

# KLASIFIKASI SENTIMEN MASYARAKAT TERHADAP PEMERINTAH TERKAIT PENERAPAN KEBIJAKAN NEW NORMAL MENGGUNAKAN METODE *K-NEAREST NEIGHBOR*

## TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Teknik  
pada Program Studi Teknik Informatika Fakultas Sains dan Teknologi



UIN SUSKA RIAU

Oleh:

**FUJI ASTUTI**

**11750124890**

UIN SUSKA RIAU

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU**

**PEKANBARU**

**2022**

## LEMBAR PERSETUJUAN

### KLASIFIKASI SENTIMEN MASYARAKAT TERHADAP PEMERINTAH TERKAIT PENERAPAN KEBIJAKAN NEW NORMAL MENGUNAKAN METODE K-NEAREST NEIGHBOR

TUGAS AKHIR

Oleh

FUJI ASTUTI

NIM. 11750124890

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir

Di Pekanbaru, pada tanggal 14 Juli 2022

Pembimbing I,



Reski Mai Candra, S.T., M.Sc

NIP. 19860505 201503 1 006

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PENGESAHAN

### KLASIFIKASI SENTIMEN MASYARAKAT TERHADAP PEMERINTAH TERKAIT PENERAPAN KEBIJAKAN NEW NORMAL MENGGUNAKAN METODE K-NEAREST NEIGHBOR


Oleh

**FUJI ASTUTI**  
NIM. 11750124890

Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik  
pada Universitas Islam Negeri Sultan Syarif Kasim Riau  
di Pekanbaru, pada tanggal 14 Juli 2022

Pekanbaru, 14 Juli 2022

Mengesahkan,  
Ketua Jurusan,


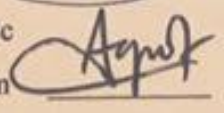


  
Awan Iskandar, M.T  
NIP. 19821216 201503 1 003



Dr. Hartono, M.Pd  
NIP. 19640301 199203 1 003

#### DEWAN PENGUJI

Ketua : Febi Yanto, S.T., M.Kom  
Pembimbing I : Reski Mai Candra, S.T., M.Sc  
Penguji I : Surya Agustian, S.T., M.Kom  
Penguji II : Siti Ramadhani, S.Pd., M.Kom

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

# SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : FUJI ASTUTI  
NIM : 11750124890  
Tempat/Tgl. Lahir : Sungai Guntung, 9 Juli 1999  
Fakultas/Pascasarjana : Fakultas Sains dan Teknologi  
Judul Skripsi :

**KLASIFIKASI SENTIMEN MASYARAKAT TERHADAP PEMERINTAH TERKAIT PENERAPAN  
KEBIJAKAN NEW NORMAL MENGGUNAKAN METODE K-NEAREST NEIGHBOR**

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Skripsi lainnya dengan judul sebagaimana tersebut diatas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Skripsi lainnya saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Skripsi saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 20 Juli 2022  
ng Membuat Pernyataan



FUJI ASTUTI  
NIM. 11750124890



## Klasifikasi Sentimen Masyarakat Terhadap Pemerintah Terkait Penerapan Kebijakan New Normal Menggunakan Metode *K-Nearest Neighbor*

Fuji Astuti<sup>1\*</sup>, Reski Mai Candra<sup>2</sup> dan Siti Ramadhani<sup>3</sup>

<sup>1-3</sup>Universitas Islam Negeri Sultan Syarif Kasim Riau, Jalan HR. Soebrantas Pekanbaru, 28293  
E-mail: [11750124890@students.uin-suska.ac.id](mailto:11750124890@students.uin-suska.ac.id)<sup>1</sup>, [reski.candra@uin-suska.ac.id](mailto:reski.candra@uin-suska.ac.id)<sup>2</sup>, [siti.ramadhani@uin-suska.ac.id](mailto:siti.ramadhani@uin-suska.ac.id)<sup>3</sup>

**Abstrak** – Instagram memiliki berbagai macam pengguna tanpa terkecuali pemerintah, akun instagram pemerintah sendiri memiliki fungsi sebagai salah satu media informasi bagi masyarakat untuk mengetahui berita terbaru. Pemerintah telah melakukan berbagai upaya untuk mengurangi dan memutus rantai penyebaran virus covid-19, salah satu langkah yang diambil pemerintah dalam menghadapi covid-19 dengan menerapkan kebijakan New Normal. Tujuan penulis dari penelitian ini adalah untuk menghitung akurasi metode *K-Nearest Neighbor* pada klasifikasi sentimen dari opini masyarakat terkait penerapan kebijakan new normal. Penelitian ini dilakukan melalui beberapa tahapan yaitu pengumpulan data dari akun yang dikelola pemerintah. Setelah itu data tersebut dilakukan proses preprocessing. Lalu membuat model dengan menggunakan word embedding FastText. Selanjutnya menggunakan feature engineering dalam melakukan improve untuk menghasilkan fitur terbaik. Dataset dalam penelitian ini berjumlah 4717 data yang terbagi dari 1456 data kelas pro, 1662 data kelas netral, 1599 data kelas kontra. Pengujian yang dilakukan menggunakan confusion matrix dengan cara menguji data latih dan data uji sebanyak 3 kali dengan perbandingan data 90:10 yaitu nilai akurasi sebesar 60,169% dan pada k-40 dengan perbandingan 80:20 yaitu nilai akurasi 60,16%, sedangkan akurasi terendah terdapat pada perbandingan data 70:30 yaitu 58,12%.

**Kata kunci:** *New Normal, Instagram, Klasifikasi, Sentimen, K-Nearest Neighbor*

**Abstract** - Instagram has various kinds of users without the exception of the government, the government's own Instagram account has a function as a medium of information for the public to find out the latest news. The government has made various efforts to reduce and break the chain of the spread of the COVID-19 virus, one of the steps taken by the government in dealing with COVID-19 by implementing the New Normal policy. The author's aim of this study is to calculate the accuracy of the *K-Nearest Neighbor* method in classifying sentiments from public opinion regarding the implementation of the new normal policy. This research was conducted through several stages, namely collecting data from accounts managed by the government. After that, the data is preprocessed. Then create a model using word embedding FastText. Then use feature engineering to improve to produce the best features. The dataset in this study amounted to 4717 data which was divided into 1456 pro class data, 1662 neutral class data, 1599 contra class data. The test was carried out using a confusion matrix by testing training data and test data 3 times with a data comparison of 90:10, namely the accuracy value of 60.169% and at k-40 with a ratio of 80:20, namely the accuracy value of 60.16%, while the lowest accuracy there is a comparison of 70:30 data that is 58.12%.

**Keywords:** *New Normal, Instagram, Classification, Sentiment, K-Nearest Neighbor*

### 1. Pendahuluan

Kemajuan teknologi dibidang komunikasi salah satunya ditandai dengan munculnya berbagai macam media sosial diantaranya adalah instagram. Instagram memiliki berbagai macam pengguna tanpa terkecuali pemerintah, akun instagram pemerintah sendiri memiliki fungsi sebagai salah satu media informasi bagi masyarakat untuk mengetahui berita terbaru. Manfaat yang bisa didapatkan dari penggunaan akun instagram pemerintah adalah dapat menganalisa sentimen masyarakat terhadap kebijakan-kebijakan yang dikeluarkan oleh pemerintah dari masyarakat pada komentar postingan akun pemerintah tersebut. Salah satunya adalah kebijakan yang berhubungan dengan pandemi yang sedang terjadi saat ini.

Presiden Joko Widodo mengumumkan secara resmi kasus pertama COVID-19 di Indonesia di Istana Negara tanggal 2 Maret 2020. COVID-19 ini terus menyebar diseluruh Indonesia. Pemerintah melakukan berbagai cara untuk mengurangi dan memutus rantai penyebaran virus covid-19, salah satu upaya yang dilakukan pemerintah dalam menghadapi covid-19 yaitu dengan menerapkan kebijakan New Normal.

Langkah yang dibuat oleh pemerintah tidak semua mendapat respon yang baik dari masyarakat mengingat adanya perbedaan pemikiran dan kondisi dari setiap orang. Respon yang muncul dapat berupa pro, netral maupun kontra, respon-respon ini tidak jarang dituangkan oleh masyarakat melalui sosial media melalui postingan akun pemerintah mengenai informasi kebijakan yang diterapkan salah satunya pada komentar

postingan dimedia sosial instagram.

Dalam penelitian ini, penulis terinspirasi dari penelitian yang berkaitan dengan latar belakang permasalahan. Seperti penelitian yang berjudul Analisis Sentimen Twitter Kebiasaan New Normal dengan hasil penelitian dari analisis sentiment twitter didapatkan 57% nilai positif, 35% netral, dan negatif 8%. Namun pada penelitian ini pengambilan data berupa tweet dalam bahasa inggris sehingga untuk kedepannya dapat dilakukan analisis dalam berbagai bahasa (Dwitiyanti & Selvia, 2021)[1].

Terdapat beberapa metode dalam melakukan klasifikasi teks, salah satunya metode K-Nearest Neighbor. Beberapa penelitian terkait sebelumnya yaitu tentang klasifikasi komentar spam pada instagram berbahasa Indonesia menggunakan K-NN. Hasil akurasi yang didapatkan sebesar 87,07% (Chrisanto & Lukito, 2017)[2] Selanjutnya penelitian tentang Perbandingan Kinerja Algoritma Support Vector Machine Dan K-Nearest Neighbor Terhadap Analisis Sentimen Kebijakan New Normal dengan hasil pengujian menunjukkan bahwa algoritma K-NN lebih baik dibandingkan SVM dalam penelitian ini dengan nilai akurasi sebesar 72,96% (Muhidin & Wibowo, 2020)[3].

Berdasarkan uraian masalah dan latar belakang diatas yang fokus permasalahan adalah “Bagaimana mengklasifikasikan sentimen masyarakat terhadap pemerintah terkait penerapan kebijakan new normal yang telah diterapkan di Indonesia menggunakan metode K-Nearest Neighbor membaginya berdasarkan kelompok sentimen pro, kontra dan netral diinstagram”. Adapun tujuan dari penelitian ini yaitu untuk mengetahui nilai akurasi metode K-Nearest Neighbor pada klasifikasikan sentimen dari opini masyarakat terkait penerapan kebijakan new normal.

## 2. Tinjauan Pustaka

### 2.1 Instagram

Instagram merupakan aplikasi untuk dapat berkomunikasi dengan orang yang berbeda dengan berbagi foto atau video di antara mereka, Instagram juga memiliki fitur lain seperti DM (pesan langsung), komentar, dan cinta. Instagram juga dapat mengambil foto dan video, menerapkan filter digital, dan membagikannya di berbagai jejaring sosial. Instagram memiliki beberapa fitur yang dapat digunakan yaitu kamera, editor, tag dan hashtag, caption, story, discovery, integrasi dengan media sosial lain dan komentar.

### 2.2 New Normal

Menurut Roger McNamee dalam jurnal [1][4], *new normal* merupakan masa di mana mengikuti aturan baru dalam jangka waktu yang cukup panjang. Di masa pandemi ini, *New normal* merupakan kebijakan yang ditetapkan oleh pemerintah setelah adanya kebijakan PSBB (Pembatasan Sosial Berskala Besar). Istilah new normal atau tatanan baru adalah kehidupan yang berlangsung dengan protokol kesehatan yang ketat untuk menghindari penularan Covid-19. Protokol kesehatan meliputi mencuci tangan setelah beraktivitas di luar ruangan, menjaga jarak, dan selalu menggunakan masker saat berada di luar ruangan.

### 2.3 Analisa Sentimen

Menurut Nasukawa & Yi dalam jurnal [2] [5] *Sentiment Analysis* dalam bahasa Indonesia adalah metode yang digunakan untuk mengidentifikasi emosi yang diungkapkan secara tertulis dan dapat diklasifikasikan sebagai sentimen positif atau sentimen negatif. Sedangkan pendapat lain menyebutkan analisis sentimen merupakan proses yang digunakan untuk menentukan opini, sikap, emosi, dan keadaan yang dicerminkan melalui teks.

### 2.4 Text Preprocessing

Menurut Feldman yang dikutip dari [3][6], *Text preprocessing* merupakan tahap awal pengolahan kata menjadi data untuk diolah. Berikut ini menjelaskan prosedur yang menjelaskan langkah-langkah preprocessing teks.

#### 1. Case Folding

Tahapan *case folding* merupakan mengubah semua huruf dalam dokumen menjadi huruf kecil. Karakter non-abad diabaikan dan diperlakukan sebagai pembatas.

#### 2. Tokenizing

Pada tahap *tokenizing* adalah tahapan pemotongan string input berdasarkan pada tiap kata yang menyusunnya.

#### 3. Tanda Baca (*Punctuation*)

Langkah ini merupakan proses pembersihan dokumen dari beberapa kata yang tidak perlu untuk mengurangi noise.. Kata yang dihapus dari data komentar adalah simbol, angka, URL link, hashtag (#), dan mention (@username).

#### 4. Stopword Removal

Tahap *stopword removal* merupakan kata yang sering muncul dalam jumlah banyak dan tidak bermakna. Beberapa contoh kata yang tergolong dalam kata langkah ini antara lain “dengan”, “ke”, “ke”, “dan”,



2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

petunjuk waktu, dan kata tanya.

**2.5 Word Embedding**

*Word Embedding* merupakan representasi pada kata yang menciptakan vektor berdasarkan konteksnya. Tahapan ini digunakan untuk mendapatkan informasi semantik dan sintaksis dari kata, dan digunakan untuk mendapatkan kesamaan. Metode *Embedding* menurut Huang [4][7] metode ini menggunakan konteks sekelompok kata dalam suatu grup, kemudian mendapatkan vektor prototipe yang berbeda untuk setiap cluster dan mengeksekusinya berdasarkan model Neuronal..[5][8].

**2.6 Feature Engineering**

*Feature engineering* atau rekayasa fitur merupakan untuk membangun fitur yang akan meningkatkan kinerja mesin. Metode ini didasarkan pada penerapan fungsi pada transformasi seperti aritmatika dan agregasi untuk mencapai fungsionalitas baru, untuk mengubah hubungan non-linier menjadi hubungan linier yang lebih mudah dipelajari untuk mesin (Nargesian et al., 2017) [9].

**2.7 Hyperparameter Tuning**

Tahapan *hyperparameter tuning* merupakan tahapan untuk menentukan nilai yang optimal untuk data input dan model prediksi untuk mendapatkan nilai yang baik. Setiap data input dibutuhkan hyperparameter yang sesuai dengan karakteristik dan dapat memberikan hasil prediksi yang akurat.

**2.8 K-Nearest Neighbor**

*K-Nearest Neighbor* merupakan metode pengenalan pola umum dan digunakan untuk proses klasifikasi kelompok dikarenakan metode ini menggunakan teknik yang sederhana[6]. Metode ini mengklasifikasikan data yang tidak berlabel menggunakan label mayoritas dari sampel tetangga terdekat pada data latih (Sari et al., 2018) [10].

**2.9 Confusion matrix**

Untuk mendapatkan hasil berdasarkan data yang digunakan, penelitian ini menggunakan matriks konfusi untuk mengetahui apakah yang dilakukan sudah tepat dan dapat menentukan tingkat akurasi tergantung pada tujuannya (Salam et al., 2018) [11].

**3. Metode Penelitian**

**3.1 Data Preparation**

Pada penelitian ini data yang digunakan adalah data berupa teks komentar instagram beberapa akun resmi milik pemerintah yang akan digunakan sebagai data latih dan data uji yang berkaitan tentang *new normal* pada sosial media instagram. Berikut akun instagram yang akan digunakan sebagai data dalam penelitian ini: @jokowi, @satgasperubahanperilaku, @kemenkominfo, @kemenkes\_ri, @bnpb\_indonesia, @lawancovid19\_id. Data komentar yang diperoleh berdasarkan komentar pada akun instagram menggunakan website *phantombuster*, kemudian data akan di *export* kedalam *microsoft excel*. Jumlah *dataset* dalam penelitian ini berjumlah 4717 data yang terbagi dari 1456 data kelas pro, 1662 data kelas netral, 1599 data kelas kontra.

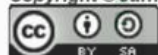
**3.2 Preprocessing**

*Text pre-processing* merupakan tahapan untuk memudahkan proses dalam pengolahan untuk tahap selanjutnya. Berikut tahapan-tahapan pada proses *text pre-processing*:

**I. Tokenizing**

Tabel 1. *Tokenizing*

No	Komentar	Hasil <i>Tokenizing</i>
1	Siap pak di laksanakan	Siap pak di laksanakan
2	Semua arahan bapa kami lakukan	Semua arahan bapa kami lakukan
3	Siap pak, saya mematuhi perintah bapak kok.	Siap pak, saya mematuhi



**2. Case folding**

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	perintah
	bapak
	kok.

Tabel 2. *Case Folding*

No	Komentar	Hasil <i>Case Folding</i>
1	Siap pak di laksanakan	siap
		pak
		di
		laksanakan
2	Semua arahan bapa kami lakukan	semua
		arahan
		bapa
		kami lakukan
3	Siap pak, saya mematuhi perintah bapak kok.	siap
		pak,
		saya
		mematuhi
		perintah
		bapak kok.

**3. Stopword**

Tabel 3. *Stopword*

No	Komen tar	Sebelum <i>Filtering</i>	Setelah <i>Filtering</i>
1	Siap pak di laksana kan	siap	siap
		pak	pak
		di	
		laksanakan	laksanakan
2	Semua arahan bapa kami lakuka n	semua	semua
		arahan	arahan
		bapak	bapak
		kami	kami
		lakukan	lakukan
3	Siap pak, saya mematuhi perinta h bapak kok.	siap	siap
		pak,	pak,
		saya	saya
		mematuhi	mematuhi
		perintah	perintah
		bapak kok.	bapak





Hak cipta milik UIN Suska Riau  
 2. Diarangi mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.  
 2. Diarangi mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.  
 2. Diarangi mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4. Punctuation

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

Tabel 4. Punctuation

No	Komentar	Sebelum Punctuation	Setelah Punctuation
1	Siap pak di laksanakan	siap	siap
		pak	pak
		laksanaka n	laksanakan
2	Semua arahan bapa kami lakukan	semua	semua
		arahan	arahan
		bapak	bapak
		kami	kami
3	Siap pak, saya mematuhi perintah bapak kok.	siap	siap
		pak,	pak
		saya	saya
		mematuhi	mematuhi
		perintah	perintah
	bapak	bapak	

3.3 Training Model

Pelatihan model merupakan langkah dalam proses pembuatan model menggunakan penyisipan kata *FastText*. Langkah selanjutnya adalah membuat model dari data dan mengimport data library yang akan digunakan. Jika impor data berhasil, seluruh komentar data menjadi satu kata untuk menjalankan proses pembuatan model menggunakan *FastText* dan simpan ke folder yang disiapkan

3.4 Feature Engineering

Pada titik ini, mari tambahkan fungsi *scaling* sebagai *MinMax* untuk mendapatkan hasil yang lebih baik dari kombinasi langkah-langkah *preprocessing* yang telah dilakukan. Hal ini sering disebut *improve data*, *scalling* dilakukan agar antara satu data dengan data lainnya tidak ada nilai yang ekstrim dan tidak ada anomali, kemudian dilakukan langkah penyesuaian untuk mendapatkan nilai k terbaik dari setiap nilai yang dicoba.

3.5 Klasifikasi K-Nearest Neighbor

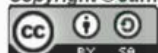
*K-Nearest Neighbor* merupakan metode klasifikasi objek untuk data latih dengan memilih kelas fitur berdasarkan jarak terdekat dengan data latih. Berdasarkan atribut dan pola dari data latih, algoritma ini dapat mengklasifikasikan objek baru. KNN merupakan metode klasifikasi dengan menghitung nilai K (Neighbor). K adalah titik tetangga objek, nilai K akan dibandingkan untuk menentukan klasifikasi objek. Model K-Nearest Neighbors pada penelitian ini diimplementasikan dalam bahasa pemrograman Python menggunakan fungsi *K-NeighborsClassifier* dari library *Sklearn*. Langkah-langkah metode K-Nearest Neighbor adalah:

- Tentukan k (jumlah tetangga paling dekat).
- Hitung jarak data *testing* dengan semua data *training*. Ukuran jarak yang digunakan adalah *Euclidean distance*:

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \quad (1) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

- Urutkan jarak tersebut dari yang terdekat atau jarak terkecil.
- Periksa kelas k tetangga terdekat
- Kelas data baru = Kelas mayoritas k tetangga terdekatnya
- Lalu didapatkan hasil pengklasifikasian.

Sebelumnya hyperparameter yang sesuai *K-Nearest Neighbor* telah dicari menggunakan *Randomized Search* untuk mencari parameter terbaik.

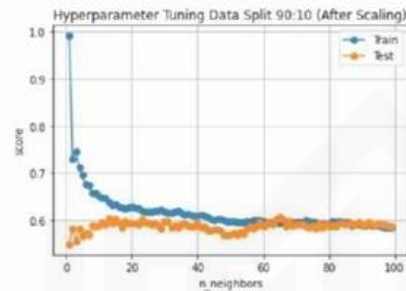


#### 4. Hasil dan Pembahasan

Klasifikasi terdiri dari 3 Kelas yaitu Pro, Netral dan Kontra. Jumlah *dataset* dalam penelitian ini berjumlah 4717 data yang terbagi dari 1456 data kelas pro, 1662 data kelas netral, 1599 data kelas kontra. Pengujian *confusion matrix* dilakukan dengan cara menguji data latih menjadi data uji sebanyak 3 kali pengujian dengan mekanisme pengujian data latih dan data uji menggunakan perbandingan 70:30, 80:20, 90:10.

##### 1. Pengujian akurasi dengan pembagian data 90:10

Pengujian pertama dilakukan perbandingan 90:10 dengan k=31 didapatkan hasil akurasi sebesar 60,169%

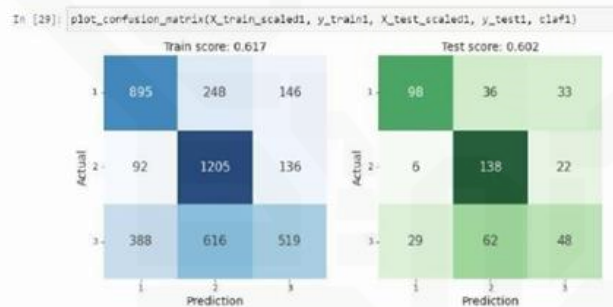


```
In [28]: clf1 = KNeighborsClassifier(n_neighbors=31, metric='minkowski')
clf1 = clf1.fit(X_train_scaled1, y_train1)
clf1_pred1 = clf1.predict(X_test_scaled1)
acc1 = (accuracy_score(y_test1, clf1_pred1)*100)
print("Akurasi untuk data split 90:10 = ", acc1)
```

Akurasi untuk data split 90:10 = 60.16949152542372

Gambar 1. Pengujian Akurasi Dengan Pembagian Data 90:10

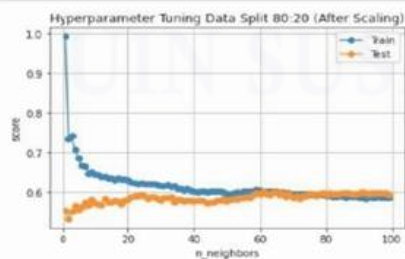
Lalu dilakukan pengujian matrix dengan pembagian data 90:10



Gambar 2. Pengujian Akurasi Dengan Pembagian Data 90:10

##### 2. Pengujian akurasi dengan pembagian data 80:20

Pengujian kedua dilakukan perbandingan 80:20 dengan k=40 didapatkan hasil akurasi sebesar 60,16%

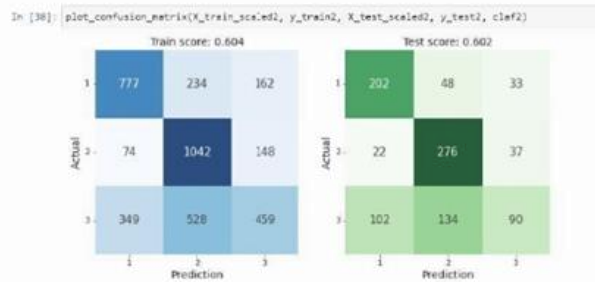


```
[37]: clf2 = KNeighborsClassifier(n_neighbors=40, metric='minkowski')
clf2 = clf2.fit(X_train_scaled2, y_train2)
clf2_pred2 = clf2.predict(X_test_scaled2)
acc2 = (accuracy_score(y_test2, clf2_pred2)*100)
print("Akurasi untuk data split 80:20 = ", acc2)
```

Akurasi untuk data split 80:20 = 60.16949152542372

Gambar 3. Pengujian Akurasi Dengan Pembagian Data 80:20

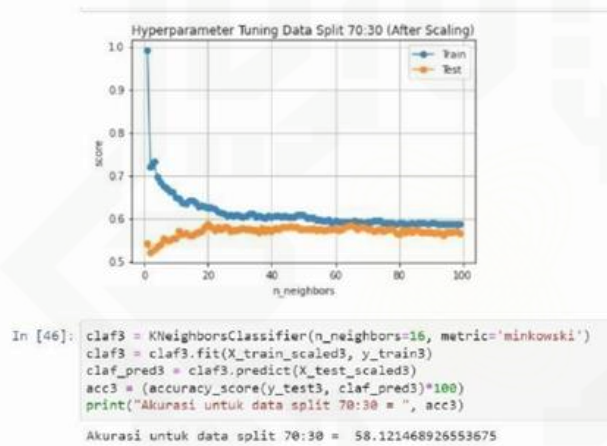
Lalu dilakukan pengujian matrix dengan pembagian data 80:20



Gambar 4. Pengujian Akurasi Dengan Pembagian Data 80:20

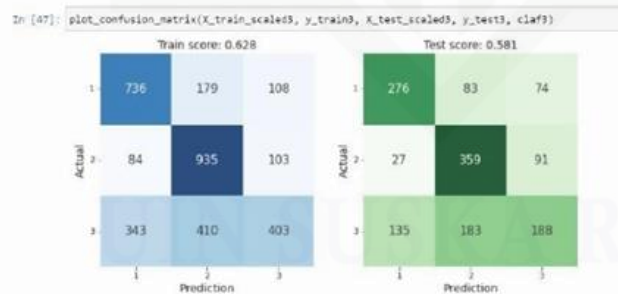
3. Pengujian akurasi dengan pembagian data 70:30

Pengujian ketiga dilakukan perbandingan 70:30 dengan k=16 didapatkan hasil akurasi sebesar 58,12%



Gambar 5. Pengujian Akurasi Dengan Pembagian Data 70:30

Lalu dilakukan pengujian matrix dengan pembagian data 70:30



Gambar 5. Pengujian akurasi dengan pembagian data 70:30

## 5. Kesimpulan

Berdasarkan tahapan-tahapan metodologi penelitian yang telah dilakukan, maka didapatkan kesimpulan sebagai berikut:

- Metode *K-Nearest Neighbor* dapat diterapkan untuk proses klasifikasi pro, kontra dan netral pada komentar instagram.
- Klasifikasi pro, kontra dan netral pada komentar instagram menggunakan *K-Nearest Neighbor* telah berhasil melakukan klasifikasi, dengan akurasi terbaik terdapat pada k-31 dengan perbandingan data 90:10

yaitu sebesar 60,169% dan pada k-40 dengan perbandingan 80:20 yaitu 60,16%, sedangkan akurasi terendah terdapat pada perbandingan data 70:30 yaitu 58,12%.

### Daftar Pustaka

- [1] A. I. Nugraha, H. Haniyah, R. A. Kusumah, S. Novitasari, L. P. Maharani, and D. Tresnawati, "Peningkatan Pengetahuan Tentang Covid-19 , Literasi Digital Dan New Normal Masyarakat," pp. 126–134, 2020.
- [2] M. S. Hadna, P. I. Santosa, and W. W. Winarno, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 57–64, 2016.
- [3] A. Rahmah, H. B. Santoso, and Z. A. Hasibuan, "Characteristics analysis for technology enhanced learning maturity: A qualitative approach," *ICCE 2019 - 27th Int. Conf. Comput. Educ. Proc.*, vol. 1, no. 1, pp. 487–489, 2019.
- [4] E. H. Huang, R. Socher, C. D. Manning, and A. Y. Ng, "and Multiple Word Prototypes," no. July, pp. 873–882, 2012.
- [5] Y. Liu, Z. Liu, T. Chua, and M. Sun, "Liuyang\_Chua," *Proc. 29th AAAI Conf. Artif. Intell.*, vol. 2, no. C, pp. 2418–2424, 2015.
- [6] R. M. Juwita, E. Haerani, S. K. Gusti, and S. Ramadhani, "Klasifikasi Berita Menggunakan Metode K-Nearest Neighbor," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 2, pp. 259–268, 2022.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.