIMPULAN

Dari proses analisis, perancangan dan pengembangan yang dilakukan untuk data tweet pada akun Palang Merah Indonesia di Twitter tanggal 01-Agustus-2021 sampai 30-November-2021, maka dapat diambil kesimpulan sebagai berikut :

1. Dengan mengimplementasikan Pengembangan Aplikasi Web Untuk Mengklasifikasikan Penyebaran Informasi Akun Twitter Palang Merah Indonesia menggunakan metode Waterfall, penulis dapat melakukan pengembangan aplikasi web dalam waktu yang lebih sesuai dan terstruktur. Dimulai dari requirement analysis sampai implementasi dilakukan sangat lancar karena secara step by step sehingga perlu menyelesaikan satu tahap baru melanjutkan ke tahap lain.
2. Proses penarikan data tweet pada akun Palang merah Indonesia dilakukan secara langsung dengan bantuan API twitter sehingga data lebih up to date. Penarikan data dapat dilakukan menggunakan script python dan koneksi ke twitter berhasil sehingga data yang ditarik dan simpan di database dapat digunakan untuk proses selanjutnya dalam mencari cluster penyebaran informasinya.
3. Proses analisis cluster dengan metode elbow berhasil dilakukan, dimana cluster optimal sebanyak 2 cluster. Proses clustering dari jumlah cluster optimal dengan metode k-means menghasilkan cluster pertama dengan jumlah 77 tweet dan cluster kedua 2 tweet. Cluster dengan jumlah penyebaran tertinggi terdapat pada cluster kedua.

5. Hasil data tweet yang telah di lakukan analisis dapat di implementasikan ke sistem aplikasi web yang telah dibuat, sehingga aplikasi web dapat menampilkan diagram, tabel, dan juga hasil analisis.

#### DAFTAR PUSTAKA

- [1] A. Febriani, R. Melyanti, and R. W. Syahputra, “Sistem Informasi Donor Darah Berbasis Android Pada Unit Transfusi Darah Palang Merah Indonesia (UTD PMI) Kota Pekanbaru,” *J. Ilmu Komput.*, vol. 9, no. 1, pp. 11–19, 2020, doi: 10.33060/jik/2020/vol9.iss1.146..
- [2] A. Profita, D. S. Utomo, and F. Fachriansyah, “Optimasi Manajemen Persediaan Darah Menggunakan Simulasi Monte Carlo,” *J. Ind. Eng. Manag.*, vol. 2, no. 1, p. 16, 2017, doi: 10.33536/jiem.v2i1.101.
- [3] L. Raufun and W. O. D. Angraini, “Rancangan Bangun Aplikasi Sistem Informasi Ketersediaan Darah Pada Palang Merah Indonesia Kabupaten Buton Berbasis Android,” *J. Inform.*, vol. 8, no. 1, 2019..
- [4] M. A. Maulana, A. Setyanto, and M. P. Kurniawan, “Analisis Sentimen Media Sosial Universitas Amikom Yogyakarta Sebagai Sarana Penyebaran Informasi Menggunakan Algoritma Klasifikasi SVM,” *Semin. Nas. Teknol. Inf. dan Multimed. 2018 Univ. AMIKOM Yogyakarta*, 10 Februari 2018, pp. 7–12, 2018.
- [5] D. T. Lukmana, S. Subanti, and Y. Susanti, “Analisis Sentimen Terhadap Calon Presiden 2019 Dengan Support Vector Machine Di Twitter,” *Semin. Nas. Penelit. Pendidik. Mat. 2019 UMT*, no. 2002, pp. 154–160, 2019.
- [6] N. T. Hartanti, “Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 82–89, 2020, doi: 10.25077/teknosi.v6i2.2020.82-89.
- [7] W. I. Loka and F. Natalia, “Perancangan dan Pembuatan Visualisasi Data Dana Penelitian Internal dan Hibah Dikti LPPM Universitas Multimedia Nusantara,” *Ultim. InfoSys*, vol. X, no. 1, pp. 61–68, 2019.