



## ANALISIS SENTIMEN DATA *TWITTER* MENGGUNAKAN METODE *K-MEANS CLUSTERING* PADA STUDI KASUS PEMINDAHAN IBU KOTA NUSANTARA (IKN)

Ritma Melati<sup>1</sup>, M. Reza Redo Islami<sup>2</sup>

<sup>1</sup>Prodi Sistem Informasi, STMIK Dharma Wacana, Kota Metro, Lampung

<sup>2</sup>Prodi Teknik Informatika, STMIK Dharma Wacana, Kota Metro, Lampung

<sup>1,2</sup>Jalan Kenanga No.3 Mulyojati Metro Barat, Kota Metro, Lampung, Indonesia

E-mail: [ritmamelati1122@gmail.com](mailto:ritmamelati1122@gmail.com)<sup>1</sup>, [reza.redo@hotmail.com](mailto:reza.redo@hotmail.com)<sup>2</sup>

### Article history:

Received: March 23, 2024

Revised: April 1, 2024

Accepted: June 7, 2024

Corresponding authors

[ritmamelati1122@gmail.com](mailto:ritmamelati1122@gmail.com)

### Keywords:

Twitter;

Sentiment Analysis;

K-Means Clustering;

Clustering;

Centroid.

### Abstract

This research aims to find out the public's response to the relocation of the Indonesian capital or IKN and to find out the emotions contained in tweets. The background in this research is the large number of users of certain social media, especially Twitter, to express opinions about the relocation of the archipelago's capital city or IKN. This research uses the k-means clustering method with text input. This research includes data collection, preprocessing, clustering. Preprocessing includes stages of cleaning, case flooding, normalization, data transformation, stemming, data reduction, data discretization, clustering using the k-means method by determining several clusters. The experiment was carried out with 2000 tweet data which was grouped into 3 clusters, namely C1, C2, C3. The highest result from sentiment analysis was 79.31% with negative responses and the highest accuracy results obtained with the k-means clustering method.



This is an open access article under the CC-BY-SA license.

### I. PENDAHULUAN

Di era saat ini anak muda, orang dewasa atau kalangan masyarakat di manapun berada tidak dapat menghindari salah satu media sosial teratas yaitu media sosial twitter yang berpengaruh terutama dalam hal pencarian informasi dan berita-berita terkini. Dan media sosial twitter sendiri menyediakan banyak berbagai fasilitas-fasilitas yang sangat menarik yang dapat di manfaatkan untuk berbagai peluang, termasuk juga di dalam mencari informasi dan berita-berita terkini yang saat ini sedang trending. Dengan adanya fitur-fitur yang canggih yang tersedia di media sosial twitter ini, mencari berbagai informasi dan berita terkini tidak lagi menjadi suatu tantangan bagi siapapun saja yang menggunakannya.

Penggunaan media sosial twitter sebagai alat untuk mencari berbagai informasi dan berita sehingga telah membuat mediasosial ini menjadi

sebuah hal yang penting dan terkait erat dengan kehidupan sehari-hari. Media sosial atau sering juga disebut sosmed adalah pelantara digital yang memfasilitasi terhadap penggunaannya untuk saling berinteraksi satu dengan yang lain atau juga bisa saling membagikan sebuah konten berupa tulisan, gambar, video, dan berita terkini, dan juga media sosial merupakan pelantara digital yang banyak menyediakan berbagai fasilitas-fasilitas lakukan suatu aktivitas sosial bagi setiap penggunanya, media sosial terdapat banyak jenis-jenis media sosial yang banyak disukai di kalangan anak muda, orang dewasa dan masyarakat pada saat ini termasuk salah satunya adalah media sosial twitter.

Twitter sendiri merupakan layanan jejaring sosial dan mikroblogging yang memberikan kemudahan-kemudahan bagi pengguna media sosial twitter untuk mengetik atau menuliskan sesuatu dan menyatakan

sesuatu tindakan atau pendapat melalui akun media sosial yang di miliki. Pengguna media sosial twitter dapat berbagi aktivitas-aktivitas sehari-hari yang di lakukan contohnya seperti mengunggah foto atau menyampaikan pandangan mengenai topik tertentu yang sedang trending pada saat ini kepada pengguna twitter lainnya. Topik topik dan informasi yang populer di media sosial twitter sering kali memicu banyaknya komentar-komentar dan diskusi hangat di kalangan masyarakat selain itu juga, opini-opini dan sentimen yang tertangkap dalam media sosial twitter juga dapat menjadi subjek atau juga bisa menjadi sebuah studi kasus dalam penelitian karya ilmiah atau skripsi.

Analisis sentimen sendiri adalah sebuah proses untuk memahami, mengekstrak dan memproses informasi-informasi perasaan seseorang yang terkait tentang pendapat tersebut, berlandaskan tentang komputasi tentang pendapat, sentimen, opini, maupun perasaan seseorang terhadap suatu topik. Di dalam penelitian ini, untuk mengumpulkan data-data di lakukan dengan menggunakan media sosial twitter untuk menyimpan pendapat publik. Gabungan pandangan atau pendapat serta komentar-komentar dan opini tersebut yang nantinya dapat di gunakan sebagai resource di dalam proses analisis sentimen. Pada saat ini terdapat banyak pengembangan teknologi sentimen yang di gunakan dalam penelitian untuk penerapan teknologi ini bertujuan untuk menganalisis topik dan pendapat serta komentar-komentar publik, yang khususnya dalam mengevaluasi sejauh mana pengguna atau pelanggan merasa puas terhadap topik tertentu yang sedang trending di masa ini. Serta analisis sentimen sendiri adalah seatu metode yang otomatis di gunakan untuk memahami, mengekstrak, dan memproses data-data teks yang di dapatkan dengan bertujuan untuk mendapatkan wawasan dan informasi-informasi yang berharga saat ini. Teknik inilah yang sering banyak di gunakan untuk mengevaluasi dan menganalisis berbagai puas pengguna atau pelanggan terhadap suatu topik dan kebijakan tertentu.

Dalam studi kasus ini mengambil suatu berita yang sedang banyak di bahas dalam masyarakat saat ini yaitu tentang pemindahan ibu kota nusantara atau IKN, IKN sendiri yaitu ibu kota nusantara dan pemindahan ibu kota saat ini bukanlah suatu fenomena yang baru dan banyak di perbincangkan tentang pengalihan pusat pemerintahan indonesia yaitu dari ibu kota jakarta di pindahkan ke wilayah kalimantan dan telah

berlangsung sejak priode pemimpin dari presiden Soekarno sampai dengan Jokowi. Meski dengan demikian, sehingga kini rencana tersebut belum terealisasi pada tanggal 29 april 2019, di dalam sebuah rapat , presiden indonesia yang telah di ambil keputusan mengenai pemindahan ibu kota yang akan di pindahkan dari pulau jawa oleh presiden Jokowi. dengan melalui banyak penelitian-penelitian dan audiensi publik, di kalimantan timur yang saat ini terpilih sebagai ibu kota yang baru di indonesia . Keputusan untuk melanjutkan proyek mengenai pemindahan ibu kota dengan senilai Rp.466,04 yang harus di jalankan, dan keputusan-keputusan ini yang memberikan kepastian politik pemerintah berencana untuk memindahkan ibu kota negara ke kabupaten penajam paser utara dan kutai kartanegara yang berada di provinsi kalimantan timur. Pemindahan ibu kota negara bertujuan untuk mempercepat pertumbuhan ekonomi dan mengurangi ketakseimbangan di luar pulau jawa, yang terutama di bagian indonesia timur. Parapemerintah percaya bahwa pemindahan ibu kota negara ke pasar modal nasional yang nantinya akan membawa banyak manfaat ekonomi. Dengan adanya analisis sentimen ini yang bertujuan untuk menentukan isi data set yang berupa teks yang bersifat positif, negatif, dan netral. dalam kemajuan teknologi dan informasi yang meningkat memberikan dampak pada pertukaran informasi-informasi serta komunikasi yang semakin baik dan mudah. Serta media sosial merupakan salah satu sarana yang di gunakan oleh netizen untuk mengakses berbagi dan berdiskusi mengenai pemindahan ibu kota nusantara atau IKN.

Media sosial twitter sendiri sebagai salah satu media sosial yang merupakan platform yang secara nyata sering di pilih untuk membicarakan atau mengkomunikasikan hal tersebut. Dengan melalui analisis sentimen ini kita dapat memahami bagaimana orang-orang menggambarkan serta mengungkapkan perasaan dan persepsi mereka terhadap pemindahan ibu kota nusantara atau IKN baik dengan secara positif, negatif, dan netral. Analisis sentimen tersebut sangat penting untuk melihat sejauh mana media sosial twitter yang di gunakan saat ini sebagai salah satu instrumen diseminasi informasi mengenai pemindahan ibu kota nusantara atau IKN. Tujuan dari penelitian yang di lakukan saat ini yaitu melakukan analisis sentimen data twitter dengan metode k-means clustering.

## **II. METODE PENELITIAN**

Metode penelitian ini di uraikan tujuannya agar memberikan gambaran telah tercapainya tentang langkah-langkah yang akan di ambil untuk menjawab pertanyaan dalam penelitian. Langkah-langkah ini

melibatkan pengumpulan data atau crawling data, pemrosesan informasi, dan juga analisis data. Pada penelitian ini untuk melakukan tahapan clustering data yaitu dengan menggunakan metode k-means clustering, guna untuk mencari nilai optimal pada silhouette scores dan juga menentukan jumlah kluster yang akan dilakukan dalam analisis di tahap selanjutnya.

Metode k-Means clustering juga memberikan gambaran lebih mendalam tentang seberapa tanggapan positif, negatif, dan netral dalam suatu teks. Pemrosesan tweet menggunakan metode yang akan menghasilkan tanggapan atau skor negatif, skor netral, skor positif. Pemberian label sentimen positif dilakukan pada tweet yang memiliki skor gabungan lebih dari nol. Sementara label sentimen negatif diberikan pada tweet dengan skor gabungan kurang dari nol. Dan label sentimen netral dilakukan pada tweet yang memiliki skor gabungan sama dengan nol.

Data yang di gunakan berasal dari salah satu media sosial twitter, pengumpulan data di lakukan menggunakan web scraping melalui aplikasi google collab dengan menggunakan bahasa pemrograman python. Dan kata kunci yang di gunakan untuk mencari data tersebut yaitu (IBU KOTA NEGARA / IKN) data yang di dapat sebanyak 2000 data tweet dengan menggunakan bahasa indonesia Pengambilan Data yang dikumpulkan berupa data teks yang diambil dengan teknik crawling data pada tanggal 29 november 2023 pengumpulan data di lakukan kurang lebih selama satu tahun twitt yang di ambil yaitu sebanyak 2000 data twitt yang berisikan pendapat, komentar, opini mengenai pemindahan ibu kotanusantara (IKN).

## **2.2. Prapemrosesan Data**

Menurut han, J., Kamber, M., & Pei, J. (2011) preprocessing data adalah tahapan awal dalam analisis data yang melibatkan serangkaian langkah-langkah untuk membersihkan, merubah, dan mempersiapkan data sebelum di gunakan dalam analisis. Tujuan dari preprocessing yaitu untuk meningkatkan kualitas data dan juga memastikan bahwa data tersebut sesuai dengan kebutuhan analisis yang akan di lakukan. Dalam pengertian yang lainnya menyebutkan bahwa data preprocessing adalah sebagai tahapan-tahapan untuk menghilangkan beberapa masalah yang dapat mengganggu di saat pemrosesan data yang sedang berlangsung. Sehingga hal tersebut mendapat banyaknya data yang formatnya tidak sesuai atau tidak konsisten. Data preprocessing sendiri yang merupakan teknik yang paling awal sebelum

melakukan data mining. Melainkan terdapat beberapa proses dalam data preprocessing, guna memungkinkan proses mining akan berjalan dengan lancar. Dikarenakan data yang sudah melalui tahap prapemrosesan data yang merupakan data sudah melalui beberapa tahapan pembersihan. Tahapan-tahapan di dalam preprocessing adalah sebagai berikut :

1. **Data Cleaning**  
Data cleaning sendiri adalah proses untuk memperbaiki atau menghapus kesalahan dalam data, ketidakkonsistenan data, dan ketidakakuratan data di dalam sekumpulan data. Data cleaning yang merupakan tahapan yang sangat di perlukan dan di lakukan sebelum melakukan analisis data dikarenakan biasanya data yang mentah mengandung informasi-informasi yang tidak akurat, tidak tersusun dengan rapi, dan data tidak lengkap. Cleaning untuk menghapus URL dan HTML, emoji, angka, mention dan hastag, serta symbol yang tidak relevan.
2. **Case Folding** : untuk mengubah teks dokumen menjadi huruf kecil.
3. **Normalisasi** : menangani kata slag atau kata tidak baku.
4. **Tokenization** : Proses memecah teks menjadi bagian-bagian kecil sebelum dianalisis.
5. **Filtering/Stopword Removal** : Proses untuk memfokuskan pada kata-kata yang lebih bermakna dengan menghapus kata-kata yang tidak memiliki makna.
6. **Stemming** : mengubah kata berhimpunan menjadi kata dasar.

## **2.2. Pemisahan data/clustering**

Clustering yang merupakan sebuah metode pengelompokan data atau pemisahan data ke dalam beberapa kluster di dalam penelitian ini membagi menjadi 3 kluster yaitu C1, C2, dan C3.

## **2.3. Analisis Sentimen**

Setelah melalui tahap preprocessing data berikutnya melakukan langkah analisis sentimen/pengklasifikasian tweet. Dalam langkah ini dilakukan untuk labeling sentiment. Dasar penentuan untuk score gabungan menggunakan vader sentiment. Skor gabungan kurang dari 0 adalah sentiment negatif, skor gabungan samadengan 0 adalah sentimen netral, dan skor gabungan lebih besar dari 0 adalah sentimen positif. Sedangkan Analisis sentimen sendiri adalah sebuah proses untuk memahami, mengekstrak dan memproses informasi-informasi perasaan seseorang yang terkait tentang pendapat tersebut, berlandaskan tentang komputasi tentang pendapat, sentimen, opini, maupun perasaan seseorang terhadap suatu topik analisis sentimen.

III. HASIL DAN PEMBAHASAN

3.1. Pengambilan Data

Data yang di gunakan berasal dari salah satu media sosial twitter, pengumpulan data di lakukan menggunakan web scraping melalui aplikasi google collab dengan menggunakan bahasa pemrograman python. Dan kata kunci yang di gunakan untuk mencari data tersebut yaitu (IBU KOTA NEGARA/IKN) data yang di dapat sebanyak 2000 data twett dengan menggunakan bahasa indonesia Pengambilan Data yang dikumpulkan berupa data teks yang diambil dengan teknik crawling data pada tanggal 29 november 2023 pengumpulan data di lakukan kurang lebih selama satu tahun twitt yang di ambil yaitu sebanyak 2000 data twitt yang berisikan pendapat,komentar,opini mengenai pemindahan ibu kotanusantara (IKN). Di bawah ini hasil proses crawling data menggunakan google collab :

created_at	full_text	username	tweet_url	id_str	quote_count	reply_count	retweet_count	favorite_count	lang	user_id_str
Feb 22 09:13	@WIndonesia -sopng	https://twitter.c...	17695912285	0	0	0	0	0	in	112149491001
Feb 22 09:13	@Sinar9 @Pa... Nuclei95	https://twitter.c...	17695908015	0	0	0	0	0	in	124338937995
Feb 22 09:12	@HandaBerC13 Jalew91	https://twitter.c...	17695902553	0	1	0	0	0	in	350433570377
Feb 22 09:08	Hegga Deante... abantra95	https://twitter.c...	17695489293	0	0	0	0	0	in	152844827366
Feb 22 09:08	@bedicon Tumb... harano177	https://twitter.c...	17695488700	0	0	0	0	0	in	833821763262
Feb 22 09:08	Hegga Deante... abantra95	https://twitter.c...	17695488790	0	0	0	0	0	in	152844827366
Feb 22 09:08	Hegga Deante... abantra95	https://twitter.c...	17695488689	0	0	0	0	0	in	152844827366
Feb 22 09:05	Mentor 479,PN... dedicon	https://twitter.c...	17695389825	0	4	0	2	0	in	6918151
Feb 22 09:03	@FayGngIptka... marjin	https://twitter.c...	17695353687	0	0	0	0	0	in	72246152
Feb 22 09:02	Gakur Masih Ter... meid_j	https://twitter.c...	17695314348	0	0	0	0	0	in	960237828
Feb 22 09:01	@bananusa96... realian2	https://twitter.c...	17695308635	0	0	0	0	0	in	238189590
Feb 22 09:01	Salam @M91 net... kurnia214	https://twitter.c...	17695306302	0	0	1	1	0	in	35339789461
Feb 22 09:01	hari ada pabri... hudo180ca	https://twitter.c...	17695278072	0	0	0	0	0	in	174769398461
Feb 22 09:00	Balpapan sap h... probalito	https://twitter.c...	17695264184	0	0	0	0	0	in	312470790
Feb 22 09:00	@endkamarca... siddanabes	https://twitter.c...	17695259376	0	0	0	0	0	in	128019138
Feb 22 09:00	@B9Koro lah kat... realian2	https://twitter.c...	17695226339	0	0	0	0	0	in	238189590
Feb 22 08:57	301 nengad saku... shrejan_2	https://twitter.c...	17695188619	0	0	0	0	0	in	178861588081
Feb 22 08:52	@WIndonesia... agubago	https://twitter.c...	17695186757	0	1	0	0	0	in	325224683
Feb 22 08:46	Charita Perdik... per_dnd	https://twitter.c...	17694977393	0	0	0	0	0	in	148643181932
Feb 22 08:46	Sayang daku ben... her_gone	https://twitter.c...	17694951824	0	0	0	0	0	in	178865047
Feb 22 08:46	@99 katon ha... kumung_rya	https://twitter.c...	17694913583	0	0	0	0	0	in	102888819

Gambar.2. Hasil Dari Crawling Data TweetB. Preprocessing data

Menurut han, J., Kamber,M.,& Pei,J. (2011) preprocessing data adalah tahapan awal dalam

analisis data yang melibatkan serangkaian langkah-langkah untuk membersihkan,merubah, dan mempersiapkan data sebelum di gunakan dalam analisis. Tujuan dari preprocessing yaitu untuk meningkatkan kualitas data dan juga memastikan bahwa data tersebut sesuai dengan kebutuhan analisis yang akan di lakukan. Dalam pengertian lain menyebutkan bahwa data preprocessing adalah sebagai tahapan untuk menghilangkan beberapa masalah yang bisa mengganggu di saat pemerosesan data sedang berlangsung. Sehingga hal tersebut karena banyaknya data yang formatnya tidak sesuai atau tidak konsisten. Data preprocessing yang merupakan teknik yang paling awal sebelum melakukan data mining. Melainkan terdapat beberapa proses dalam data preprocessing, guna memungkinkan proses mining akan berjalan dengan lancar. Dikarenakan data yang sudah melalui tahap prapemerosesan data yang merupakan data sudah melalui beberapa tahapan pembersihan. Tahapan- tahapan di dalam preprocessing adalah sebagai berikut :

1. Data Cleaning

Data cleaning sendiri adalah proses untuk memperbaiki atau menghapus kesalahan dalam data,ketidakkonsistenan data, dan ketidakakuratan data di dalam sekumpulan data. Data cleaning yang merupakan tahapan yang sangat di perlukan dan di lakukan sebelum melalukan analisis data di karenakan biasanya data yang mentah mengandung informasi-informasi yang tidak akurat, tidak tersusun dengan rapi, dan data tidak lengkap.Cleaning untuk menghapus URL dan HTML, emoji, angka, mention dan hastag, serta symbol yang tidak relevan.

Table 1. table hasil cleaning data

Sebelum	Sesudah
@PakarINTElek Ntr kalo pindah keIKN demonya giman min?	Entarkaluapindah ke IKN demonya gimana min.
@tahaRDDddin Angket itu ujungnya DPR Keluarkan rekomendas pd pelaksana dan penyelenggara Pemilu dan pemakzulan krn cukup alasan hukum utk itu. Soal jazah, Hutang dan penambahan anggaran Bansos tanpa persetujuan DPR, soal Gibran, IKN, PSN dan kecurangan.	Angket itu ujungnya DPR keluarkan rekomendasi pd pelaksana dan penyelenggara pemilu dan pemakzulan krn cukup alasan hukum utk itu. Soal ijazah, hutang dan penambahan anggaran bansos tanpa persetujuan DPR, soal Gibran, IKN, PSN dan kecurangan.
@Penyuka_ombak@KPU_ID @mohmahfudmd Presiden di IKNðŸ˜„ðŸ˜„ðŸ˜„, Yang masih warasikut presiden yang di jakarta aja	Yang masih Waras ikut presiden yang di Jakarta aja
@Bentir16 @wantahgdn___ Padahal infrastruktur IKNada ngembat tanahwarga disitu. Saya nonton beritanya	Padahal infrastuktur IKNada ngembat tanah warga situ. Saya nonton beritanya.
@Tirtoid argetnya carisumber duit untuk menyicil utang, membiayai IKN tapi dengan cara mencekik	Targetnya cari sumber duituntuk menyicil utang, membiayai IKN tapi dengan cara mencekik.

2. Case Folding : untuk mengubah teks dokumen menjadi huruf kecil.

Table.2. Case Folding

Sebelum	Sesudah
@RaisaTjokrodnta @destyawana MENDING EKSPOR KELUARGA IJAZAH PALSU DAN PENDUKUNGNYA MEKI GOCENGANREZIM... KE IKN AJAYANG DEKET ITUKLO ORANG SINGAPUR MAUBAU SELANGKANGAN ITEM JAMURAN	mending ekspor keluarga ijazah palsu dan pendukungnyameki gocenganrezim... ke ikn aja yang dekat itu klo orang singapur maubau selangkangan item jamurand
@dewapuci Sori...gwketawa2 aja sih liat IKN IKN ini. Jkw itu udh terkenal dg janji2nya. Liat aja Esemka, janji gakngutang, janji gakbakal impor....halah. Kl nanti ini jg akanberakibat tinggal janji, plg dia cm cengarcengir doang sambil ngomong lha	sori...gw ketawa2 ajasih liat ikn ikn ini.jkw itu udh terkenal dg janji2nya. liat aja esemka, janji gak ngutang, janji gak bakal impor....halah. kl nanti ini jg akan berakibat tinggaljanji, plg dia cm cengar cengir doang sambil ngomong lha
@DokterTifa Faktanya Anies Tidak berani terang terangan Menolak IKN dalam kampanyenya. Mau bagaimanapun jugajika Anies Terpilihmenjadi Presiden Aniesakan melanjutkan IKN	faktanya anies tidakberani terangterangan menolak ikn dalam kampanyenya. mau bagaimanapun juga jika anies terpilih menjadi presiden anies akan melanjutkan ikn
@PDemokrat Termasuk penyelewengan FOOD ESTATE, REMPANG, IKN juga kah..â• %oi•	termasuk penyelewengan food estate, rempang, ikn juga kah..â• %oi•
@windaherdisa Lanjut ke Kalimantan aja gimana, Pak? Start fresh di IKN. Asa lebar DKI yg udah ditata sama Pak Anies (dan udah ada yg diotak-atik sama Heru) ngomong lha	lanjut ke kalimantan aja gimana, pak? start fresh di ikn. asa lebar dki yg udah ditata sama pak anies (dan udah ada yg diotak-atik sama heru)
@DokterTifa Faktanya Anies Tidak berani terang terangan Menolak IKN dalam kampanyenya. Mau bagaimanapun jugajika Anies Terpilihmenjadi Presiden Aniesakan melanjutkan IKN	faktanya anies tidakberani terangterangan menolak ikn dalam kampanyenya. mau bagaimanapun juga jika anies terpilih menjadi presiden anies akan melanjutkan ikn
@PDemokrat Termasuk penyelewengan FOOD ESTATE, REMPANG, IKN juga kah..â• %oi•	termasuk penyelewengan foodestate, rempang, ikn juga kah..â• %oi•
@windaherdisa Lanjut ke Kalimantan aja gimana, Pak? Startfresh di IKN. Asa lebar DKI yg udah ditata sama Pak Anies (dan udah ada yg diotak-atik sama Heru)	lanjut ke kalimantan aja gimana, pak? start fresh di ikn. asa lebar dki yg udah ditata sama pak anies(dan udah ada yg diotak-atik sama heru)
@RaisaTjokrodnta @destyawana MENDING EKSPOR KELUARGA IJAZAH PALSU DAN PENDUKUNGNYA MEKI GOCENGANREZIM... KE IKN AJAYANG DEKET ITUKLO ORANG SINGAPUR MAUBAU SELANGKANGAN ITEM JAMURAN	mending ekspor keluarga ijazah palsu dan pendukungnyameki gocenganrezim... ke ikn aja yang dekat itu klo orang singapur maubau selangkangan item jamurand
@dewapuci Sori...gwketawa2 aja sih liat IKN IKN ini. Jkw itu udh terkenal dg janji2nya. Liat aja Esemka, janji gak ngutang, janji gak	sori...gw ketawa2 ajasih liat ikn ikn ini. jkw itu udh terkenal dg janji2nya. liat aja esemka, janji gak ngutang, janji gak

3. Normalisasi : menangani kata slag atau katatidak baku.

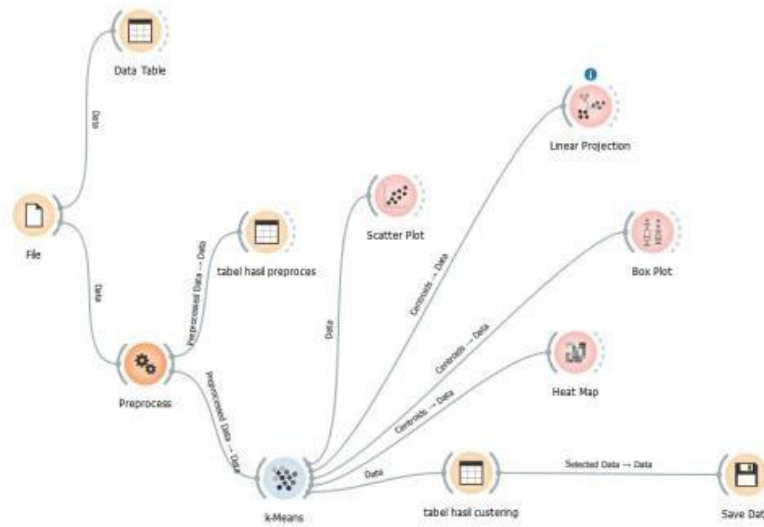
Table.3. Normalisasi

Sebelum	Sesudah
@RaisaTjokrodnta @destyawana MENDI EKSPOR KELUARGA IJAZAH PALSU DAN PENDUKUNGNYA MEKI GOCENGANREZIM... KE IKN AJAYANG DEKET ITUKLO ORANG SINGAPUR MAUBAU SELANGKANGAN ITEM JAMURAN	mending ekspor keluarga ijazah palsu dan pendukungnyameki gocenganrezim... ke ikn aja yang dekat itu klo orang singapur maubau selangkangan item jamurand
@dewapuci Sori...gwketawa2 aja sih liat IKN IKN ini. Jkw itu udh terkenal dg janji2nya. Liat aja Esemka, janji gak ngutang, janji gakbakal impor....halah. Kl nanti ini jg akanberakibat tinggal janji,plg dia cm cengar	sori...gw ketawa2 ajasih liat ikn ikn ini.jkw itu udh terkenal dg janji2nya. liat ajaesemka, janji gak ngutang, janji gakbakal impor....halah. kl nanti ini jg akan berakibat tinggaljanji, plg dia cm cengar

cengir doang sambil ngutang, janji gak bakal impor....halah. Kl nanti ini jg akan berakibat tinggal janji, plg dia cm cengar cengir doang sambil ngomong lha	cengir doang sambil ngomong lha bakal impor....halah. kl nanti ini jg akan berakibat tinggal janji, plg dia cm cengar cengir doang sambil ngomong lha
@DokterTifa Faktanya Anies Tidak berani terang terangan Menolak IKN dalam kampanyenya. Mau bagaiamanapun juga jika Anies Terpilih menjadi Presiden Anies akan melanjutkan IKN	faktanya anies tidak berani terang terangan menolak ikn dalam kampanyenya. mau bagaiamanapun juga jika anies terpilih menjadi presiden anies.....akan melanjutkan ikn
@PDemokrat Termasuk penyelewengan FOOD ESTATE, REMPANG, IKN juga kah..â• %öï,•	termasuk penyelewengan food estate, rempang, ikn juga kah..â• %öï,•
@windaherdisa Lanjut ke Kalimantan aja gimana, Pak? Start fresh di IKN. Asa lebar DKI yg udah ditata sama Pak Anies (dan udah ada yg diotak-atik sama Heru)	lanjut ke kalimantan aja gimana, pak? start fresh di ikn. asa lebar dki yg udah ditata sama pak anies (dan udah ada yg diotak-atik sama heru)

**3.2. Pemisahan Data**

Clustering yang merupakan sebuah metode pengelompokan data atau pemisahan data ke dalam beberapa kluster di dalam penelitian ini membagi menjadi 3 kluster yaitu C1,C2,dan C3. Dalam penelitian ini clustering menggunakan metode k- means dan menggunakan orange data mining. Berikut hasil clustering yang di peroleh menggunakan orange data mining :



Gambar.7. hasil pemerosean clustering

Table hasil clustering data menggunakan orangedata mining :

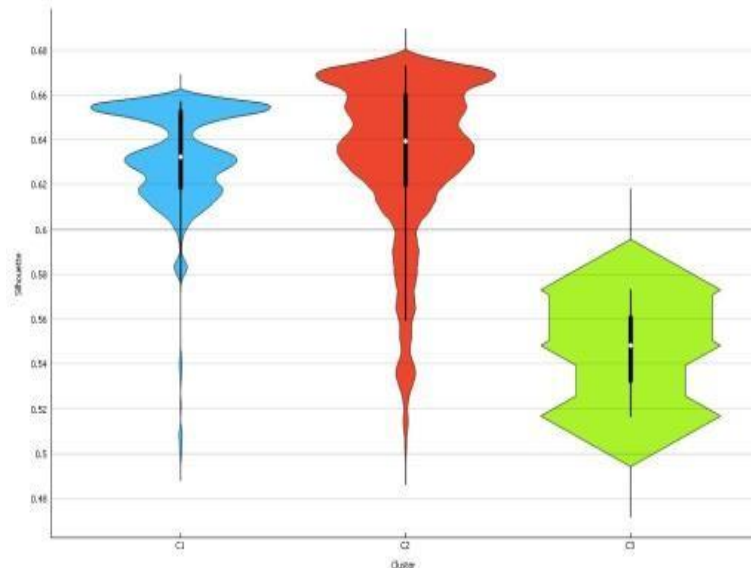


Table. 8 Hasil Clustering

Dari hasil table di atas kluster tertinggi yaitu C2 .

	created_at	full_text	username	tweet_url	Cluster	Silhouette	id_str	quote_count	reply_count	retweet_count
452	Fri Feb 23 00:00:...	Pemerintah meny...	Saadah1234	https://twitter.co...	C2	0.61225	1760816886663	2	3	19
451	Fri Feb 23 00:04:...	Ternyata Sri Muly...	OneLove4Human	https://twitter.co...	C2	0.646177	1760817953887	0	0	0
450	Fri Feb 23 00:05:...	Lahan 90% dibab...	NRIndonesia79	https://twitter.co...	C1	0.633276	1760818041737	1	0	3
449	Fri Feb 23 00:05:...	@geloraco Lagi...	bluebingreyabu	https://twitter.co...	C2	0.647142	1760818199946	0	0	0
448	Fri Feb 23 00:06:...	Hanya kenderaa...	Melvindicze	https://twitter.co...	C2	0.542875	1760818439164	0	0	0
447	Fri Feb 23 00:07:...	IKN Nusantara m...	nsingrings	https://twitter.co...	C1	0.634154	1760818517002	0	0	0
446	Fri Feb 23 00:08:...	Mengan RB Mata...	portalbontang_	https://twitter.co...	C2	0.643314	1760818921278	0	0	0
445	Fri Feb 23 00:09:...	Ortu gw rihony...	velvedist	https://twitter.co...	C1	0.633889	1760819241195	0	1	0
444	Fri Feb 23 00:10:...	Mari kita usaha...	velvedist	https://twitter.co...	C1	0.633882	1760819309632	0	1	0
443	Fri Feb 23 00:11:...	IKN Nusantara m...	terindukan	https://twitter.co...	C2	0.644478	1760819674068	0	0	0
442	Fri Feb 23 00:12:...	@bhagiarus Ma...	bventymarcusai	https://twitter.co...	C2	0.623613	1760819959385	0	1	0
441	Fri Feb 23 00:13:...	Jenderal Marul S...	tempodotco	https://twitter.co...	C1	0.633969	1760820043306	0	0	0
440	Fri Feb 23 00:14:...	@m00711 @sak...	halo_jentera	https://twitter.co...	C2	0.61686	1760820351913	0	0	0
439	Fri Feb 23 00:15:...	@kompascom Da...	Dedy44369434	https://twitter.co...	C2	0.578044	1760820565001	0	0	0
438	Fri Feb 23 00:16:...	Jujurly aku tm y...	vrgglass	https://twitter.co...	C2	0.565844	1760820946515	0	1	0
437	Fri Feb 23 00:18:...	@agustian_y17 T...	qainta1217	https://twitter.co...	C2	0.553955	1760821503585	0	1	0
436	Fri Feb 23 00:20:...	Otonita BKN Pape...	kikwaria_com	https://twitter.co...	C2	0.535387	1760821781026	0	0	0
435	Fri Feb 23 00:20:...	@intonyusupr Ik...	ironjeons	https://twitter.co...	C2	0.536857	1760821836655	0	0	0
434	Fri Feb 23 00:20:...	@qainta1217 A...	agustian_y17	https://twitter.co...	C2	0.644268	1760821848533	0	1	0
433	Fri Feb 23 00:21:...	Plus memberikan...	msed_bdu	https://twitter.co...	C2	0.578822	1760822063386	0	5	35
432	Fri Feb 23 00:25:...	ASN yang akan p...	kompascom	https://twitter.co...	C1	0.633365	1760823179462	0	1	1
430	Fri Feb 23 00:26:...	@panca66 @Tan...	batu_candra	https://twitter.co...	C1	0.634074	1760823441647	0	0	0
431	Fri Feb 23 00:26:...	@sugar_baby97 p...	pacarseungyun_	https://twitter.co...	C1	0.627038	1760823441282	0	1	0
429	Fri Feb 23 00:28:...	@Hansumiko Ma...	Anggita_jung	https://twitter.co...	C1	0.633706	1760823493221	0	0	0
428	Fri Feb 23 00:27:...	@nurur_naufal S...	SD6eeeeeeee_e	https://twitter.co...	C2	0.635628	1760823617624	0	0	0
427	Fri Feb 23 00:29:...	@geloraco Dlu a...	asgrayu	https://twitter.co...	C2	0.637435	1760824122492	0	0	2
426	Fri Feb 23 00:30:...	@yuuuuee_@el... Kemataperisai		https://twitter.co...	C2	0.637779	1760824316634	0	0	0
425	Fri Feb 23 00:31:...	@PDenokrat Ter...	Cahyo_Stronger	https://twitter.co...	C1	0.633712	1760824768570	0	0	0
424	Fri Feb 23 00:32:...	@geloraco Jd bu...	RMooh	https://twitter.co...	C2	0.618715	1760824860505	0	0	0

Gambar.9. hasil dari violin plot

Violin plot yaitu gabungan dari kotak dan plot kepadatan karnel,yang menunjukkan puncaknya dalam suatu data. Dari hasil clustering di atas bahwasalahasil C2 hasil yang tertinggi hingga mencapai 0.68. Hasil Analisis Sentimen dapat dilihat pada data gambar berikut ini:

	A	B	C	D	E	F	G	H	I
1	Source text	Sentiment	Sentiment Score	Magnitude	Subjectivity	Entities	Themes	Keywords	Auto Categori
2	@CNNIndonesia sepe... malah jadi 0, km presiden dan wa	neutral	0.158	0.367	unknown				
3	@siane95 @HandalBen12382 Vio IKN tidak memakmurkan	negative	-0.603	0.75	unknown		sempat cerita bisa, negative, -0.515		
4	@HandalBen12382 @Machos90E Dibungkam dgn IKN sng	negative	-0.608	0.782	unknown		dibungkam dgn kn sng, negative, -0.543		
5	Hingga Desember 2024, sebanyak kurang lebih 12 ribu pegawai yang terdiri dari ja		0	0					
6	@etikcom Tumben searang @Agustudhoyono mau berkunjung ke proyek IKN yg		0	0					
7	Hingga Desember 2024, sebanyak kurang lebih 12 ribu pegawai yang terdiri dari ja		0	0					
8	Hingga Desember 2024, sebanyak kurang lebih 12 ribu pegawai yang terdiri dari ja		0	0					
9	Menteri ATN/BPN AHY akan berkunjung ke IKN Nusantara pekan depan. AHY diseb		0	0					
10	@PacDing Ngebayangin kalo istana negara uadh di kn, mau demo jauh banget. Apa		0	0					
11	Golkar Masih Tertalu Perkas di IKN Nusantara, PPP Terdepak	negative	-0.584	0.725	unknown	(PERSON)	Gok masih tertalu per golkar, neutral, -0.247		
12	@ranalmusyaffak lah katanya indo negara maju ada kereta cepat, mrt, ibu kota is		0	0					
13	Saham BUMN melejit dan investor IKN membuat QRIS kusri	negative	-0.584	0.725	unknown		jumat siang aang, negative, -0.515		
14	Nanti ada pabrik bikin sapu lidi investor dr china bahan baku dr vietnam Dan didirik		0	0					
15	Baligpapan siap hadapi pemindahan IKN dengan rencana pembangunan Rumah Si		0	0					
16	@andikamalreza Trauma Jabar culup di Jabar aja. Atau taro IKN sana ajalah. Dya l		0	0					
17	@ElyKoro lah katanya indo negara maju ada kereta cepat, mrt, ibu kota baru IKN		0	0					
18	IKN meladi solusi untuk menghindari rakut vs secertri mi	neutral	-0.807	0.782	unknown		rakut vs secertri, negative, -0.541		

Gambar.10. hasil sentimen analisis

Sentiment	COUNTA of Sen
neutral	20,69%
negative	79,31%
	0,00%
<b>Grand Total</b>	<b>100.00%</b>

Gambar.11. tabel pivote

### 3.3. Analisis Sentimen

Dari hasil analisis sentimen dan clustering menggunakan metode *k-means clustering* di atas bahwa hasil tertinggi sebanyak 79,31% dengan tanggapan masyarakat mengenai pemindahan ibu kota nusantara atau IKN yaitu secara negatif dan yang beranggapan netral sebanyak 20,69% dan hasil tertinggi dari clustering menggunakan metode *k-means clustering* yaitu cendroid 2 atau C2 hasil dari tabel violin plot sebesar 0.68.

## IV . KESIMPULAN

Untuk melakukan pengelompokan data dapat dilakukan dengan menerapkan metode *K-means clustering* .data yang di peroleh di peroses dengan menggunakan orange data mining . nilai-nilai cendroid di tentukan dengan membaginya menjadi 3 kelompok yaitu C1,C2,dan C3. Dari hasil analisis sentimen dan clustering menggunakan metode *k-means clustering* di atas bahwa hasil tertinggi sebanyak 79,31% dengan tanggapan masyarakat mengenai pemindahan ibu kota nusantara atau IKN yaitu secara negatif dan yang beranggapan netral sebanyak 20,69%. Dan hasil tertinggi dari clustering menggunakan metode *k-means clustering* yaitu cendroid 2 atau C2 hasil dari tabel violin plot sebesar 0.68. Saran untuk penelitian ini pada analisis ini sangat di perlukan dan di perhatikan dalam pengambilan data,clustering data dan analisis sentimen di karenakan agar mendapat hasil yang baik .

## DAFTAR PUSTAKA

- [1] Andaru Liken Anggoro,Lulasnova Violprameswari ,”*Analisis Media Text Clustering Pada Twitter Akan Kasus Selebriti Menggunakan Orange Data Minin*”, Remik,2023.
- [2] Krisna Jhonatan,Sitorus,Anita Mulia Wati, Sarika,” *Klasifikasidan Analisis sentiment Pada Data Twitter Menggunakan Algoritma Naiv Bayes .(Studi Kasus : Pekan Olahraga Nasional Xx 2021)*” Jakarta 20 Agustus 2022.
- [3] Tania Syifa Utami, “ *Analisis Sentimen Judul Berita Mengenai Pemindahan Ibu Kota Negara (IKN) Pada Portal Berita Detik.Com Menggunakan Support Vector Machine (SVM)*”. Jakarta 2023.
- [4] Sheila Shevira, I Made Agus Dwi Suarjaya,Putu Wira Buana,” *Pengaruh Kombinasi Dan Urutan Pre-Processing Pada Tweets Bahasa Indonesia*”. Universitas Udayana Agustus 2022.
- [5] Hafizh Dery Al Assyam, Firman Noor Hasan,”*Analisis Sentimen Twitter Terhadap Pemindahan Ibu Kota Negara Ke IKN Nusantara Menggunakan Orange Data Mining*”. Jakarta 2023.
- [6] Sulstri, Dwi Agus Diartono,”*Analisis Jejaring Sosial Twitter Menggunakan Kalstering K-Means Dan Hirarki Agglomeratif*” Semarang 2019.
- [7] Ni Putu Eka Merlina, Ernawati,Alb. Joko Santoso,”*Analisa Penentuan Jumlas Cluster Terbaik Pada Metode K-Mean Clustering*”. Universitas Atmajaya , Yogyakarta.
- [8] Ratu Aghnia Raffaidy Wiguna,Andri Irfan Rifai,”*Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibuslaw Menggunakan Orange Data Mining* “.Bandung 2021.
- [9] Muhammad Ilham Ramadhon,” Analisis Sentimen Terhadap Pemindahan Ibu Kota Indonesia Pada Media Sosial Twitter Menggunakan Metode Algoritma K-Nearst Neighbor (K-NN). UINSH Jakarta 2020.
- [10] Rizky Kurnia Pratama, Putry Wahyu Setyaningsih,”*Analisis Komentar Pada Twitter Terhadap Lapangan Kerja Dengan Metode Naiv Bayes*”.UMBY Yogyakarta 2022.
- [11] Raihanda Luthfiansyah, Budi Wasito,” *Analisis Sentimen Terhadap Para Kandidat Presiden 2024 Berdasarkan Netizen Pengguna Twitter Dengan Metode Data Mining Dan Text Mining*” .2024.
- [12] Miranti Alysha Zulia Larasati,Nurul Anisa Sari Winarsih, Muammad Syaifur Rohman, Galuh Wilujeng Saraswati,” *Penerapan Metode K-Means Clustering Dalam Menganalisis Sentimen Masyarakat Terhadap K-Popers Pada Twitter*”.UDN Semarang 2 Agustus 2022.
- [13] Endahns,”*Klasifikasi Data Menggunakan Metode K-Means Clustering Dengan Aplikasi Orang*”.27 Juni 2020.
- [14] Joel Carron, Ilmuan Data Di Mode, “*Pengertian Violin Plot*”. Februari 2024.