

**IMPLEMENTASI SELEKSI FITUR *INFORMATION GAIN* UNTUK
KLASIFIKASI BERITA *HOAX* DENGAN MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR***

SKRIPSI



MARISSA AUDINA

NIM. 1808561020

**PROGRAM STUDI INFORMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS UDAYANA**

JIMBARAN

2022

SURAT PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini menyatakan bahwa naskah Skripsi dengan judul:
**“IMPLEMENTASI SELEKSI FITUR *INFORMATION GAIN* UNTUK
KLASIFIKASI BERITA *HOAX* DENGAN MENGGUNAKAN METODE *K-NEAREST NEIGHBOR*”.**

Nama : Marissa Audina
NIM : 1808561020
Program Studi : Informatika
E-mail : marissaaudina@gmail.com
Nomor telp/HP : 081935400770
Alamat : Jl. Gunung Welirang Gang Pura Luhur Sandat I No. 21,
Pemecutan Klod, Denpasar Barat

Belum pernah dipublikasikan dalam dokumen skripsi, jurnal nasional maupun internasional atau dalam prosiding manapun, dan tidak sedang atau akan diajukan untuk publikasi di jurnal atau prosiding manapun. Apabila di kemudian hari terbukti terdapat pelanggaran kaidah-kaidah akademik pada karya ilmiah saya, maka saya bersedia menanggung sanksi-sanksi yang dijatuhkan karena kesalahan tersebut, sebagaimana diatur oleh Peraturan Menteri Pendidikan Nasional Nomor 17 Tahun 2010 tentang Pencegahan dan Penanggulangan Plagiat di Perguruan Tinggi.

Demikian Surat Pernyataan ini saya buat dengan sesungguhnya untuk dapat dipergunakan bilamana diperlukan.

Bukit Jimbaran, 6 April 2022

Yang membuat pernyataan,

Marissa Audina

NIM. 1808561020

LEMBAR PENGESAHAN TUGAS AKHIR

Judul : Implementasi Seleksi Fitur *Information Gain* untuk
Klasifikasi Berita *Hoax* dengan Menggunakan Metode *K-Nearest Neighbor*
Nama : Marissa Audina
NIM : 1808561020
Tanggal Seminar : 6 April 2022

Disetujui oleh:

Pembimbing I



Dr. Anak Agung Istri Ngurah Eka
Karyawati, S.Si., M.Eng.
NIP. 197404071998022001

Penguji I



Dr. Ir. I Ketut Gede Suhartana,
S.Kom., M.Kom
NIP. 197201102008121001

Pembimbing II



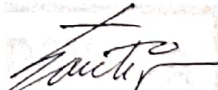
I Wawan Supriana, S.Si., M.Cs
NIP. 1984082920181113001

Penguji II



I Gede Santi Astawa, S.T., M.Cs.
NIP. 198012062006041003

Penguji III



Drs. I Wawan Santiyasa, M.Si
NIP. 196704141992031002

Mengetahui,

Program Studi Informatika
FMIPA UNUD
KPS,



Dr. Ir. I Ketut Gede Suhartana, S.Kom., M.Kom.
NIP. 197201102008121001

Judul : Implementasi Seleksi Fitur *Information Gain* untuk Klasifikasi Berita *Hoax* dengan Menggunakan Metode *K-Nearest Neighbor*

Nama : Marissa Audina (NIM: 1808561020)

Pembimbing : 1. Dr. Anak Agung Istri Ngurah Eka Karyawati,
S.Si.,M.Eng.
2. I Wayan Supriana, S.Si., M.Cs

ABSTRAK

Berita adalah salah satu sumber informasi yang digunakan oleh masyarakat. Namun, tidak semua berita yang beredar di media digital adalah fakta. Beberapa oknum mengambil kesempatan untuk membagikan berita-berita yang tidak berdasar dan tidak dapat dipertanggungjawabkan kebenarannya. Sejak pandemi Covid-19 melanda Indonesia, berita *hoax* mengenai pandemi tersebut semakin banyak beredar di media digital. Pada penelitian kali ini penulis membangun sebuah model yang dapat mengklasifikasikan berita *hoax* dengan menggunakan metode *K-Nearest Neighbor* yang dikombinasikan dengan seleksi fitur *Information Gain*. Data yang digunakan adalah data berita fakta dan data berita *hoax* berbahasa Indonesia. Evaluasi dilakukan dengan mengukur performa model *K-Nearest Neighbor* tanpa seleksi fitur dan performa model dengan implementasi seleksi fitur *Information Gain*. Model *K-Nearest Neighbor* tanpa seleksi fitur dengan nilai $k=5$ menghasilkan performa *precision*, *recall*, *F1-Score*, dan akurasi secara berturut-turut yaitu 87.5%, 96,5%, 91,8%, dan 91,6%. Sedangkan model *K-Nearest Neighbor* dengan kombinasi seleksi fitur *Information Gain threshold* 0.5% dengan nilai $k=3$ menghasilkan performa *precision*, *recall*, *F1-Score*, dan akurasi secara berturut-turut yaitu 93.3%, 96.6%, 95%, dan 95%.

Kata kunci: *K-Nearest Neighbor*, *Information Gain*, TF-IDF, Klasifikasi Teks, Berita *Hoax*

Title : Implementation of Information Gain Feature Selection for
Hoax News Classification using K-Nearest Neighbor
Method

Name : Marissa Audina (NIM: 1808561020)

Supervisor : 1. Dr. Anak Agung Istri Ngurah Eka Karyawati,
S.Si.,M.Eng.
2. I Wayan Supriana, S.Si., M.Cs

ABSTRACT

News is one of information resources that is being used by the public. However, not all news circulating in digital media are facts. Some people take the opportunity to share unfounded and irresponsible news. Since the Covid-19 pandemic hit Indonesia, hoax news about the pandemic has increasingly circulated in digital media. In this study, the author builds a model that can classify hoax news using the K-Nearest Neighbor method combined with the Information Gain feature selection. The data used are factual news data and hoax news data in Indonesian language. Evaluation is done by measuring the performance of the K-Nearest Neighbor model without feature selection and model performance by implementing Information Gain feature selection. The K-Nearest Neighbor model without feature selection with a value of $k=5$ obtained precision, recall, F1-Score, and accuracy performance of 87.5%, 96.5%, 91.8%, and 91.6%, respectively. While the K-Nearest Neighbor model with a combination of 0.5% Information Gain threshold feature selection with a value of $k=3$ obtained precision, recall, F1-Score, and accuracy performance of 93.3%, 96.6%, 95%, and 95%, respectively.

Keywords: K-Nearest Neighbor, Information Gain, TF-IDF, Text Classification, Hoax News

KATA PENGANTAR

Puji syukur kepada Allah SWT yang selama ini telah melancarkan seluruh proses penelitian sehingga penulis dapat menyelesaikan penelitian dengan judul “Implementasi Seleksi Fitur *Information Gain* untuk Klasifikasi Berita *Hoax* dengan Menggunakan Metode *K-Nearest Neighbor*”. Penelitian ini disusun dalam rangkaian kegiatan pelaksanaan Tugas Akhir di Program Studi Informatika FMIPA UNUD. Penelitian ini dilaksanakan pada periode Agustus 2021 hingga April 2022 di Universitas Udayana.

Sehubungan dengan telah terselesaikannya penelitian ini, maka diucapkan terima kasih dan penghargaan kepada berbagai pihak yang telah membantu penulis, antara lain:

1. Bapak Dr. Ir. I Ketut Gede Suhartana, S.Kom., M.Kom sebagai Koordinator Program Studi Informatika Universitas Udayana;
2. Bapak I Gusti Ngurah Anom Cahyadi Putra, ST., M.Cs sebagai Ketua Komisi Tugas Akhir Program Studi Informatika Universitas Udayana;
3. Ibu Dr. Anak Agung Istri Ngurah Eka Karyawati, S.Si., M.Eng. sebagai Pembimbing I sekaligus Pembimbing Akademik yang telah banyak meluangkan waktu untuk membantu pelaksanaan penelitian ini;
4. Bapak I Wayan Supriana, S.Si., M.Cs sebagai Pembimbing II yang telah bersedia mengkritisi, memeriksa, dan menyempurnakan penulisan tugas akhir ini;
5. Bapak-bapak dan Ibu-ibu dosen di Program Studi Informatika yang telah meluangkan waktu turut memberikan saran dan masukan dalam pelaksanaan penelitian;
6. Seluruh keluarga penulis khususnya orang tua, Ibu Elvi Sutarsih dan Bapak Saprawi, serta adik Rama Raditya Ahmad yang selalu memberikan kasih sayang, doa, nasihat, dukungan, dan kesabarannya yang luar biasa dalam setiap langkah hidup penulis;

7. Teman-teman di Program Studi Informatika, Safira, Farin Istighfarizky, Ni Putu Ambalika Dewi, Sang Putu Febri Wira Pratama, dan Ayu Kadek Nadya Oktaviana, yang telah memberikan dukungan dan berjuang bersama sejak awal perkuliahan hingga terselesaikannya penelitian ini;
8. Seluruh sejawat dan kerabat yang tidak mungkin penulis sebutkan satu-persatu yang telah membantu dan memberikan dukungan moral dalam penyelesaian tugas akhir ini;
9. Diri sendiri yang selalu berusaha dan tidak pernah menyerah dalam keadaan apapun;

Disadari pula bahwa sudah tentu hasil-hasil dari penelitian ini masih mengandung kelemahan dan kekurangan. Memperhatikan hal ini, maka masukan dan saran-saran penyempurnaan sangat diharapkan.

Bukit Jimbaran, 6 April 2022

Penyusun,

Marissa Audina

DAFTAR ISI

LEMBAR JUDUL	i
SURAT PERNYATAAN KEASLIAN KARYA ILMIAH.....	ii
LEMBAR PENGESAHAN TUGAS AKHIR	iii
ABSTRAK	iv
ABSTRACT.....	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	2
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Tinjauan Teori	4
2.1.1 Berita <i>Hoax</i>	4
2.1.2 Klasifikasi Teks.....	4
2.1.3 <i>Preprocessing</i>	4
2.1.4 <i>Term Frequency Inverse Document Frequency (TF-IDF)</i>	5
2.1.5 <i>K-Nearest Neighbor</i>	6
2.1.6 <i>Information Gain</i>	7
2.1.7 <i>Confusion Matrix</i>	8

2.2	Tinjauan Empiris	9
2.2.1	Detecting Hoaxes in Indonesian News using TF/TDM and K-Nearest Neighbor (Zuliarso et al, 2020).....	9
2.2.2	The Effect of Information Gain Feature Selection for Hoax Identification in Twitter Using Classification Method Support Vector Machine (Mubarok & Setiawan, 2020).....	10
2.2.3	Analisis Sentimen Terhadap Ulasan Pengguna MRT Jakarta Menggunakan <i>Information Gain</i> dan <i>Modified K-Nearest Neighbor</i> (Paramitha, Indriati, dan Sari, 2020)	10
2.2.4	Pemanfaatan Klasifikasi Soal Biologi <i>Cognitive Domain Bloom's Taxonomy</i> Menggunakan KNN <i>Chi-Square</i> Sebagai Penyusunan Naskah Soal (Listiowarni & Dewi, 2020).....	11
2.2.5	Klasifikasi Berita <i>Clickbait</i> Menggunakan <i>K-Nearest Neighbor</i> (Sagita, Enri, dan Primajaya, 2020)	11
BAB III METODOLOGI PENELITIAN.....		13
3.1	Data dan Metode Pengumpulan Data	13
3.2	Alur Penelitian.....	13
3.3	<i>Preprocessing</i>	15
3.4	Pembobotan TF-IDF.....	16
3.5	Seleksi Fitur <i>Information Gain</i>	18
3.6	<i>K-Nearest Neighbor</i>	19
3.7	Desain Evaluasi Metode	22
3.7.1	Skenario Pengujian.....	22
3.7.2	Evaluasi Metode.....	24
BAB IV HASIL DAN PEMBAHASAN		26
4.1	Implementasi Sistem	26
4.2	Implementasi Memuat dan Membaca Dataset	26

4.3	Implementasi <i>Preprocessing</i>	27
4.3.1	Implementasi <i>Case Folding</i>	27
4.3.2	Implementasi <i>Cleansing</i>	27
4.3.3	Implementasi Tokenisasi.....	28
4.3.4	Implementasi <i>Stopword Removal</i>	28
4.3.5	Implementasi Normalisasi.....	28
4.3.6	Implementasi <i>Stemming</i>	29
4.4	Implementasi Pembobotan TF-IDF	30
4.5	Implementasi Seleksi Fitur <i>Information Gain</i>	30
4.6	Implementasi Metode <i>K-Nearest Neighbor</i>	34
4.6.1	Implementasi <i>Euclidean Distance</i>	34
4.6.2	Implementasi <i>K-Nearest Neighbor</i>	35
4.7	Hasil Penelitian.....	36
4.7.1	Hasil Pengujian Metode <i>K-Nearest Neighbor</i> tanpa Seleksi Fitur..	36
4.7.2	Hasil Pengujian Metode <i>K-Nearest Neighbor</i> dengan Seleksi Fitur <i>Information Gain</i>	38
4.7.3	Pengaruh <i>Threshold Information Gain</i> terhadap Pengujian Metode <i>K-Nearest Neighbor</i>	45
4.7.4	Pengujian Model Terbaik.....	47
BAB V KESIMPULAN DAN SARAN.....		50
5.1	Kesimpulan.....	50
5.2	Saran.....	50
DAFTAR PUSTAKA		51
LAMPIRAN.....		53

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	8
Tabel 4.1 Penggalan Kode Memuat Dokumen	26
Tabel 4.2 Penggalan Kode Proses <i>Case Folding</i>	27
Tabel 4.3 Penggalan Kode Proses <i>Cleansing</i>	27
Tabel 4.4 Penggalan Kode Proses Tokenisasi.....	28
Tabel 4.5 Penggalan Kode Proses <i>Stopword Removal</i>	28
Tabel 4.6 Penggalan Kode Proses Normalisasi.....	29
Tabel 4.7 Penggalan Kode Proses <i>Stemming</i>	29
Tabel 4.8 Penggalan Kode Proses Pembobotan TF-IDF	30
Tabel 4.9a Penggalan Kode Proses <i>Information Gain</i>	31
Tabel 4.9b Penggalan Kode Proses <i>Information Gain</i>	32
Tabel 4.9c Penggalan Kode Proses <i>Information Gain</i>	32
Tabel 4.9d Penggalan Kode Proses <i>Information Gain</i>	33
Tabel 4.9e Penggalan Kode Proses <i>Information Gain</i>	34
Tabel 4.10 Penggalan Kode Proses <i>Euclidean Distance</i>	35
Tabel 4.11 Penggalan Kode Proses <i>K-Nearest Neighbor</i>	36
Tabel 4.12 Hasil Pengujian <i>K-Nearest Neighbor</i> tanpa Seleksi Fitur.....	37
Tabel 4.13 Pengujian <i>Information Gain Threshold</i> 50%	38
Tabel 4.14 Pengujian <i>Information Gain Threshold</i> 25%	39
Tabel 4.15 Pengujian <i>Information Gain Threshold</i> 20%	40
Tabel 4.16 Pengujian <i>Information Gain Threshold</i> 10%	40
Tabel 4.17 Pengujian <i>Information Gain Threshold</i> 5%	41
Tabel 4.18 Pengujian <i>Information Gain Threshold</i> 2%	42
Tabel 4.19 Pengujian <i>Information Gain Threshold</i> 1%	42
Tabel 4.20 Pengujian <i>Information Gain Threshold</i> 0.5%	43
Tabel 4.21 Pengujian <i>Information Gain Threshold</i> 0.2%	44
Tabel 4.22 Pengujian <i>Information Gain Threshold</i> 0.1%	44
Tabel 4.22 Perbandingan Hasil Pengujian Model.....	48

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	14
Gambar 3.2 Alur <i>Preprocessing</i>	16
Gambar 3.3 Pembobotan TF-IDF	17
Gambar 3.4 Seleksi Fitur <i>Information Gain</i>	19
Gambar 3.5 <i>Training K-Nearest Neighbor</i>	21
Gambar 3.6 <i>Testing K-Nearest Neighbor</i>	22
Gambar 3.7 <i>N-Fold Cross Validation</i>	23
Gambar 4.1 Tampilan <i>Data Training</i>	26
Gambar 4.2 Hasil Pengujian <i>K-Nearest Neighbor</i> tanpa Seleksi Fitur.....	37
Gambar 4.3 Pengaruh <i>Threshold</i> terhadap Akurasi	45
Gambar 4.4 Pengaruh <i>Threshold</i> terhadap <i>F1-Score</i>	46
Gambar 4.5 Hasil Pengujian Model Terbaik.....	48

DAFTAR LAMPIRAN

Lampiran :

1. Contoh Prediksi Data *Testing* Berita

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berita adalah salah satu sumber informasi yang digunakan oleh masyarakat. Menurut Kamus Besar Bahasa Indonesia, berita merupakan cerita atau keterangan mengenai kejadian atau peristiwa yang hangat. Berita dapat ditemukan melalui berbagai media, contohnya koran, majalah, dan beberapa media digital seperti portal berita, media sosial, dan lain-lain. Berita-berita yang diakses dapat memuat topik seperti kesehatan, politik, pendidikan, olahraga, dan lain-lain.

Tidak semua berita yang beredar di media digital adalah fakta. Beberapa oknum mengambil kesempatan untuk membagikan berita-berita yang tidak berdasar dan tidak dapat dipertanggungjawabkan kebenarannya. Menurut Kamus Besar Bahasa Indonesia, hoaks (bahasa Inggris: *hoax*) memiliki makna tidak benar atau bohong (tentang berita, pesan, dan sebagainya). Sejak pandemi COVID-19 melanda Indonesia, berita *hoax* mengenai pandemi tersebut semakin banyak beredar di media digital. Data terbaru dari Kementerian Komunikasi dan Informatika, sebanyak 5457 sebaran *hoax* Covid-19 sudah ditindaklanjuti sejak 23 Januari 2020 hingga 18 Maret 2022.

Berita-berita yang didapatkan dari media dapat diklasifikasikan menjadi berita *hoax* dan berita fakta. Pengklasifikasian berita tersebut membutuhkan suatu metode atau algoritma agar tidak menggunakan cara manual dan menghabiskan waktu yang lama. Peranan informatika dibutuhkan dalam hal ini untuk membuat suatu model klasifikasi yang dapat mengkategorinya dua jenis berita tersebut. Penelitian mengenai klasifikasi berita *hoax* ini pernah dilakukan oleh beberapa peneliti sebelumnya seperti penelitian klasifikasi berita *clickbait* menggunakan *K-Nearest Neighbor* oleh (Sagita, Enri, dan Primajaya, 2020) yang menghasilkan akurasi terbaik 71% dengan parameter nilai $k=11$ pada skenario 80% data latih dan 20% data uji. Kemudian penelitian *detecting hoaxes in indonesian news using tf/tdm and k-nearest neighbor* oleh (Zuliarso et al, 2020) dengan akurasi yang dihasilkan yaitu 75% pada k value bernilai 4. Penelitian selanjutnya adalah analisis sentimen

terhadap ulasan pengguna MRT Jakarta menggunakan *Information Gain* dan *Modified K-Nearest Neighbor* dengan peningkatan akurasi 4-5% setelah menggunakan seleksi fitur *Information Gain* (Paramitha, Indriati, dan Sari, 2020).

Berdasarkan paparan penelitian yang sudah pernah dilakukan sebelumnya, pada penelitian kali ini penulis melakukan klasifikasi berita *hoax* dengan menggunakan metode *K-Nearest Neighbor* yang dikombinasikan dengan seleksi fitur *Information Gain*. Penulis berharap bahwa dengan menambahkan metode seleksi fitur *Information Gain* pada metode *K-Nearest Neighbor* ini dapat menghasilkan performa *precision*, *recall*, *F1-Score*, dan akurasi yang lebih baik dibandingkan penelitian sebelumnya.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan yang sudah dipaparkan di atas, maka dirumuskan permasalahan penelitian yaitu bagaimana pengaruh seleksi fitur *Information Gain* terhadap performa *precision*, *recall*, *F1-Score*, dan akurasi metode *K-Nearest Neighbor* dalam klasifikasi berita *hoax*?

1.3 Batasan Masalah

Agar pembahasan lebih terarah, berdasarkan rumusan masalah dan sesuai dengan batasan kemampuan penulis, maka ruang lingkup lingkungan pembahasan dibatasi pada:

- a. Pengklasifikasian berita menjadi dua kategori yaitu *hoax* dan fakta.
- b. Data yang digunakan adalah berita dalam Bahasa Indonesia.
- c. Domain data berita spesifik tentang Covid-19.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, tujuan penelitian ini yaitu untuk mengetahui pengaruh seleksi fitur *Information Gain* terhadap performa *precision*, *recall*, *F1-Score*, dan akurasi metode *K-Nearest Neighbor* dalam klasifikasi berita *hoax*.

1.5 Manfaat Penelitian

Manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut :

- a. Pemanfaatan teknologi untuk klasifikasi berita *hoax* secara otomatis.
- b. Hasil penelitian dapat dimanfaatkan bagi masyarakat untuk lebih mudah mengenali berita *hoax*.

1.6 Sistematika Penulisan

Sistematika penulisan bertujuan untuk mempermudah melihat dan mengetahui pembahasan yang dikerjakan secara menyeluruh. Adapun sistematika penulisan pada tugas akhir ini sebagai berikut:

1. **BAB I PENDAHULUAN**

Bab ini terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan skripsi.

2. **BAB II TINJAUAN PUSTAKA**

Bab ini meliputi tinjauan teori mengenai berita *hoax*, klasifikasi teks, *preprocessing*, TF-IDF, *K-Nearest Neighbor*, *Information Gain*, dan *confusion matrix*, serta tinjauan empiris penelitian terdahulu mengenai metode *K-Nearest Neighbor* dan seleksi fitur *Information Gain*.

3. **BAB III METODOLOGI PENELITIAN**

Bab ini mengemukakan pengumpulan data, alur penelitian, alur *preprocessing*, pembobotan TF-IDF, alur *K-Nearest Neighbor* dan *Information Gain*, serta desain evaluasi metode.

4. **BAB IV HASIL DAN PEMBAHASAN**

Bab ini menguraikan implementasi sistem, hasil pengujian metode, pengaruh seleksi fitur terhadap metode, dan analisis hasil yang diperoleh.

5. **BAB V KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan dan saran dari seluruh penelitian yang telah dilakukan.

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Teori

2.1.1 Berita Hoax

Menurut Kamus Besar Bahasa Indonesia, hoaks (bahasa Inggris: *hoax*) memiliki makna tidak benar atau bohong (tentang berita, pesan, dan sebagainya). Informasi atau berita hoax adalah berita palsu yang dibuat seolah-olah benar. *Hoax* adalah informasi palsu, usang atau tidak dapat diverifikasi yang secara spontan yang disebarkan oleh pengguna Internet. *Hoax* menyangkut subjek apapun yang cenderung memicu emosi positif atau negatif pada pengguna (Assiroj *et al*, 2018). Berita *hoax* dapat memicu pembacanya agar terpengaruh atau termanipulasi oleh hal-hal bertentangan yang terdapat dalam berita tersebut.

2.1.2 Klasifikasi Teks

Klasifikasi adalah suatu proses yang sangat sering terjadi dalam kehidupan sehari-hari. Pada dasarnya, klasifikasi melibatkan pembagian objek sehingga masing-masing ditetapkan ke salah satu dari sejumlah kategori yang dikenal sebagai kelas. Klasifikasi pada data tekstual yang juga disebut kategorisasi teks adalah suatu penugasan untuk secara otomatis menetapkan bagian teks ke satu atau lebih kelas yang telah ditentukan. Menurut definisi kategori, klasifikasi teks termasuk klasifikasi topik, klasifikasi genre, klasifikasi sentimen, deteksi spam, dan lain-lain (Zong, Xia dan Zhang, 2021).

2.1.3 Preprocessing

Preprocessing adalah tahapan pertama yang dilakukan terhadap data teks sebelum data tersebut diolah lebih lanjut. *Preprocessing* teks bertujuan untuk membuat dokumen masukan lebih konsisten untuk memfasilitasi representasi teks, yang diperlukan untuk sebagian besar tugas analisis teks (Aggarwal dan Zhai, 2012). Tahapan *preprocessing* adalah sebagai berikut :

a. *Case Folding*

Case Folding merupakan proses penyetaraan teks yang di dalamnya termasuk proses mengubah semua huruf menjadi huruf kecil (*lowercase*).

b. *Cleansing*

Proses *cleansing* adalah proses penghapusan seluruh karakter yang berupa html ataupun web yang tidak memiliki makna atau kaitan terhadap klasifikasi berita. Pada proses ini juga dilakukan proses penghapusan *punctuation* atau tanda baca, penghapusan *multiple space*, *tab*, *new line*, *back slice* dan penghapusan karakter yang tidak termasuk ke dalam ASCII.

c. Tokenisasi

Tokenisasi adalah proses pembagian atau pemisahan kata dalam suatu paragraf atau kalimat sehingga terbagi menjadi token-token tertentu.

d. *Stopword Removal*

Stopword Removal adalah proses penghapusan atau penghilangan kata-kata yang tidak berpengaruh dan tidak penting terhadap klasifikasi atau kategorisasi. Kata-kata dalam daftar *stopword* dianggap terlalu umum dan memiliki nilai yang tidak besar dalam teks (Aggarwal dan Zhai, 2012).

e. Normalisasi

Normalisasi dilakukan pada kata yang tidak baku untuk mengubah dan mengembalikannya kembali ke bentuk penulisan yang sesuai dengan Kamus Besar Bahasa Indonesia (KBBI).

f. *Stemming*

Proses *stemming* berfungsi agar kata-kata berimbuhan (awalan dan akhiran) dapat diekstraksi ke bentuk akarnya (*root extraction*) atau dapat dikatakan sebagai kata dasar (Aggarwal dan Zhai, 2012).

2.1.4 *Term Frequency Inverse Document Frequency (TF-IDF)*

Term Frequency Inverse Document Frequency (TF-IDF) adalah model statistik untuk mengevaluasi signifikansi kata-kata dalam kumpulan dokumen (Das,

Kamalanathan dan Alphonse, 2021). Data yang diproses akan diubah menjadi data numerik dengan metode pembobotan TF-IDF. Dalam metode TF-IDF ini menggabungkan dua konsep yaitu *Term Frequency* dan *Inverse Document Frequency*. *Term Frequency* merupakan frekuensi dari kemunculan sebuah *term* atau kata dalam sebuah dokumen sedangkan *Inverse Document Frequency* merupakan sebuah perhitungan dari bagaimana *term* atau kata didistribusikan secara luas pada koleksi dokumen yang bersangkutan. Kata-kata yang muncul di sebagian besar dokumen akan memiliki nilai IDF mendekati nol. Bobot kata TF-IDF adalah indikator kepentingan yang baik, dan mudah serta cepat untuk dihitung (Aggarwal dan Zhai, 2012). Adapun tahapan dari *Term Frequency Inverse Document Frequency* (TF-IDF) adalah sebagai berikut :

- a. Menghitung jumlah kemunculan *term* i dalam dokumen j ($tf_{i,j}$).
- b. Menghitung jumlah dokumen yang mengandung *term* i (df_i)
- c. Menghitung nilai bobot *inverse document frequency* (idf) dengan menggunakan persamaan :

$$idf_i = \log \left(\frac{N}{df_i} \right) \quad (2.1)$$

Keterangan :

N = jumlah dokumen secara keseluruhan

- d. Menghitung nilai bobot TF-IDF dengan menggunakan persamaan :

$$w_{i,j} = tf_{i,j} \times idf_i \quad (2.2)$$

Keterangan :

$w_{i,j}$ = bobot *term* i terhadap dokumen j

$tf_{i,j}$ = frekuensi *term* i pada dokumen j

idf_i = nilai bobot IDF *term* i

2.1.5 K-Nearest Neighbor

K-Nearest Neighbor (KNN) adalah salah satu algoritma klasifikasi serbaguna dan terkenal yang sering digunakan dalam *data mining* dan *text mining*.

KNN memiliki waktu kalkulasi yang lebih rendah dibandingkan dengan algoritma lain seperti Random Forest, yang berguna untuk deteksi otomatis yang cepat (Zuliarso et al, 2020). KNN merupakan algoritma sederhana yang setiap pengamatannya diprediksi berdasarkan kemiripannya dengan pengamatan lainnya (Boehmke dan Greenwell, 2019). KNN memberikan keanggotaan kelas ke data berdasarkan mayoritas tetangganya, dengan objek yang ditetapkan ke kelas yang paling umum di antara k tetangga terdekatnya (k adalah bilangan bulat positif yang biasanya bernilai kecil). Adapun tahapan dari *K-Nearest Neighbor* adalah sebagai berikut:

- a. Menentukan jumlah tetangga k yang akan digunakan.
- b. Melakukan perhitungan jarak objek terhadap masing-masing data kelompok menggunakan rumus persamaan *Euclidean distance* di bawah:

$$D(x_a, x_b) = \sqrt{\sum_{j=1}^P (x_{aj} - x_{bj})^2} \quad (2.3)$$

Keterangan :

D = jarak

x_a = data uji (*data testing*)

x_b = data latih (*data training*)

- c. Mendapatkan hasil pengklasifikasian

2.1.6 Information Gain

Information gain adalah salah satu metode seleksi fitur yang digunakan dalam klasifikasi teks. *Information gain* melakukan perhitungan mengenai pengaruh suatu fitur terhadap keseragaman kelas pada data. Data tersebut dipecah menjadi sub data dengan nilai fitur tertentu. Semakin besar nilai perolehan *information gain* suatu fitur, maka semakin besar pengaruh atau kekuatan dari fitur tersebut (Aggarwal dan Zhai, 2012). Persamaan dari *Information Gain* adalah sebagai berikut (Aggarwal, 2018) :

$$\text{Information Gain } I(t_j) \text{ of } t_j = -\sum_{r=1}^k \frac{n_r}{n} \log \left(\frac{n_r}{n} \right) - E(t_j) \quad (2.4)$$

Dimana $-\sum_{r=1}^k \frac{n_r}{n} \log \left(\frac{n_r}{n} \right)$ adalah *entropy* pada distribusi kelas, dan $E(t_j)$ adalah *entropy* bersyarat yang dihitung dengan persamaan :

$$E(t_j) = -\sum_{r=1}^k \left\{ \left[\frac{n(t_j)}{n} \right] P(c_r|t_j) \cdot \log[P(c_r|t_j)] + \left[\frac{n-n(t_j)}{n} \right] P(c_r|\neg t_j) \cdot \log[P(c_r|\neg t_j)] \right\} \quad (2.5)$$

Keterangan :

n_r : jumlah total dokumen dengan kelas r.

n : jumlah total dokumen.

$n(t_j)$: banyaknya dokumen yang mengandung term t_j dari korpus berukuran $n \geq n(t_j)$.

$P(c_r|t_j)$: peluang *term* t_j terdapat pada kelas r dalam dokumen.

$P(c_r|\neg t_j)$: peluang *term* t_j tidak terdapat pada kelas r dalam dokumen.

2.1.7 Confusion Matrix

Confusion Matrix digunakan untuk menghitung akurasi, *recall*, *precision*, dan *error rate*. *Confusion matrix* dapat digunakan untuk mengevaluasi kualitas *classifier*. Pada *confusion matrix* dua kelas, matriks menunjukkan *true positives*, *true negatives*, *false positives*, dan *false negatives*. *Confusion matrix* untuk dua kelas ditunjukkan pada Tabel 2.1 (Han *et al.*, 2012).

Tabel 2.1 *Confusion Matrix*

Kelas Sebenarnya	Prediksi Kelas	
	Ya	Tidak
Ya	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
Tidak	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

a. *True Positive* = Jumlah prediksi benar dari data positif

b. *False Positive* = Jumlah prediksi salah dari data negatif

c. *False Negative* = Jumlah prediksi salah dari data positif

d. *True Negative* = Jumlah prediksi benar dari data negatif

Efektivitas atau akurasi sistem dapat diketahui dengan menghitung *precision*, *recall*, dan *F1-Score*. *Precision* mengukur presentase dokumen bernilai positif benar di antara seluruh dokumen yang diidentifikasi positif (Aggarwal dan Zhai, 2012). *Precision* dapat dihitung menggunakan persamaan:

$$Precision = \frac{TP}{(TP+FP)} \quad (2.6)$$

Recall mengukur presentase dokumen bernilai positif benar yang dapat diidentifikasikan di antara seluruh dokumen yang berlabel positif. *Recall* dapat dihitung menggunakan persamaan:

$$Recall = \frac{TP}{(TP+FN)} \quad (2.7)$$

F1-Score adalah rata-rata geometris dari *precision* dan *recall* (Aggarwal dan Zhai, 2012). *F1-Score* dapat dihitung menggunakan persamaan:

$$F1-Score = \frac{2 \times recall \times precision}{(recall+precision)} \quad (2.8)$$

Akurasi adalah kedekatan nilai aktual atau nilai sebenarnya dengan nilai yang diprediksi. Akurasi dapat dihitung menggunakan persamaan:

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (2.9)$$

2.2 Tinjauan Empiris

2.2.1 Detecting Hoaxes in Indonesian News using TF/TDM and K-Nearest Neighbor (Zuliarso et al, 2020)

Penelitian ini membahas tentang implementasi metode atau algoritma dalam mendeteksi *hoax* pada berita berbahasa Indonesia. Data yang digunakan adalah 74 artikel berita *hoax* dan 74 artikel berita asli berbahasa Indonesia. Metode yang digunakan adalah metode *K-Nearest Neighbor* dan TF/TDM. Setelah dilakukan pengujian akurasi terhadap metode tersebut, didapatkan hasil performa akurasi tertinggi 83,6% pada saat parameter nilai *k* adalah 1. Penelitian Zuliarso et al (2020) mengimplementasikan metode *K-Nearest Neighbor* tanpa melakukan proses seleksi fitur. Perbedaan penelitian penulis dengan penelitian yang dilakukan

oleh Zuliarso et al (2020) adalah *dataset* berita yang dikumpulkan, serta penerapan metode seleksi fitur *Information Gain* terhadap algoritma *K-Nearest Neighbor* yang diharapkan dapat meningkatkan performa akurasi.

2.2.2 The Effect of Information Gain Feature Selection for Hoax Identification in Twitter Using Classification Method Support Vector Machine (Mubaroq & Setiawan, 2020)

Penelitian ini membahas tentang klasifikasi berita *hoax* yang beredar di media sosial Twitter. Data yang digunakan adalah 25329 *tweet hoax* dan 26029 *tweet non-hoax*. Algoritma atau metode yang digunakan pada penelitian ini adalah *Support Vector Machine* dengan seleksi fitur *Information Gain*. Setelah dilakukan pengujian akurasi terhadap metode tersebut, didapatkan hasil performa akurasi sebesar 70,21% tanpa *Information Gain* dan 95,66% dengan *Information Gain*. Penelitian Mubaroq & Setiawan (2020) mengimplementasikan metode *Support Vector Machine* dengan *Information Gain* sebagai seleksi fitur. Perbedaan penelitian penulis dengan penelitian yang dilakukan oleh Mubaroq & Setiawan (2020) adalah *dataset* berita yang dikumpulkan, serta penerapan metode seleksi fitur *Information Gain* terhadap metode yang berbeda yaitu *K-Nearest Neighbor* yang diharapkan dapat meningkatkan performa akurasi.

2.2.3 Analisis Sentimen Terhadap Ulasan Pengguna MRT Jakarta Menggunakan Information Gain dan Modified K-Nearest Neighbor (Paramitha, Indriati, dan Sari, 2020)

Penelitian ini membedakan pendapat pengguna yang berupa pendapat positif dan negatif berbasis analisis sentimen. Data yang digunakan adalah ulasan pengguna melalui media sosial mengenai MRT Jakarta sejumlah 650 ulasan dari media sosial yang dilabeli oleh 4 pakar. Metode yang digunakan pada penelitian ini adalah metode *Modified K-Nearest Neighbor* yang dikombinasikan dengan seleksi fitur *Information Gain*. Setelah dilakukan pengujian akurasi terhadap metode tersebut, didapatkan hasil akurasi tertinggi yaitu 86,77% dengan parameter nilai k adalah 3 dengan nilai *threshold* seleksi fitur adalah 25%. Akurasi yang dihasilkan

dengan menggunakan seleksi fitur *Information Gain* meningkat sebesar 4-5% dari akurasi metode tanpa melibatkan seleksi fitur. Penelitian yang dilakukan oleh Paramitha, Indriati, dan Sari (2020) menggunakan kombinasi metode *Modified K-Nearest Neighbor* dan seleksi fitur *Information Gain*. Perbedaan penelitian penulis dengan penelitian yang dilakukan oleh Paramitha, Indriati, dan Sari (2020) adalah data yang digunakan penulis berupa artikel atau berita, serta metode yang digunakan penulis adalah *K-Nearest Neighbor* dengan kombinasi seleksi fitur *Information Gain* yang diharapkan dapat meningkatkan performa akurasi.

2.2.4 Pemanfaatan Klasifikasi Soal Biologi *Cognitive Domain Bloom's Taxonomy* Menggunakan KNN *Chi-Square* Sebagai Penyusunan Naskah Soal (Listiowarni & Dewi, 2020)

Penelitian ini mengklasifikasikan naskah soal biologi ke dalam 6 kelas ranah *Cognitive Taxonomy Bloom* yaitu pemahaman, aplikasi, analisis, sintesis, dan evaluasi. Data yang digunakan adalah 600 soal uraian biologi yang sudah dilabeli dengan 6 kelas tersebut. Metode yang digunakan adalah kombinasi *K-Nearest Neighbor* dengan metode seleksi fitur *Chi-Square*. Setelah dilakukan pengujian akurasi terhadap metode tersebut, didapatkan hasil akurasi tertinggi yaitu 79,36% dengan parameter nilai k adalah 5 dan *threshold* seleksi fitur *chi-square* adalah 0,5. Penelitian yang dilakukan oleh Listiowarni & Dewi (2020) menggunakan kombinasi metode *K-Nearest Neighbor* dan seleksi fitur *Chi-Square*. Perbedaan penelitian penulis dengan penelitian yang dilakukan oleh Listiowarni & Dewi (2020) adalah data yang digunakan penulis berupa artikel atau berita, serta penulis mengkombinasikan metode *K-Nearest Neighbor* dengan seleksi fitur *Information Gain*.

2.2.5 Klasifikasi Berita *Clickbait* Menggunakan *K-Nearest Neighbor* (Sagita, Enri, dan Primajaya, 2020)

Penelitian ini mengklasifikasikan berita ke dalam 2 kategori yaitu *clickbait* dan bukan *clickbait*. Data yang digunakan adalah 1000 yang dilabeli secara manual. Metode yang digunakan adalah metode *K-Nearest Neighbor*. Setelah dilakukan

pengujian akurasi terhadap metode tersebut, didapatkan hasil akurasi terbaik yaitu 71% dengan parameter nilai k adalah 11 pada skenario 80% data latih dan 20% data uji. Penelitian yang dilakukan oleh Sagita, Enri, dan Primajaya (2020) menggunakan metode *K-Nearest Neighbor* tanpa melakukan proses seleksi fitur. Perbedaan penelitian penulis dengan penelitian yang dilakukan oleh Sagita, Enri, dan Primajaya (2020) adalah data yang digunakan penulis berupa artikel atau berita *hoax* dan fakta, serta penulis mengkombinasikan metode *K-Nearest Neighbor* dengan seleksi fitur *Information Gain*.

BAB III

METODOLOGI PENELITIAN

3.1 Data dan Metode Pengumpulan Data

3.1.1 Data Penelitian

Data yang digunakan pada penelitian ini adalah data dalam bentuk berita yang berkaitan dengan Covid-19. Bagian berita yang digunakan adalah isi berita. Data berjumlah 300 dengan format *file* *.csv yang meliputi 150 berita *hoax* dan 150 berita fakta yang bersumber dari media internet. Seluruh data sudah dilabeli oleh lembaga media internet tersebut. Data berita kemudian dibagi menjadi dua yaitu data latih dan data uji, dengan presentase 80% data latih dan 20% data uji. Data latih tersebut kemudian dibagi lagi menjadi data latih dan data validasi untuk digunakan dalam proses pelatihan model dengan menggunakan *N-fold cross validation*.

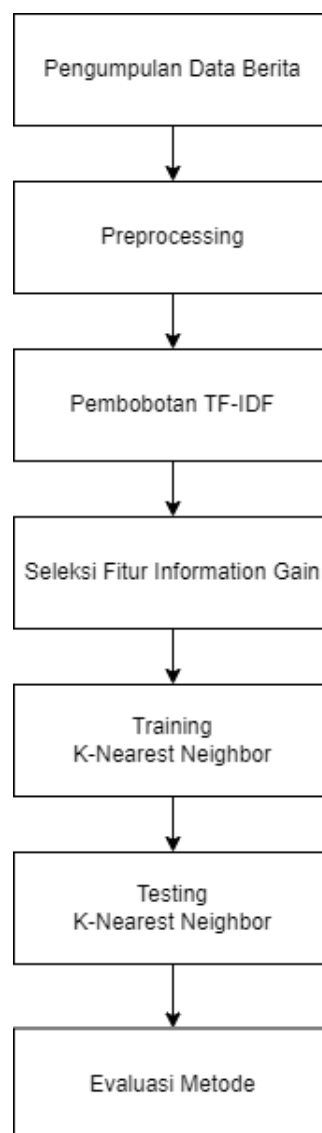
3.1.2 Metode Pengumpulan Data

Data pada penelitian ini adalah data sekunder. Data berita *hoax* bersumber dari <https://cekfakta.com>, sedangkan data berita fakta bersumber dari <https://detik.com>. Data berita didapatkan dengan menggunakan proses *web scrapping* dengan menggunakan bantuan *library* BeautifulSoup4 dengan Bahasa pemrograman Python. Hasil berita yang berhasil didapatkan disimpan dengan format *.csv lalu ditambahkan label sesuai dengan kategori berita tersebut, yaitu *hoax* dan fakta.

3.2 Alur Penelitian

Pada bagian ini menggambarkan alur penelitian yang dilakukan oleh penulis. Langkah pertama yang dilakukan adalah melakukan proses pengumpulan data yaitu berita *hoax* dan fakta. Setelah data didapatkan, maka proses selanjutnya adalah *preprocessing* untuk membersihkan data dari *noise* dan mempersiapkan data sebelum diolah lebih lanjut, seperti yang dijelaskan pada Gambar 3.2. Kemudian, data melalui tahap pembobotan menggunakan *Term Frequency-Inverse Document*

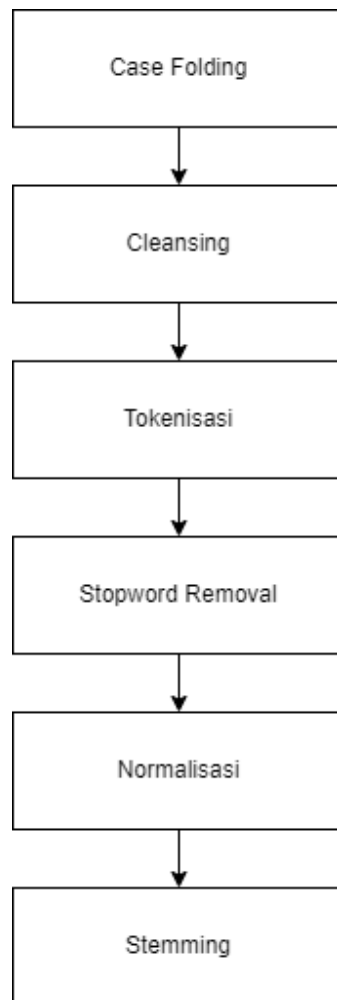
Frequency (TF-IDF) yang dijelaskan pada Gambar 3.3. Data yang sudah diberi bobot tadi kemudian melalui tahap seleksi fitur *Information Gain* untuk memilih fitur-fitur yang memiliki nilai *Information Gain* terbaik, seperti yang dijelaskan pada Gambar 3.4. Setelah seleksi fitur, data melalui tahap *training* dan *testing* klasifikasi menggunakan metode *K-Nearest Neighbor* seperti pada Gambar 3.5. Tahap terakhir adalah evaluasi performa metode dengan ukuran evaluasi yang sudah ditentukan. Alur penelitian dapat dilihat pada Gambar 3.1.



Gambar 3.1 Alur Penelitian

3.3 *Preprocessing*

Sebelum melakukan tahap pembobotan, data terlebih dahulu melalui tahapan *preprocessing*. *Preprocessing* adalah proses dilakukan untuk mempersiapkan data-data teks yang belum terstruktur menjadi data terstruktur dan siap digunakan untuk proses klasifikasi. Pada tahap ini, terdapat beberapa proses. Proses pertama adalah *case folding*, yaitu proses mengubah semua huruf menjadi huruf kecil. Proses ini dilakukan dengan menggunakan *function* yang disediakan oleh Python. Kemudian *cleansing*, pada proses ini dilakukan penghapusan karakter yang tidak memiliki kaitan terhadap klasifikasi berita, termasuk penghapusan tanda baca. Pada proses ini digunakan *library string* dan implementasi *regular expression*. Selanjutnya adalah tokenisasi, yaitu pemisahan kata dalam suatu paragraf atau kalimat menjadi token-token tertentu. Pada proses ini digunakan *library NLTK (Natural Language Toolkit)*. *Library* yang sama juga digunakan pada proses *stopword removal*, yaitu penghapusan kata-kata yang tidak berpengaruh terhadap klasifikasi berita. Pada proses ini, terdapat tambahan korpus *stopword* yang bersumber dari kompilasi *stopword* oleh Diaz (2016). Kemudian proses normalisasi, yaitu mengubah dan mengembalikan bentuk penulisan tidak baku ke bentuk penulisan yang sesuai dengan KBBI. Pada proses ini digunakan korpus yang berisi kumpulan kata tidak baku dan bentuk baku dari kata tersebut. Korpus ini bersumber dari *dataset* pada penelitian yang dilakukan oleh Salsabila *et al* (2018). Proses terakhir adalah *stemming*, yaitu proses ekstraksi kata-kata berimbuhan menjadi kata dasar. Pada proses ini digunakan *library PySastrawi*. Seluruh proses *preprocessing* ditunjukkan pada Gambar 3.2.

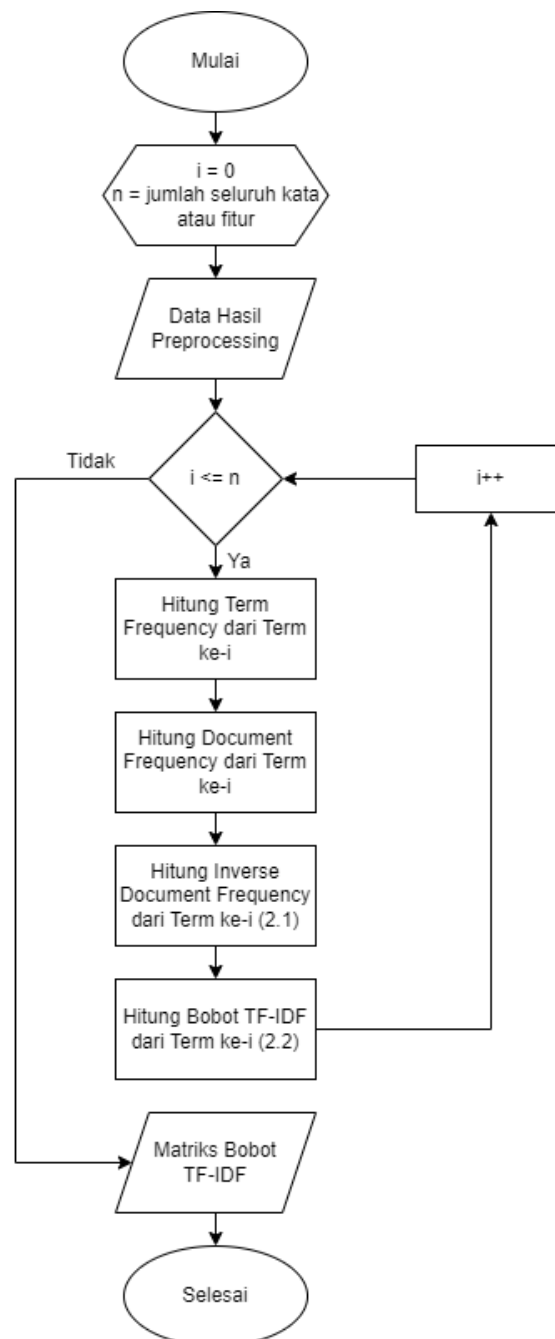


Gambar 3.2 Alur *Preprocessing*

3.4 Pembobotan TF-IDF

Data hasil *preprocessing* kemudian melalui tahap pembobotan dengan metode *Term Frequency Inverse Document Frequency*. Pembobotan TF-IDF diawali dengan input data hasil *preprocessing*. Kemudian, dilakukan perulangan hingga jumlah seluruh kata. Proses selanjutnya adalah perhitungan *term frequency* yaitu menghitung frekuensi kemunculan kata pada suatu dokumen. Kemudian dilanjutkan dengan perhitungan *document frequency* yaitu menghitung jumlah dokumen yang mengandung suatu kata tertentu. Kemudian proses dilanjutkan dengan perhitungan bobot *inverse document frequency*, yaitu menghitung distribusi kata pada koleksi dokumen dengan menggunakan persamaan (2.1). Terakhir, proses

perhitungan bobot TF-IDF dilakukan dengan mengalikan nilai *term frequency* dan nilai *inverse document frequency* dengan menggunakan persamaan (2.2).

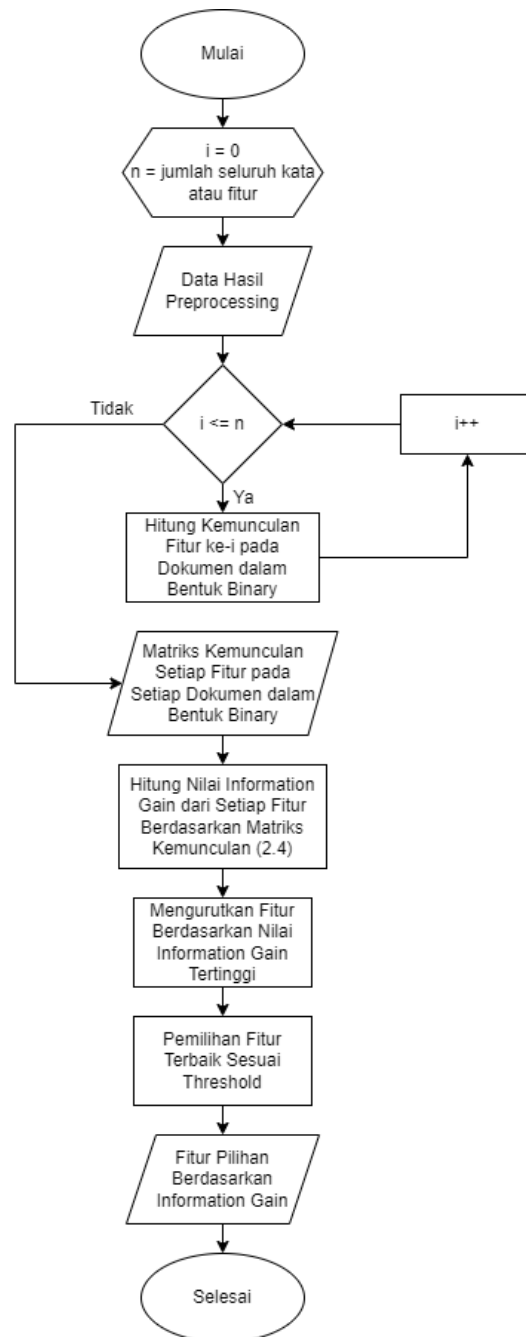


Gambar 3.3 Pembobotan TF-IDF

3.5 Seleksi Fitur *Information Gain*

Seleksi fitur digunakan untuk mengurangi fitur-fitur yang tidak memiliki pengaruh besar dalam tahap klasifikasi. Pada seleksi fitur ini menggunakan seleksi fitur *Information Gain* seperti yang ditunjukkan pada Gambar 3.4. Berdasarkan tinjauan empiris, *threshold* yang menghasilkan akurasi tertinggi adalah 25%. Namun pada penelitian ini terdapat perubahan nilai *threshold* dengan tujuan untuk menghasilkan performa terbaik terhadap data digunakan.

Input dari metode seleksi fitur *Information Gain* ini adalah label dan isi data berita yang sudah dilakukan *preprocessing*. Kemudian, dilakukan perulangan hingga jumlah seluruh kata. Pada setiap kata atau fitur, kemunculan fitur pada setiap dokumen dihitung dalam bentuk *binary*, 0 jika fitur tersebut tidak muncul pada dokumen, dan 1 jika fitur tersebut muncul dalam dokumen. Setelah mendapatkan seluruh nilai kemunculan fitur, maka akan didapatkan matriks kemunculan fitur dalam bentuk *binary*. Berdasarkan matriks tersebut, dilakukan proses perhitungan *Information Gain* dari setiap fitur dengan menggunakan persamaan (2.4) untuk mengetahui fitur-fitur yang memiliki pengaruh penting pada suatu dokumen berita. Proses selanjutnya adalah pengurutan nilai-nilai *Information Gain* tersebut dari nilai tertinggi hingga terendah agar dapat diketahui fitur-fitur terbaik. Fitur-fitur terbaik tersebut dipilih dan digunakan karena fitur tersebut adalah yang paling berpengaruh terhadap proses klasifikasi berdasarkan nilai-nilai yang memenuhi batasan *threshold* yang telah diberikan.

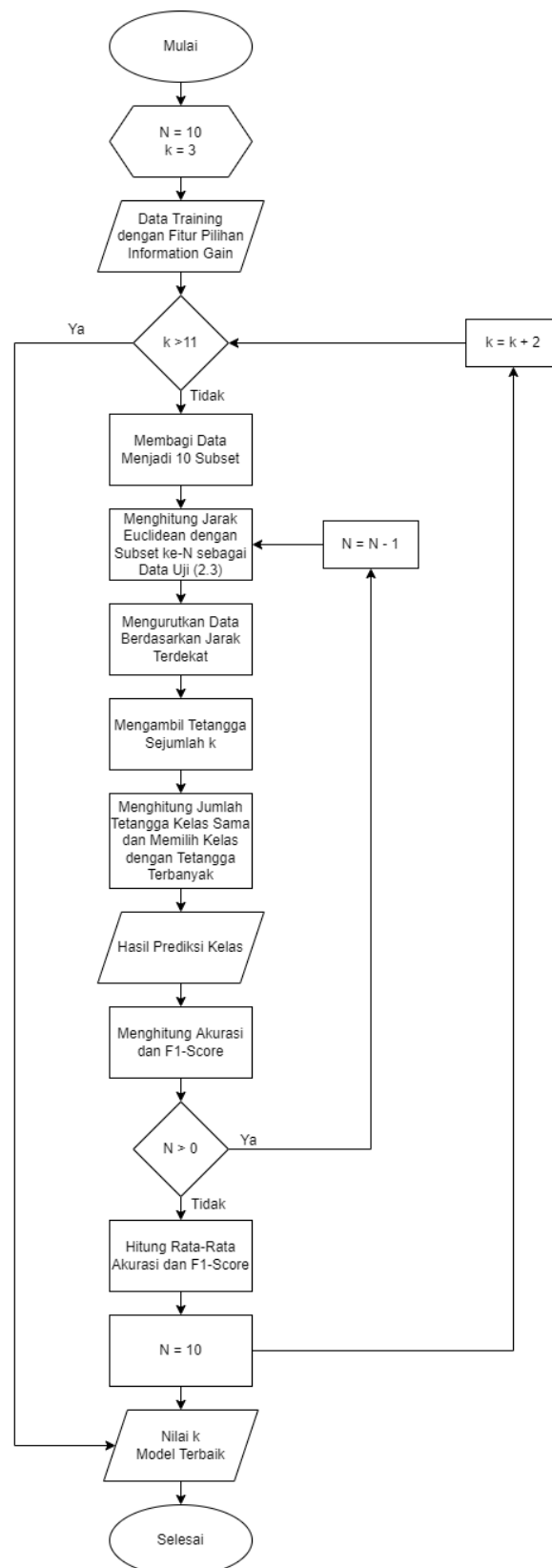
Gambar 3.4 Seleksi Fitur *Information Gain*

3.6 *K-Nearest Neighbor*

Pada tahap klasifikasi, digunakan metode *K-Nearest Neighbor* untuk menentukan kategori dari data berita yaitu *hoax* dan fakta. Berdasarkan salah satu tinjauan empiris, nilai k yang menghasilkan akurasi tertinggi adalah $k = 5$. Namun

pada penelitian ini dilakukan perubahan parameter nilai k dengan tujuan untuk menghasilkan performa terbaik terhadap data yang digunakan. Input dari metode *K-Nearest Neighbor* ini adalah fitur hasil pembobotan TF-IDF yang kemudian telah dipilih fitur terbaiknya berdasarkan seleksi fitur *Information Gain*.

Pada proses *training* dan validasi, tahap pertama yang dilakukan adalah menginisialisasikan nilai N dari *N-Fold Cross Validation* dan nilai k yang akan diuji. Lalu, proses selanjutnya adalah membagi data *training* menjadi 10 Subset. Kemudian, menghitung jarak *Euclidean* dengan menggunakan persamaan (2.3) dengan subset ke- N sebagai data uji untuk menemukan jarak mana yang paling terdekat. Kemudian, tahapan selanjutnya yaitu jarak-jarak tersebut diurutkan berdasarkan jarak terdekat. Tahap berikutnya adalah mengambil tetangga sejumlah k yang sudah ditentukan sebelumnya pada tahap pertama. Selanjutnya sejumlah k tetangga terdekat tersebut dihitung jumlah tetangga dengan kategori yang sama (*hoax* atau fakta) dan menentukan kategori dengan tetangga terbanyak. Selanjutnya, kategori dengan tetangga terbanyak tersebut akan menjadi prediksi kategori dari data yang diuji dengan metode *K-Nearest Neighbor*. Hasil dari klasifikasi ini berupa prediksi kategori dari berita yaitu *hoax* atau fakta. Performa akurasi dan *F1-Score* kemudian dihitung pada setiap iterasi. Setelah iterasi ke-10 selesai, dihitung rata-rata akurasi dan *F1-Score* dari 10 iterasi yang sudah dihitung sebelumnya. Ketika seluruh eksperimen nilai k telah dilakukan, nilai k dengan performa *F1-Score* tertinggi dipilih menjadi model terbaik untuk kemudian diuji kembali dengan data baru.



Gambar 3.5 Training K-Nearest Neighbor



Gambar 3.6 *Testing K-Nearest Neighbor*

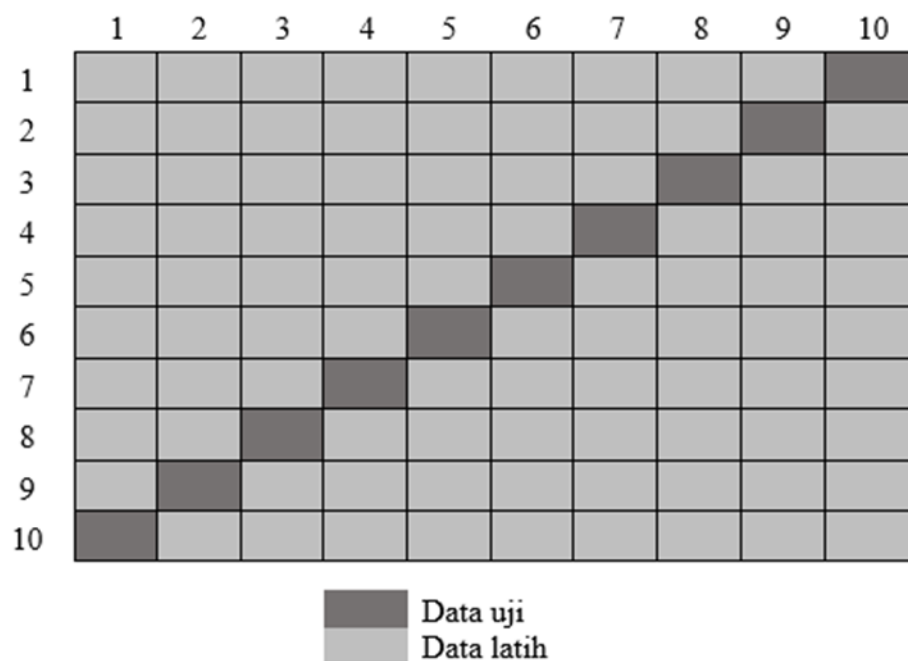
3.7 Desain Evaluasi Metode

3.7.1 Skenario Pengujian

Pengujian sistem berfungsi untuk mengetahui kinerja dari sistem dalam melakukan tugasnya yaitu mengklasifikasikan berita *hoax* dan fakta. Pengujian ini dilakukan untuk memilih model terbaik melalui proses pelatihan dan validasi.

Pada tahap pelatihan dan validasi, digunakan *N-Fold Cross Validation* dengan nilai $N = 10$. *N-Fold Cross Validation* digunakan untuk mencari performa

terbaik dari model dengan membagi dataset dan menambah variasi data, sehingga berpengaruh pada performa model dengan data yang berjumlah sedikit. Pada tahap pertama, diinisialisasikan jumlah iterasi yaitu 10. Dataset dibagi menjadi 10 bagian, yang kemudian dilakukan iterasi sebanyak 10 kali. Pada setiap iterasi, pengujian menggunakan satu subset sebagai data uji, dan 9 subset lainnya sebagai data latih. Pada setiap iterasi, dihitung rata-rata performa *F1-Score* dan akurasi dengan menggunakan persamaan (2.8) dan (2.9).



Gambar 3.7 *N-Fold Cross Validation*

Pengujian dilakukan pada metode *K-Nearest Neighbor* dan seleksi fitur *Information Gain* dengan menggunakan data yang sudah disiapkan. Adapun skenario pengujian yang dilakukan pada penelitian ini adalah sebagai berikut:

a. Pengujian Metode *K-Nearest Neighbor* tanpa Seleksi Fitur

Skenario pertama adalah menguji metode *K-Nearest Neighbor* dan memilih model terbaik tanpa mengimplementasikan seleksi fitur. Pada skenario ini, tahap pelatihan dan validasi dilakukan dengan menggunakan *N-Fold Cross Validation*. Perubahan nilai k dilakukan pada metode *K-Nearest Neighbor*. Nilai k yang diuji adalah nilai $k=3$, $k=5$, $k=7$, $k=9$, dan $k=11$. Nilai k yang menghasilkan performa

F1-Score tertinggi dipilih sebagai model terbaik yang akan digunakan pada proses *testing* data baru. Nilai *k* yang menghasilkan *F1-Score* tertinggi memiliki makna bahwa hasil prediksi berita *hoax* akan lebih akurat kebenarannya.

b. Pengujian Metode *K-Nearest Neighbor* dengan Seleksi Fitur *Information Gain*

Skenario selanjutnya adalah menguji metode *K-Nearest Neighbor* dan memilih model terbaik dengan menggunakan seleksi fitur *Information Gain*. Pada skenario ini, tahap pelatihan dan validasi dilakukan dengan menggunakan *N-Fold Cross Validation*. Pengujian terhadap *threshold* dilakukan pada seleksi fitur *Information Gain*. *Threshold* yang digunakan adalah 50%, 25%, 20%, 10%, 5%, 2%, 1%, 0.5%, 0.2%, dan 0.1%. *Threshold* digunakan untuk menentukan jumlah fitur dengan nilai *Information Gain* tertinggi yang digunakan pada proses klasifikasi. Pada setiap skenario perubahan *threshold*, perubahan nilai *k* juga dilakukan dengan menggunakan nilai *k*=3, *k*=5, *k*=7, *k*=9, dan *k*=11. Kombinasi *threshold* dan nilai *k* yang menghasilkan performa *F1-Score* tertinggi dipilih sebagai model terbaik yang akan digunakan pada proses *testing* data baru. Kombinasi yang menghasilkan *F1-Score* tertinggi memiliki makna bahwa hasil prediksi berita *hoax* akan lebih akurat kebenarannya.

c. Pengujian Model Terbaik

Skenario pengujian berikutnya adalah menguji dua model terbaik yang sudah dipilih sebelumnya, yaitu model *K-Nearest Neighbor* dan model *K-Nearest Neighbor* dengan kombinasi seleksi fitur *Information Gain*. Pengujian dilakukan dengan menggunakan data baru yang belum pernah melalui tahap pelatihan dan validasi. Kemudian kedua model dibandingkan untuk mengetahui pengaruh seleksi fitur *Information Gain* terhadap performa metode *K-Nearest Neighbor*.

3.7.2 Evaluasi Metode

Evaluasi ditujukan untuk mengukur performa model terbaik yang sudah dipilih dari proses validasi. Ukuran evaluasi yang digunakan adalah *precision*, *recall*, *F1-Score*, dan akurasi. *Precision* mengukur presentase dokumen bernilai

positif benar di antara seluruh dokumen yang diidentifikasi positif. *Precision* dihitung dengan menggunakan persamaan (2.6). *Recall* mengukur presentase dokumen bernilai positif benar yang dapat diidentifikasikan di antara seluruh dokumen yang relevan. *Recall* dihitung menggunakan persamaan (2.7). *F1-Score* adalah kombinasi hasil dari *precision* dan *recall*. *F1-Score* dapat dihitung menggunakan persamaan (2.8). Akurasi adalah kedekatan nilai aktual atau nilai sebenarnya dengan nilai yang diprediksi. Akurasi dapat dihitung menggunakan persamaan (2.9).

BAB IV

HASIL DAN PEMBAHASAN

4.1 Implementasi Sistem

Sistem diimplementasikan menggunakan bahasa pemrograman Python dengan versi 3.7.12 serta menggunakan Google Colaboratory Notebook. Beberapa perangkat digunakan sebagai alat bantu dalam penelitian ini, yaitu:

a. Hardware

- AMD Ryzen 3 2200U 2.5Ghz
- Memory 4096MB RAM
- SSD 512GB

b. Software

- Google Colaboratory Notebook (Python versi 3.7.12)

4.2 Implementasi Memuat dan Membaca Dataset

Sesuai dengan alur penelitian pada Gambar 3.1, tahap pertama adalah *input dataset*. Pada tahap ini, data *training* dimuat ke dalam sistem untuk kemudian digunakan pada tahap *preprocessing*. Library Pandas digunakan untuk membantu membaca dataset dalam bentuk *dataframe*. Pada baris ke-1, dataset dibaca dalam bentuk *dataframe* dan dimuat ke dalam variabel `df`. Adapun isi dari *dataframe* `df` adalah tabel dengan “Label” dan “Content” sebagai kolom, dan “index” data sebagai baris.

Tabel 4.1 Penggalan Kode Memuat Dokumen

No	Penggalan Kode
1	<code>df = pd.read_csv('data_training.csv')</code>

	Label	Content
0	0	Jakarta - Ketua DPRD Kabupaten Bogor Rudy Susm...
1	0	Jakarta - Vaksin 1 dan 2 Sinovac, vaksin 3 apa...
2	0	Bandung - Pemkot Bandung menyiapkan 226 tempat...
3	0	Makassar - Kasus aktif COVID-19 di Kota Makass...
4	0	Kabupaten Bandung - Polisi menerapkan ganjil g...

Gambar 4.1 Tampilan *Data Training*

4.3 Implementasi *Preprocessing*

Pada tahap ini, data yang sudah dimuat diproses melalui 6 tahapan *preprocessing* sesuai dengan alur pada Gambar 3.2, yaitu *case folding*, *cleansing*, tokenisasi, *stopword removal*, normalisasi, dan *stemming*. Pada tahap ini digunakan beberapa *library* bantuan yaitu NLTK, *string*, *regular expression*, dan PySastrawi.

4.3.1 Implementasi *Case Folding*

Pada tahap *case folding* dilakukan proses penyetaraan teks dengan mengubah semua huruf menjadi huruf kecil (*lowercase*). Penggalan kode dapat dilihat pada tabel 4.2.

Tabel 4.2 Penggalan Kode Proses *Case Folding*

No	Penggalan Kode
1	<code>df['Content'] = df['Content'].str.lower()</code>

4.3.2 Implementasi *Cleansing*

Pada tahap *cleansing* dilakukan proses penghapusan karakter yang tidak memiliki makna atau kaitan terhadap klasifikasi berita. Baris ke-1 hingga baris ke-5 merupakan proses pembersihan *noise*. Baris ke-2 menghapus *tab*, baris baru, dan *backslash*. Baris ke-3 menghapus seluruh karakter yang tidak termasuk ke dalam karakter ASCII, baris ke-4 menghapus *link*, *mention*, dan tanda pagar, baris ke-5 menghapus angka, dan baris ke-6 mengembalikan data yang sudah dibersihkan. Baris ke-7 hingga baris ke-9 adalah tahap penghapusan tanda baca pada data. Baris ke-10 dan baris ke-11 adalah proses menghapus spasi ganda.

Tabel 4.3 Penggalan Kode Proses *Cleansing*

No	Penggalan Kode
1	<code>def remove_noise(text):</code>
2	<code>text = text.replace('\t', " ").replace('\n', " ").replace</code>
	<code>('\\', "")</code>
3	<code>text = text.encode('ascii', 'replace').decode('ascii')</code>
4	<code>text = ' '.join(re.sub("([@#][A-Za-z0-</code>
	<code>9])+ (\w+:\//\S+)", " ", text).split())</code>
5	<code>text= re.sub('[0-9]+', '', text)</code>
6	<code>return text</code>
7	<code>def remove_punctuation(text):</code>
8	<code>transtable = str.maketrans(string.punctuation, ' '*len</code>
	<code>(string.punctuation))</code>
9	<code>return text.translate(transtable)</code>
10	<code>def remove_multiple_space(text):</code>
11	<code>return re.sub('\s+', ' ', text)</code>

4.3.3 Implementasi Tokenisasi

Pada tahap ini dilakuakn proses pembagian atau pemisahan kata dalam setiap kalimat pada data sehingga terbagi menjadi token-token atau fitur-fitur tertentu. Baris ke-1 dan ke-2 melakukan proses tokenisasi dan mengembalikan sejumlah fitur sesuai dengan data yang diinputkan.

Tabel 4.4 Penggalan Kode Proses Tokenisasi

No	Penggalan Kode
1	<code>def word_tokenization(text):</code>
2	<code> return word_tokenize(text)</code>

4.3.4 Implementasi *Stopword Removal*

Pada tahap ini dilakukan penghapusan atau penghilangan kata-kata yang tidak berpengaruh dan tidak penting terhadap proses klasifikasi. Baris ke-1 memuat kata-kata yang termasuk ke dalam korpus *stopword* Bahasa Indonesia pada *library* NLTK. Daftar kata-kata tersebut dimuat ke dalam variabel `list_sw`. Kemudian pada baris ke-2 terdapat penambahan korpus *stopword* yaitu *file* dengan format **.txt* yang dimuat ke dalam variabel `txt_sw`. Pada baris ke-3 dan ke-4, daftar kata pada korpus *stopword* ditambahkan dengan daftar kata yang ada pada *file*. Pada baris ke-5 dan baris ke-6 mengembalikan kata-kata yang tidak termasuk ke dalam daftar *stopword*.

Tabel 4.5 Penggalan Kode Proses *Stopword Removal*

No	Penggalan Kode
1	<code>list_sw = stopwords.words('indonesian')</code>
2	<code>txt_sw = pd.read_csv("stopwords-id.txt", names=</code> <code> ["stopwords"], header = None)</code>
3	<code>list_sw.extend(txt_sw["stopwords"][0].split(' '))</code>
4	<code>list_sw = set(list_sw)</code>
5	<code>def stopword_removal(words):</code>
6	<code> return [word for word in words if word not in list_sw]</code>

4.3.5 Implementasi Normalisasi

Proses normalisasi mengubah dan mengembalikan bentuk penulisan kata yang tidak baku ke bentuk penulisan yang sesuai dengan KBBI. Baris ke-1 membaca dan memuat korpus dengan format *file* **.csv* ke dalam variabel `normalisasi`. Korpus berisikan kata-kata tidak baku dan bentuk baku dari kata

tersebut. Pada baris ke-2 hingga baris ke-4, sebuah *dictionary* dibuat untuk menampung kata tidak baku sebagai *key*, dan bentuk baku dari kata tersebut sebagai *value*. Baris ke-5 dan ke-6 merupakan fungsi yang mengembalikan bentuk baku (*value*) jika *term* yang diinputkan terdapat pada *dictionary*, atau mengembalikan *term* tersebut tanpa mengubahnya jika *term* tidak terdapat pada *dictionary*.

Tabel 4.6 Penggalan Kode Proses Normalisasi

No	Penggalan Kode
1	<code>normalisasi = pd.read_csv('new_colloquial-indonesian-lexicon.csv')</code>
2	<code>for index, row in normalisasi.iterrows():</code>
3	<code> if row[0] not in normalisasi_kata_dict:</code>
4	<code> normalisasi_kata_dict[row[0]] = row[1]</code>
5	<code>def normalized_term(document):</code>
6	<code> return [normalisasi_kata_dict[term] if term in normalisasi_kata_dict else term for term in document]</code>

4.3.6 Implementasi Stemming

Tahap *stemming* dilakukan untuk mengubah kata-kata berimbuhan (awalan dan akhiran) ke bentuk akarnya yaitu kata dasar. Baris ke-1 dan ke-2 adalah inisiasi untuk memanggil fungsi *stemming* dari *library* PySastrawi. Baris ke-3 dan ke-4 adalah fungsi untuk mengembalikan hasil *stemming* dari kata yang diinputkan. Pada baris ke-5 hingga baris ke-9, sebuah *dictionary* dibuat untuk menampung kata-kata pada dokumen yang diinputkan sebagai *key*. Lalu pada baris ke-10 dan ke-11 dibuat *value* dari setiap *key* tadi yaitu hasil *stemming* dari kata-kata tersebut. Baris ke-12 dan ke-13 merupakan sebuah fungsi yang mengembalikan bentuk kata dasar yang terdapat pada *dictionary* yang sudah dibuat.

Tabel 4.7 Penggalan Kode Proses Stemming

No	Penggalan Kode
1	<code>factory = StemmerFactory()</code>
2	<code>stemmer = factory.create_stemmer()</code>
3	<code>def stemmed_wrapper(term):</code>
4	<code> return stemmer.stem(term)</code>
5	<code>term_dict = {}</code>
6	<code>for document in df['Content']:</code>
7	<code> for term in document:</code>
8	<code> if term not in term_dict:</code>
9	<code> term_dict[term] = ' '</code>
10	<code>for term in term_dict:</code>
11	<code> term_dict[term] = stemmed_wrapper(term)</code>
12	<code>def get_stemmed_term(document):</code>
13	<code> return [term_dict[term] for term in document]</code>

4.4 Implementasi Pembobotan TF-IDF

Pada tahap ini, data teks yang diproses diubah menjadi data numerik dengan metode pembobotan TF-IDF sesuai alur pada gambar Gambar 3.3. Baris ke-1 menginisialisasikan dan memanggil *class* `TfidfVectorizer()` ke dalam variabel `tf`. Kemudian pada baris ke-2, dilakukan pemanggilan *method* `fit_transform` yang di dalamnya terdapat proses perhitungan *term frequency* yaitu jumlah kemunculan *term* di setiap dokumen, proses perhitungan *document frequency* yaitu jumlah dokumen yang mengandung suatu *term*, lalu perhitungan *inverse document frequency* atau IDF dari suatu *term* dengan cara membagi jumlah seluruh dokumen dengan jumlah dokumen yang mengandung *term* tersebut. IDF digunakan untuk mengukur apakah suatu *term* umum atau jarang dalam korpus dokumen tertentu. Kemudian pada *method* ini dihitung hasil akhir dari perkalian nilai TF dengan nilai IDF dan membuatnya ke dalam bentuk matriks. Baris ke-3 mengecek ukuran matriks dimana ukuran matriks data *training* adalah (240, 4934). Baris ke-4 adalah mengubah bentuk matriks `tfidf_matrix` ke dalam bentuk matriks *dense* dan memuatnya ke dalam variabel `matriks`. Lalu, pada baris ke-5, dibuat satu dataframe baru yang memuat nilai-nilai TF-IDF pada `matriks`, dengan fitur atau *term* sebagai kolom, dan *index* dari dokumen sebagai baris.

Tabel 4.8 Penggalan Kode Proses Pembobotan TF-IDF

No	Penggalan Kode
1	<code>tf = TfidfVectorizer()</code>
2	<code>tfidf_matrix = tf.fit_transform(df['Content'].values)</code>
3	<code>tfidf_matrix.shape</code>
4	<code>matriks = tfidf_matrix.todense()</code>
5	<code>df_new = pd.DataFrame(matriks, columns= tf.get_feature_names_out(), index=df.index)</code>

4.5 Implementasi Seleksi Fitur *Information Gain*

Seleksi fitur *Information Gain* digunakan untuk mengurangi fitur-fitur yang tidak memiliki pengaruh besar dalam tahap klasifikasi. Implementasi kode disesuaikan dengan alur pada Gambar 3.4. Metode *Information Gain* dibuat dalam bentuk *class*, yang didalamnya terdapat *object*, *attribute*, dan *method*. *Class* `TextFeatureSelection()` memiliki beberapa *method* di dalamnya. Baris ke-2

hingga ke-4 adalah inisialisasi *object* dan *attribute* dimana `self` merupakan *object*, `target` merupakan *attribute* yang menyimpan nilai label data, dan `input_doc_list` merupakan *attribute* yang menyimpan data tanpa label.

Baris ke-5 hingga baris ke-7 adalah *method* untuk menghitung nilai *Information Gain*, dengan A, B, C, D, N sebagai input. Berdasarkan *Entropy based Category Coverage Difference* (ECCD), A adalah jumlah dokumen dengan kelas 0 (fakta) yang mengandung *term* t_j . B adalah jumlah dokumen dengan kelas 1 (*hoax*) yang mengandung *term* t_j . C adalah jumlah dokumen dengan kelas 0 (fakta) yang tidak mengandung *term* t_j . D adalah jumlah dokumen dengan kelas 1 (*hoax*) yang tidak mengandung *term* t_j . N adalah jumlah keseluruhan dokumen (Lageron, Moulin, dan Géry, 2011). Perhitungan nilai *Information Gain* disesuaikan dengan persamaan (2.4). $(A+C)/N$ adalah peluang kelas 0 (fakta) pada seluruh dokumen. $(B+D)/N$ adalah peluang kelas 1 (*hoax*) pada seluruh dokumen. $(A+B)/N$ adalah peluang *term* t_j terdapat pada keseluruhan dokumen. $(C+D)/N$ adalah peluang *term* t_j tidak terdapat pada keseluruhan dokumen. $A/(A+B)$ adalah peluang *term* t_j terdapat pada kelas 0 (fakta). $C/(C+D)$ adalah peluang *term* t_j tidak terdapat pada kelas 0 (fakta). $B/(A+B)$ adalah peluang *term* t_j terdapat pada kelas 1 (*hoax*). $D/(C+D)$ adalah peluang *term* t_j tidak terdapat pada kelas 1 (*hoax*).

Tabel 4.9a Penggalan Kode Proses *Information Gain*

No	Penggalan Kode
1	<code>class TextFeatureSelection():</code>
2	<code>def __init__(self, target, input_doc_list):</code>
3	<code>self.target=target</code>
4	<code>self.input_doc_list=input_doc_list</code>
5	<code>def _InformationGain(self, A, B, C, D, N):</code>
6	<code>with np.errstate(divide='ignore', invalid='ignore'):</code>
7	<code>return -((A+C)/N * np.log((A+C)/N) + ((A+B)/N</code>
	<code>* (A/(A+B)) * np.log(A/(A+B)) + ((C+D)/N * (C/(C+D))</code>
	<code>* np.log(C/(C+D)) + -((B+D)/N * np.log((B+D)/N))</code>
	<code>+ ((A+B)/N * (B/(A+B)) * np.log(B/(A+B)) + ((C+D)/N</code>
	<code>* (D/(C+D)) * np.log(D/(C+D))</code>

Baris ke-8 hingga baris ke-15 adalah *method* `_get_binary_label` yang digunakan untuk mengubah label menjadi *binary* apabila label data belum berbentuk *binary*.

Tabel 4.9b Penggalan Kode Proses *Information Gain*

No	Penggalan Kode
8	<code>def _get_binary_label(self,label_array):</code>
9	<code> label_array=np.array(label_array)</code>
10	<code> unique_label=np.unique(label_array)</code>
11	<code> if 0 in unique_label and 1 in unique_label:</code>
12	<code> pass</code>
13	<code> else:</code>
14	<code> label_array=np.where(label_array==unique_label[0],1,0)</code>
15	<code> return label_array</code>

Baris ke-16 hingga baris ke-24 adalah *method* untuk mendapatkan representasi matriks dari data, jumlah dokumen yang mengandung setiap *term* atau fitur, dan daftar fitur-fitur yang ada. Pada baris ke-17 dan ke-18, *library* CountVectorizer diinisialisasikan dan digunakan untuk menghitung jumlah kemunculan *term* pada setiap dokumen. Baris ke-19 digunakan untuk mendapatkan dan menampung seluruh fitur yang ada. Baris ke-20 dan ke-21 menggunakan *library* yang sama, namun untuk mendapatkan kemunculan *term* dalam bentuk *binary*, jika suatu *term* muncul di dalam dokumen, maka nilainya adalah 1, dan sebaliknya. Baris ke-22 digunakan untuk mengubah nilai-nilai tadi ke dalam bentuk *array* dan menampungnya ke dalam variabel `word_binary_matrix`. Baris ke-23 menjumlahkan setiap kolom pada `word_binary_matrix` untuk mendapatkan jumlah dokumen yang mengandung setiap *term*.

Tabel 4.9c Penggalan Kode Proses *Information Gain*

No	Penggalan Kode
16	<code>def _get_term_binary_matrix(self,input_doc_list):</code>
17	<code> vectorizer = CountVectorizer()</code>
18	<code> X = vectorizer.fit_transform(input_doc_list)</code>
19	<code> word_list = vectorizer.get_feature_names_out()</code>
20	<code> vectorizer = CountVectorizer(binary=True)</code>
21	<code> X = vectorizer.fit_transform(input_doc_list)</code>
22	<code> word_binary_matrix = X.toarray()</code>
23	<code> count_list = word_binary_matrix.sum(axis=0)</code>
24	<code> return word_list,count_list,word_binary_matrix</code>

Baris ke-25 hingga baris ke-41 adalah *method* untuk mendapatkan nilai dari A,B,C,D, dan N. Input dari *method* ini adalah `word_binary_matrix` dan `label_array` yang nilainya telah didapatkan sebelumnya. Baris ke-26 hingga baris ke-29 adalah inisiasi dari variabel A,B,C,dan D yang masih belum berisikan nilai

apapun. Baris ke-30 hingga ke-35 adalah perulangan untuk mendapatkan nilai A, B, C, dan D. Untuk setiap *i* pada *range* matriks yaitu 4394, nilai pada *label_array* dikalikan dua dan ditambah dengan nilai setiap kolom fitur pada matriks. Nilai 0 didapatkan jika suatu *term* tidak terdapat pada dokumen dengan kelas 0 (fakta). Nilai 1 didapatkan jika suatu *term* terdapat pada dokumen dengan kelas 0 (fakta). Nilai 2 akan didapatkan jika suatu *term* tidak terdapat pada dokumen dengan kelas 1 (*hoax*). Nilai 3 akan didapatkan jika suatu *term* terdapat pada dokumen dengan kelas 1 (*hoax*). Dengan bantuan *class* Counter dari *library collections*, berdasarkan hasil perhitungan tadi, dihitung jumlah dokumen yang memiliki nilai 0, 1, 2, dan 3. Baris ke-32 hingga ke-35 menampung nilai-nilai tersebut ke dalam variabel. A untuk jumlah dokumen dengan nilai 1. B untuk jumlah dokumen dengan nilai 3. C untuk jumlah dokumen dengan nilai 0. D untuk jumlah dokumen dengan nilai 2. Kemudian pada baris ke-36 hingga ke-40, nilai-nilai pada variabel diubah menjadi bilangan *array*. Baris ke-40 menghitung nilai N yaitu jumlah keseluruhan dokumen.

Tabel 4.9d Penggalan Kode Proses *Information Gain*

No	Penggalan Kode
25	<code>def _get_ABCD(self, word_binary_matrix, label_array):</code>
26	<code> A=[]</code>
27	<code> B=[]</code>
28	<code> C=[]</code>
29	<code> D=[]</code>
30	<code> for i in range(word_binary_matrix.shape[1]):</code>
31	<code> computed_result=Counter(label_array * 2 +</code>
	<code> word_binary_matrix[:,i])</code>
32	<code> A.append(computed_result[1])</code>
33	<code> B.append(computed_result[3])</code>
34	<code> C.append(computed_result[0])</code>
35	<code> D.append(computed_result[2])</code>
36	<code> A=np.array(A)</code>
37	<code> B=np.array(B)</code>
38	<code> C=np.array(C)</code>
39	<code> D=np.array(D)</code>
40	<code> N=A+B+C+D</code>
41	<code> return A,B,C,D,N</code>

Baris ke-42 hingga baris ke-59 adalah *method* untuk mendapatkan nilai-nilai dari beberapa *method* lainnya. Baris ke-43 untuk mendapatkan bentuk *array* dari label data, baris ke-44 untuk mendapatkan daftar fitur, jumlah dokumen yang

mengandung setiap *term* atau fitur, dan representasi matriks *binary* kemunculan *term* di setiap dokumen, baris ke-45 untuk mendapatkan nilai A, B, C, D, dan N. Pada baris ke-46 hingga baris ke-48 adalah proses pembuatan sebuah *dataframe* baru yang menampung fitur, jumlah kemunculannya, dan nilai *Information Gain* dari fitur tersebut. Baris ke-50 hingga baris ke-52 adalah *method* untuk memanggil `_getvalues` dan menampung nilai-nilai pada variabel `values_df`.

Tabel 4.9e Penggalan Kode Proses *Information Gain*

No	Penggalan Kode
42	<code>def _getvalues(self):</code>
43	<code> label_array=self._get_binary_label(self.target)</code>
44	<code> word_list,count_list,word_binary_matrix=</code>
	<code> self._get_term_binary_matrix(self.input_doc_list)</code>
45	<code> A,B,C,D,N=self._get_ABCD(word_binary_matrix,label_array)</code>
46	<code> out_df=pd.DataFrame({'word list':word_list,</code>
	<code> 'word occurrence count':count_list})</code>
47	<code> out_df['Information Gain']=self._InformationGain</code>
	<code> (A,B,C,D,N)</code>
48	<code> out_df['Information Gain'].replace(np.nan,0,inplace=True)</code>
49	<code> return out_df</code>
50	<code>def getScore(self):</code>
51	<code> values_df=self._getvalues()</code>
52	<code> return values_df</code>

4.6 Implementasi Metode *K-Nearest Neighbor*

4.6.1 Implementasi *Euclidean Distance*

Pada tahap klasifikasi, digunakan metode *K-Nearest Neighbor* untuk menentukan kategori dari data berita yaitu *hoax* dan fakta. Sebelum menentukan label tetangga mana yang menjadi label dari data, maka dihitung terlebih dahulu jarak antara data latih dan data validasi menggunakan *Euclidean Distance* sesuai dengan Gambar 3.5. Sebuah *class* dibangun untuk menghitung jarak *Euclidean*. Pada baris ke-2 dan baris ke-3, diinisialisasikan *object* yang akan digunakan pada *class*. Baris ke-4 hingga baris ke-11 adalah proses perhitungan jarak. Inputan dari *method* ini adalah dua buah vektor yang didapatkan dari matriks TF-IDF yang sudah dihitung sebelumnya. Baris ke-5 adalah inisialisasi vector sebagai *attribute* dari *object*. Baris ke-6 dan baris ke-7 adalah pengecekan panjang vector pertama dan kedua, jika panjang tidak sama maka akan diberikan pesan peringatan. Baris ke-8 adalah inisialisasi variabel `distance`. Baris ke-9 dan baris ke-10 adalah

perulangan, dengan *range* panjang vektor pertama dikurang satu yaitu 4393, nilai pada vektor pertama akan dikurangi oleh nilai pada vektor kedua lalu dipangkatkan, hasil perhitungan tersebut akan dijumlahkan dan ditampung pada variabel *distance*. Kemudian baris ke-11 mengembalikan nilai pada variabel *distance* yang sudah diakarkan.

Tabel 4.9 Penggalan Kode Proses *Euclidean Distance*

No	Penggalan Kode
1	<code>class distanceMetrics:</code>
2	<code>def __init__(self):</code>
3	<code>Pass</code>
4	<code>def euclideanDistance(self, vector1, vector2):</code>
5	<code>self.vectorA, self.vectorB = vector1, vector2</code>
6	<code>if len(self.vectorA) != len(self.vectorB):</code>
7	<code>raise ValueError("Panjang vektor tidak sama")</code>
8	<code>distance = 0.0</code>
9	<code>for i in range(len(self.vectorA)-1):</code>
10	<code>distance += (self.vectorA[i] - self.vectorB[i])**2</code>
11	<code>return (distance)**0.5</code>

4.6.2 Implementasi *K-Nearest Neighbor*

Implementasi metode KNN disesuaikan dengan Gambar 3.5. Diawali dengan membangun *class* KNN, baris ke-2 dan baris ke-3 adalah inisialisasi *object* pada *class*. Baris ke-4 hingga baris ke-7 adalah inisialisasi data *training* dan label data *training* sebagai *attribute object*. Baris ke-8 hingga baris ke-17 adalah *method* untuk mendapatkan jumlah tetangga dengan jarak terdekat sesuai dengan nilai *k* yang diinputkan. Jarak antara data latih dan data uji dihitung dengan memanggil *class euclidean distance* yang sudah dibuat sebelumnya pada baris ke-9. Baris ke-10 adalah inisialisasi variabel *distances*. Baris ke-11 hingga baris ke-13 adalah perulangan, untuk setiap baris pada data latih, setiap kolom pada data latih akan dihitung jaraknya dengan kolom pada data uji dengan baris sesuai dengan inputan *testRow*, sehingga didapatkan jarak satu baris data uji terhadap seluruh data latih. Baris ke-13 mengurutkan jarak tersebut dari yang terdekat. Baris ke-14 menginisialisasikan variabel *neighbors*. Baris ke-15 hingga baris ke-17 mengembalikan nilai *neighbors* sejumlah *k* tetangga terdekat.

Baris ke-18 hingga baris ke-28 adalah *method* untuk memprediksi label dari data uji. Baris ke-19 hingga baris ke-21 adalah inisialisasi data uji, nilai *k*, dan

perhitungan jarak sebagai *attribute object*. Baris ke-22 menginisialisasikan variabel `predictions` untuk menampung hasil prediksi label. Baris ke-23 hingga baris ke-27 adalah perulangan, untuk setiap baris pada data uji, perhitungan jarak dan pengambilan tetangga dilakukan dengan memanggil *method* `getNeighbors`. Baris ke-25 untuk mengambil nilai label dari `k` tetangga terdekat. Lalu, pada baris ke-26 didapatkan label kelas terbanyak pada `k` tetangga terdekat. Baris ke-27 dan baris ke-28 masing-masing untuk menampung hasil prediksi dan mengembalikan seluruh nilai prediksi dari data uji.

Tabel 4.10 Penggalan Kode Proses *K-Nearest Neighbor*

No	Penggalan Kode
1	<code>class kNNClassifier:</code>
2	<code>def __init__(self):</code>
3	<code>Pass</code>
4	<code>def fit(self, xTrain, yTrain):</code>
5	<code>assert len(xTrain) == len(yTrain)</code>
6	<code>self.trainData = xTrain</code>
7	<code>self.trainLabels = yTrain</code>
8	<code>def getNeighbors(self, testRow):</code>
9	<code>calcDM = distanceMetrics()</code>
10	<code>distances = []</code>
11	<code>for i, trainRow in enumerate(self.trainData):</code>
12	<code>distances.append([trainRow,</code>
	<code>calcDM.euclideanDistance(testRow, trainRow),</code>
	<code>self.trainLabels[i]])</code>
13	<code>distances.sort(key=operator.itemgetter(1))</code>
14	<code>neighbors = []</code>
15	<code>for index in range(self.k):</code>
16	<code>neighbors.append(distances[index])</code>
17	<code>return neighbors</code>
18	<code>def predict(self, xTest, k, distanceMetric='euclidean'):</code>
19	<code>self.testData = xTest</code>
20	<code>self.k = k</code>
21	<code>self.distanceMetric = distanceMetric</code>
22	<code>predictions = []</code>
23	<code>for i, testCase in enumerate(self.testData):</code>
24	<code>neighbors = self.getNeighbors(testCase)</code>
25	<code>output= [row[-1] for row in neighbors]</code>
26	<code>prediction = max(set(output), key=output.count)</code>
27	<code>predictions.append(prediction)</code>
28	<code>return predictions</code>

4.7 Hasil Penelitian

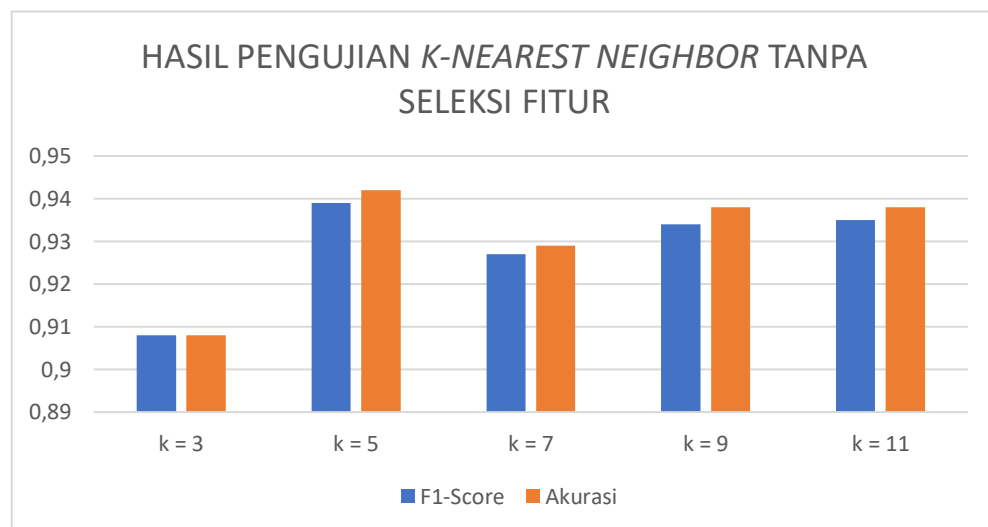
4.7.1 Hasil Pengujian Metode *K-Nearest Neighbor* tanpa Seleksi Fitur

Pada pengujian ini, metode *K-Nearest Neighbor* diuji tanpa menggunakan seleksi fitur. Proses pelatihan dan validasi menggunakan metode *N-Fold Cross*

Validation untuk mengetahui nilai k yang menghasilkan performa akurasi terbaik. Setelah melakukan pengujian dengan beberapa nilai parameter k yaitu $k=3$, $k=5$, $k=7$, $k=9$, dan $k=11$ dengan menggunakan 10 *fold*. Didapatkan nilai k dengan performa *F1-Score* terbaik yaitu $k = 5$ dengan nilai *F1-Score* 93.9% serta akurasi 94.2%. Sehingga, nilai $k = 5$ dipilih sebagai model terbaik dan akan digunakan pada proses pengujian data uji dengan menggunakan data baru yang belum digunakan pada saat proses pelatihan dan validasi.

Tabel 4.11 Hasil Pengujian *K-Nearest Neighbor* tanpa Seleksi Fitur

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.833	0.958	0.958	0.875	0.875	0.958	0.875	0.958	0.916	0.833	0.908
	F1-Score	0.846	0.96	0.96	0.96	0.857	0.957	0.88	0.916	0.923	0.818	0.908
5	Akurasi	0.916	1.0	0.958	0.875	1.0	0.916	1.0	0.916	0.916	0.916	0.942
	F1-Score	0.916	1.0	0.96	0.842	1.0	0.923	1.0	0.916	0.923	0.909	0.939
7	Akurasi	0.916	0.958	0.958	0.875	0.916	0.916	1.0	0.916	0.958	0.875	0.929
	F1-Score	0.916	0.96	0.96	0.842	0.916	0.923	1.0	0.916	0.963	0.87	0.927
9	Akurasi	0.875	1.0	1.0	0.875	0.916	0.916	1.0	0.916	0.958	0.916	0.938
	F1-Score	0.87	1.0	1.0	0.842	0.916	0.923	1.0	0.916	0.963	0.909	0.934
11	Akurasi	0.916	1.0	0.916	0.875	0.916	0.958	1.0	0.916	0.958	0.916	0.938
	F1-Score	0.916	1.0	0.923	0.842	0.916	0.96	1.0	0.916	0.963	0.909	0.935



Gambar 4.2 Hasil Pengujian *K-Nearest Neighbor* tanpa Seleksi Fitur

4.7.2 Hasil Pengujian Metode *K-Nearest Neighbor* dengan Seleksi Fitur *Information Gain*

Pada pengujian ini, metode *K-Nearest Neighbor* diuji menggunakan seleksi fitur *Information Gain*. Beberapa skenario perubahan *threshold* digunakan pada pengujian ini. Total keseluruhan fitur adalah sebanyak 4934 fitur. *Threshold* akan menyeleksi fitur-fitur dengan nilai *Information Gain* tertinggi sesuai dengan presentase yang diberikan. Mengacu pada penelitian (Paramitha, Indriati, dan Sari, 2020) yang mendapatkan akurasi tertinggi pada *threshold* 25%, pada penelitian ini digunakan batas *threshold* yaitu 50%, 25%, 20%, 10%, 5%, 2%, 1%, 0.5%, 0.2%, dan 0.1%. Proses pelatihan dan validasi menggunakan metode *N-Fold Cross Validation* untuk mengetahui nilai *k* yang menghasilkan performa *F1-Score* terbaik. Nilai parameter *k* yang digunakan yaitu *k*=3, *k*=5, *k*=7, *k*=9, dan *k*=11 dengan menggunakan 10 *fold*.

a. Pengujian *Threshold* 50%

Pada pengujian yang ditunjukkan pada Tabel 4.12, fitur-fitur yang digunakan adalah sebesar 50% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 2467 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 25.5% dengan parameter nilai *k* = 3.

Tabel 4.12 Pengujian *Information Gain Threshold* 50%

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.583	0.458	0.667	0.667	0.583	0.5	0.583	0.583	0.542	0.583	0.575
	F1-Score	0.167	0.0	0.556	0.429	0.285	0.143	0.286	0.167	0.353	0.167	0.255
5	Akurasi	0.583	0.5	0.583	0.667	0.542	0.5	0.583	0.625	0.5	0.542	0.563
	F1-Score	0.167	0.143	0.375	0.429	0.154	0.143	0.286	0.308	0.25	0.154	0.153
7	Akurasi	0.583	0.5	0.583	0.583	0.542	0.5	0.583	0.625	0.458	0.542	0.55
	F1-Score	0.167	0.143	0.375	0.167	0.154	0.143	0.286	0.308	0.133	0.154	0.203
9	Akurasi	0.583	0.458	0.542	0.583	0.542	0.542	0.583	0.625	0.5	0.5	0.546
	F1-Score	0.167	0.0	0.267	0.167	0.154	0.267	0.286	0.308	0.25	0.0	0.186
11	Akurasi	0.542	0.542	0.625	0.583	0.542	0.542	0.542	0.625	0.5	0.542	0.558
	F1-Score	0.0	0.267	0.471	0.167	0.154	0.267	0.154	0.308	0.25	0.154	0.219

b. Pengujian *Threshold* 25%

Pada pengujian yang ditunjukkan pada Tabel 4.13, fitur-fitur yang digunakan adalah sebesar 25% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 1234 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 27.1% dengan parameter nilai $k = 3$.

Tabel 4.13 Pengujian *Information Gain Threshold* 25%

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.583	0.5	0.583	0.667	0.542	0.542	0.625	0.583	0.583	0.583	0.579
	F1-Score	0.167	0.143	0.375	0.429	0.154	0.267	0.4	0.167	0.444	0.167	0.271
5	Akurasi	0.583	0.458	0.5	0.625	0.542	0.5	0.625	0.625	0.542	0.583	0.558
	F1-Score	0.167	0.0	0.143	0.308	0.154	0.143	0.4	0.308	0.353	0.167	0.214
7	Akurasi	0.542	0.458	0.583	0.542	0.5	0.5	0.583	0.542	0.542	0.583	0.537
	F1-Score	0.0	0.0	0.375	0.0	0.0	0.143	0.286	0.0	0.353	0.167	0.167
9	Akurasi	0.542	0.458	0.542	0.542	0.5	0.5	0.542	0.583	0.5	0.583	0.53
	F1-Score	0.0	0.0	0.267	0.0	0.0	0.143	0.153	0.167	0.25	0.167	0.115
11	Akurasi	0.542	0.458	0.5	0.542	0.5	0.5	0.5	0.542	0.458	0.542	0.508
	F1-Score	0.0	0.0	0.143	0.0	0.0	0.143	0.0	0.0	0.133	0.0	0.042

c. Pengujian *Threshold* 20%

Pada pengujian yang ditunjukkan pada Tabel 4.14, fitur-fitur yang digunakan adalah sebesar 20% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 987 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 26.6% dengan parameter nilai $k = 3$.

Tabel 4.14 Pengujian *Information Gain Threshold* 20%

k	Ukuran	Fold										Mean
	Evaluasi	1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.583	0.5	0.625	0.667	0.542	0.5	0.625	0.625	0.583	0.542	0.58
	F1-Score	0.167	0.143	0.471	0.429	0.154	0.143	0.4	0.308	0.444	0.0	0.266
5	Akurasi	0.542	0.458	0.542	0.667	0.542	0.5	0.625	0.667	0.542	0.583	0.567
	F1-Score	0.0	0.0	0.267	0.429	0.154	0.143	0.4	0.429	0.353	0.167	0.234
7	Akurasi	0.542	0.458	0.5	0.625	0.5	0.5	0.542	0.542	0.542	0.583	0.533
	F1-Score	0.0	0.0	0.143	0.308	0.0	0.143	0.154	0.0	0.353	0.167	0.127
9	Akurasi	0.542	0.458	0.542	0.542	0.5	0.5	0.542	0.583	0.542	0.583	0.533
	F1-Score	0.0	0.0	0.267	0.0	0.0	0.143	0.154	0.167	0.353	0.167	0.125
11	Akurasi	0.542	0.458	0.5	0.542	0.5	0.5	0.5	0.542	0.5	0.583	0.517
	F1-Score	0.0	0.0	0.143	0.0	0.0	0.143	0.0	0.0	0.25	0.167	0.07

d. Pengujian *Threshold* 10%

Pada pengujian yang ditunjukkan pada Tabel 4.15, fitur-fitur yang digunakan adalah sebesar 10% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 493 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 86.9% dengan parameter nilai $k = 5$.

Tabel 4.15 Pengujian *Information Gain Threshold* 10%

k	Ukuran	Fold										Mean
	Evaluasi	1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.917	0.833	0.958	0.875	0.875	0.875	0.833	0.833	0.917	0.75	0.86
	F1-Score	0.9	0.833	0.963	0.857	0.857	0.87	0.818	0.846	0.929	0.75	0.862
5	Akurasi	0.875	0.875	0.917	0.917	0.833	0.833	0.875	0.875	0.917	0.833	0.875
	F1-Score	0.857	0.88	0.929	0.9	0.818	0.818	0.857	0.88	0.929	0.818	0.869
7	Akurasi	0.875	0.875	0.958	0.875	0.833	0.833	0.833	0.875	0.917	0.833	0.871
	F1-Score	0.842	0.88	0.963	0.842	0.8	0.818	0.8	0.88	0.929	0.818	0.857
9	Akurasi	0.917	0.833	0.917	0.875	0.792	0.833	0.875	0.875	0.958	0.792	0.867
	F1-Score	0.9	0.833	0.923	0.842	0.762	0.818	0.857	0.88	0.966	0.783	0.856
11	Akurasi	0.917	0.833	0.917	0.833	0.75	0.75	0.917	0.917	0.917	0.792	0.854
	F1-Score	0.9	0.833	0.923	0.778	0.7	0.7	0.909	0.917	0.929	0.783	0.837

e. Pengujian *Threshold* 5%

Pada pengujian yang ditunjukkan pada Tabel 4.16, fitur-fitur yang digunakan adalah sebesar 5% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 247 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 91.2% dengan parameter nilai $k = 5$.

Tabel 4.16 Pengujian *Information Gain Threshold* 5%

k	Ukuran	Fold										Mean
	Evaluasi	1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.958	0.958	0.958	0.917	0.792	0.875	0.917	0.875	0.917	0.875	0.904
	F1-Score	0.957	0.963	0.963	0.9	0.783	0.889	0.909	0.87	0.929	0.88	0.904
5	Akurasi	0.958	0.958	0.917	0.875	0.833	0.917	0.917	0.917	0.917	0.917	0.913
	F1-Score	0.957	0.963	0.929	0.857	0.818	0.923	0.909	0.917	0.929	0.917	0.912
7	Akurasi	0.917	0.958	0.958	0.833	0.833	0.875	0.917	0.917	0.917	0.875	0.9
	F1-Score	0.917	0.963	0.963	0.818	0.818	0.88	0.909	0.917	0.929	0.88	0.899
9	Akurasi	0.958	0.958	0.958	0.833	0.875	0.875	0.833	0.917	0.917	0.792	0.892
	F1-Score	0.957	0.963	0.963	0.8	0.87	0.88	0.818	0.917	0.929	0.8	0.89
11	Akurasi	0.917	0.958	0.958	0.875	0.833	0.875	0.792	0.917	0.917	0.708	0.875
	F1-Score	0.917	0.963	0.963	0.842	0.818	0.88	0.762	0.917	0.929	0.72	0.871

f. Pengujian *Threshold* 2%

Pada pengujian yang ditunjukkan pada Tabel 4.17, fitur-fitur yang digunakan adalah sebesar 2% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 99 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 95.2% dengan parameter nilai $k = 9$.

Tabel 4.17 Pengujian *Information Gain Threshold 2%*

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.917	0.958	1.0	0.958	0.958	0.875	0.958	0.708	1.0	0.792	0.912
	F1-Score	0.917	0.96	1.0	0.957	0.96	0.88	0.957	0.696	1.0	0.8	0.913
5	Akurasi	0.958	0.917	1.0	0.917	0.917	0.958	0.958	0.875	1.0	0.958	0.946
	F1-Score	0.957	0.917	1.0	0.917	0.917	0.96	0.957	0.857	1.0	0.957	0.944
7	Akurasi	1.0	0.917	1.0	0.917	0.875	0.958	0.958	0.917	1.0	0.958	0.95
	F1-Score	1.0	0.917	1.0	0.917	0.87	0.96	0.957	0.9	1.0	0.957	0.948
9	Akurasi	0.917	0.917	0.958	0.958	0.958	1.0	0.958	0.917	1.0	0.958	0.954
	F1-Score	0.917	0.917	0.963	0.957	0.957	1.0	0.957	0.9	1.0	0.957	0.952
11	Akurasi	0.833	0.917	1.0	0.958	0.917	0.958	0.958	0.917	1.0	0.917	0.937
	F1-Score	0.846	0.917	1.0	0.957	0.909	0.963	0.957	0.9	1.0	0.909	0.936

g. Pengujian *Threshold 1%*

Pada pengujian yang ditunjukkan pada Tabel 4.18, fitur-fitur yang digunakan adalah sebesar 1% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 49 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 95.8% dengan parameter nilai $k = 11$.

Tabel 4.18 Pengujian *Information Gain Threshold 1%*

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.917	0.958	0.958	0.917	1.0	0.917	1.0	0.875	0.958	0.875	0.938
	F1-Score	0.917	0.96	0.963	0.9	1.0	0.923	1.0	0.88	0.966	0.87	0.938
5	Akurasi	0.917	0.917	0.917	0.917	0.958	0.875	1.0	0.875	0.958	0.875	0.921
	F1-Score	0.917	0.917	0.929	0.9	0.96	0.889	1.0	0.88	0.966	0.87	0.923
7	Akurasi	0.875	0.917	1.0	0.917	0.917	0.917	0.958	0.917	1.0	0.917	0.933
	F1-Score	0.88	0.917	1.0	0.9	0.917	0.917	0.957	0.909	1.0	0.909	0.93
9	Akurasi	0.958	0.917	0.958	0.958	0.958	0.958	0.958	0.917	0.917	0.917	0.942
	F1-Score	0.957	0.917	0.963	0.952	0.957	0.96	0.957	0.909	0.929	0.909	0.941
11	Akurasi	0.875	0.958	0.958	0.958	0.958	0.958	0.958	1.0	1.0	0.958	0.958
	F1-Score	0.88	0.96	0.963	0.952	0.957	0.96	0.957	1.0	1.0	0.952	0.958

h. Pengujian *Threshold* 0.5%

Pada pengujian yang ditunjukkan pada Tabel 4.19, fitur-fitur yang digunakan adalah sebesar 0.5% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 25 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 97.3% dengan parameter nilai $k = 3$.

Tabel 4.19 Pengujian *Information Gain Threshold* 0.5%

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.917	1.0	1.0	0.958	0.958	0.958	1.0	1.0	0.958	1.0	0.975
	F1-Score	0.9	1.0	1.0	0.952	0.957	0.96	1.0	1.0	0.963	1.0	0.973
5	Akurasi	0.958	1.0	1.0	0.958	0.917	0.958	0.958	1.0	0.958	1.0	0.971
	F1-Score	0.952	1.0	1.0	0.952	0.917	0.96	0.96	1.0	0.963	1.0	0.97
7	Akurasi	0.875	1.0	1.0	0.917	0.917	0.917	0.958	1.0	0.958	0.958	0.95
	F1-Score	0.87	1.0	1.0	0.9	0.917	0.923	0.96	1.0	0.963	0.957	0.949
9	Akurasi	0.833	1.0	0.917	0.917	0.917	0.917	0.958	1.0	0.958	0.958	0.938
	F1-Score	0.833	1.0	0.929	0.9	0.923	0.923	0.96	1.0	0.963	0.957	0.939
11	Akurasi	0.833	1.0	0.917	0.875	0.917	0.833	0.958	0.958	0.958	0.917	0.917
	F1-Score	0.846	1.0	0.929	0.857	0.923	0.857	0.96	0.957	0.963	0.917	0.921

i. Pengujian *Threshold* 0.2%

Pada pengujian yang ditunjukkan pada Tabel 4.20, fitur-fitur yang digunakan adalah sebesar 0.2% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 10 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 93.7% dengan parameter nilai $k = 3$ dan $k=5$.

Tabel 4.20 Pengujian *Information Gain Threshold* 0.2%

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.917	0.917	1.0	0.917	0.875	0.958	0.958	0.958	0.958	0.958	0.942
	F1-Score	0.909	0.917	1.0	0.9	0.857	0.96	0.957	0.952	0.963	0.952	0.937
5	Akurasi	0.958	0.917	1.0	0.917	0.875	0.958	0.917	0.958	0.958	0.958	0.942
	F1-Score	0.957	0.917	1.0	0.9	0.857	0.96	0.909	0.952	0.963	0.952	0.937
7	Akurasi	0.958	0.917	1.0	0.917	0.875	0.958	0.916	0.958	0.958	0.917	0.937
	F1-Score	0.957	0.917	1.0	0.9	0.857	0.96	0.909	0.952	0.963	0.9	0.931
9	Akurasi	0.958	0.917	1.0	0.917	0.875	0.958	0.917	0.958	0.958	0.917	0.937
	F1-Score	0.957	0.917	1.0	0.9	0.857	0.96	0.909	0.952	0.963	0.9	0.931
11	Akurasi	0.958	0.917	1.0	0.917	0.875	0.958	0.917	0.958	0.958	0.917	0.937
	F1-Score	0.957	0.916	1.0	0.9	0.857	0.96	0.909	0.952	0.963	0.9	0.931

j. Pengujian *Threshold* 0.1%

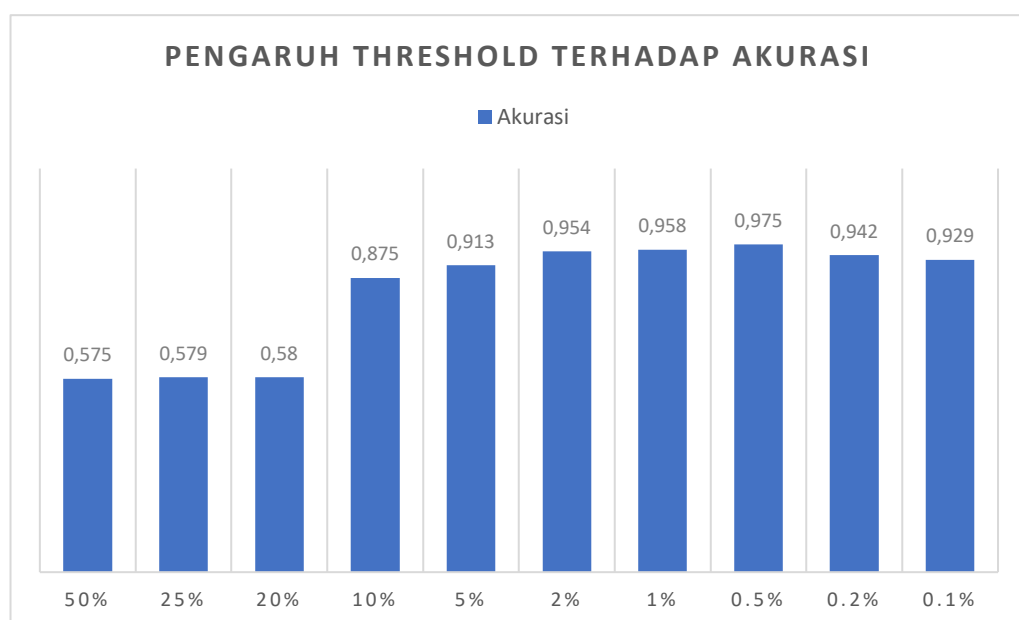
Pada pengujian yang ditunjukkan pada Tabel 4.21, fitur-fitur yang digunakan adalah sebesar 0.1% fitur dengan nilai *Information Gain* tertinggi, dengan jumlah sekitar 5 fitur. Didapatkan hasil performa *F1-Score* terbesar yaitu 92.5% dengan parameter nilai $k = 7$, $k = 9$, dan $k = 11$.

Tabel 4.21 Pengujian *Information Gain Threshold* 0.1%

k	Ukuran Evaluasi	Fold										Mean
		1	2	3	4	5	6	7	8	9	10	
3	Akurasi	0.917	0.833	0.958	0.917	0.917	0.917	0.958	0.958	0.917	0.958	0.925
	F1-Score	0.909	0.833	0.963	0.9	0.909	0.917	0.957	0.952	0.923	0.952	0.922
5	Akurasi	0.917	0.833	0.958	0.917	0.917	0.917	0.958	0.958	0.917	0.958	0.925
	F1-Score	0.909	0.833	0.963	0.9	0.909	0.917	0.957	0.952	0.923	0.952	0.922
7	Akurasi	0.917	0.875	0.958	0.917	0.917	0.917	0.958	0.958	0.917	0.958	0.929
	F1-Score	0.909	0.87	0.963	0.9	0.909	0.917	0.957	0.952	0.923	0.952	0.925
9	Akurasi	0.917	0.875	0.958	0.917	0.917	0.917	0.958	0.958	0.917	0.958	0.929
	F1-Score	0.909	0.87	0.963	0.9	0.909	0.917	0.957	0.952	0.923	0.952	0.925
11	Akurasi	0.917	0.875	0.958	0.917	0.917	0.917	0.958	0.958	0.917	0.958	0.929
	F1-Score	0.909	0.87	0.963	0.9	0.909	0.917	0.957	0.952	0.923	0.952	0.925

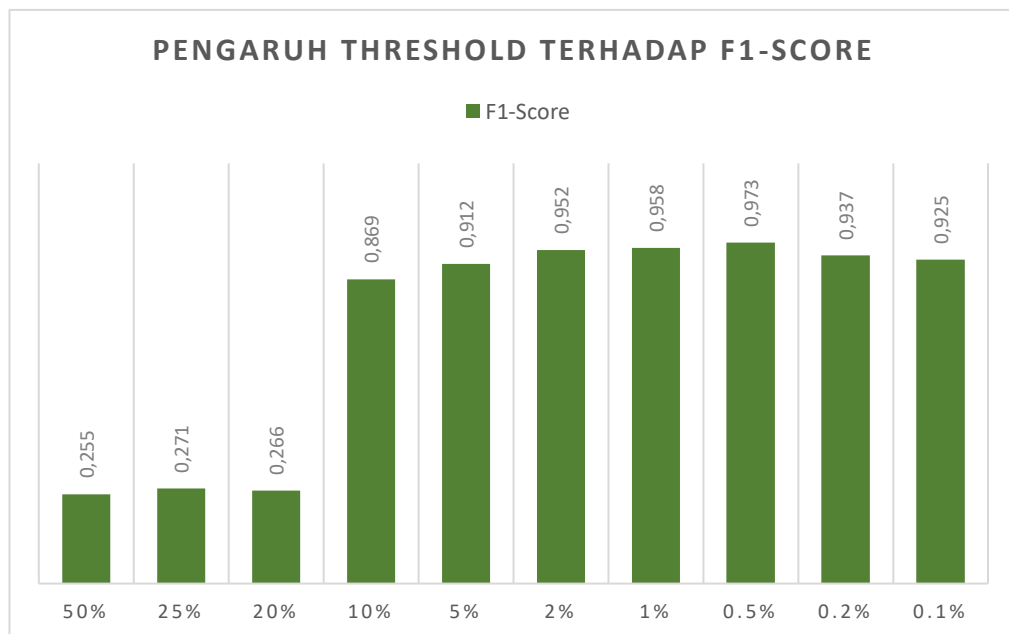
4.7.3 Pengaruh *Threshold Information Gain* terhadap Pengujian Metode *K-Nearest Neighbor*

Pengujian kombinasi *threshold Information Gain* dan nilai k yang sudah dilakukan menghasilkan beragam performa yang berbeda. Dari hasil pengujian yang diperoleh dapat diketahui pengaruh *threshold Information Gain* yang digunakan terhadap evaluasi kinerja metode *K-Nearest Neighbor*, hal tersebut dapat dilihat pada Gambar 4.3 sampai Gambar 4.4.



Gambar 4.3 Pengaruh *Threshold* terhadap Akurasi

Akurasi yang tertera pada Gambar 4.3 adalah akurasi tertinggi dari kombinasi nilai k yang sudah diuji. Nilai *threshold* diketahui memiliki pengaruh terhadap akurasi. Untuk dataset yang digunakan pada penelitian ini, rata-rata hasil akurasi dari 10 eksperimen *threshold* adalah sebesar 82.8%. Model dengan *threshold* yang menghasilkan akurasi di atas nilai rata-rata tersebut dapat dikatakan sebagai model dengan performa yang baik dalam mengklasifikasikan berita *hoax*. Pada Gambar 4.3 dapat dilihat adanya perubahan akurasi yang dihasilkan. Semakin menurun nilai *threshold*, akurasi dari metode cenderung meningkat, dan mencapai puncak akurasi tertinggi yang pada *threshold* 0.5%. Nilai *threshold* yang kecil berarti bahwa jumlah fitur-fitur yang digunakan hanya sedikit namun memiliki nilai *Information Gain* yang tinggi.



Gambar 4.4 Pengaruh *Threshold* terhadap *F1-Score*

Gambar 4.4 menunjukkan pengaruh *threshold* terhadap performa *F1-Score* metode *K-Nearest Neighbor*. *F1-Score* yang tertera pada Gambar 4.4 adalah nilai *F1-Score* dari kombinasi nilai *k* dengan akurasi tertinggi. Nilai *threshold* diketahui memiliki pengaruh terhadap *F1-Score*. Untuk dataset yang digunakan pada penelitian ini, rata-rata hasil *F1-Score* dari 10 eksperimen *threshold* adalah sebesar 73.18%. Model dengan *threshold* yang menghasilkan *F1-Score* di atas nilai rata-rata tersebut dapat dikatakan sebagai model dengan performa yang baik dalam mengklasifikasikan berita *hoax*. Pada Gambar 4.4 dapat dilihat adanya perubahan akurasi yang dihasilkan. Semakin menurun nilai *threshold*, nilai *F1-Score* dari metode cenderung meningkat dan puncak *F1-Score* tertinggi yang dihasilkan adalah pada *threshold* 0.5%.

Pada Gambar 4.3 sampai Gambar 4.4, diketahui bahwa semakin menurun presentase *threshold*, performa akurasi dan *F1-Score* cenderung meningkat. Hal ini disebabkan oleh presentase *threshold* yang kecil menyeleksi fitur menjadi lebih sedikit namun memiliki nilai *Information Gain* yang tinggi. Puncak *F1-Score* dan akurasi tertinggi yang dihasilkan adalah pada *threshold* 0.5% dengan kombinasi parameter nilai *k*=3. Jumlah fitur yang terseleksi pada *threshold* 0.5%

adalah sekitar 25 fitur. Kombinasi model inilah yang dipilih menjadi model terbaik dan akan digunakan pada proses pengujian data baru. Beberapa fitur yang terpilih seperti kata “klaim”, “unggah”, dan “akun” memiliki nilai *Information Gain* yang tinggi dan muncul di banyak dokumen pada kelas *hoax*. Beberapa fitur terpilih lainnya seperti “data”, “positif”, dan “tambah” juga memiliki nilai *Information Gain* yang tinggi dan muncul di banyak dokumen pada kelas fakta. Fitur-fitur tersebut menjadi suatu ciri khas dari kedua kelas berita, dimana berita *hoax* biasanya berkaitan dengan suatu klaim pada unggahan palsu di akun media sosial, dan berita fakta biasanya berkaitan dengan data sebaran virus serta penambahan angka kasus positif. Sehingga hal ini menyebabkan model dapat mengenali kelas dari berita dengan baik walaupun hanya dengan menggunakan fitur yang sedikit.

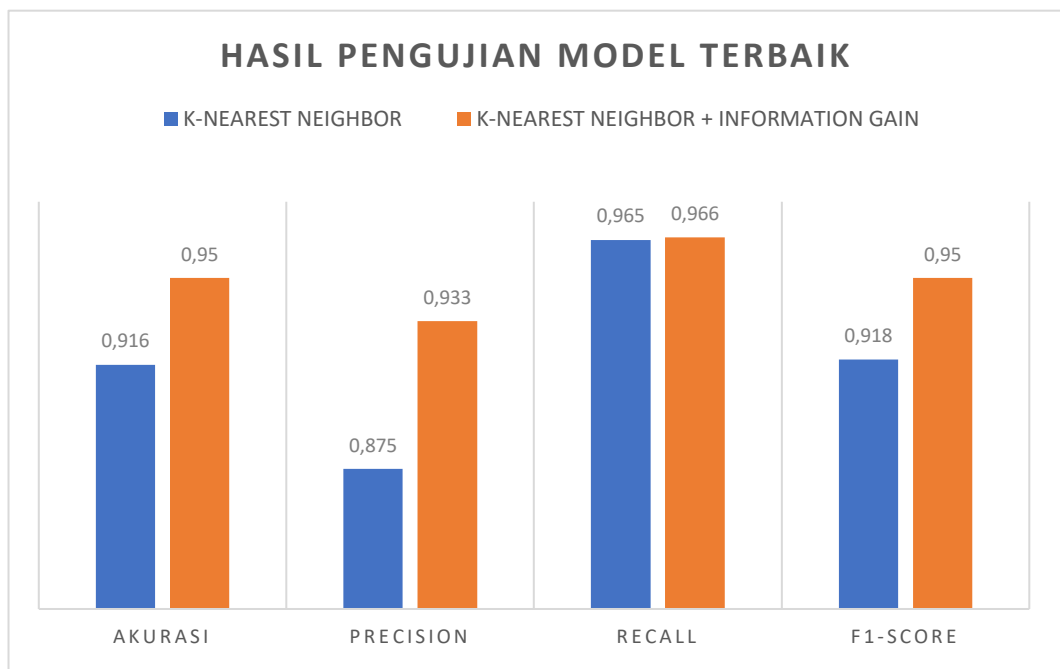
4.7.4 Pengujian Model Terbaik

Dua model terbaik yang dipilih adalah model yang menghasilkan akurasi terbaik, yaitu model *K-Nearest Neighbor* tanpa seleksi fitur dengan nilai $k = 5$, dan model *K-Nearest Neighbor* dengan kombinasi seleksi fitur *Information Gain threshold* 0.5% dengan nilai $k=3$. Kedua model tersebut kemudian diuji kembali menggunakan data baru yang belum pernah melewati tahap pelatihan dan validasi sebelumnya. Kemudian, performa model dibandingkan terlebih dahulu dengan performa pada saat pelatihan dan validasi untuk melihat adanya *overfitting* atau tidak. *Overfitting* adalah suatu keadaan jika model menghasilkan akurasi yang tinggi pada saat pelatihan dan validasi namun akurasinya menurun secara signifikan ketika diuji dengan data baru pada proses *testing*. Setelah dilakukan perbandingan, dapat dilihat bahwa kedua model tidak mengalami *overfitting* dilihat dari penurunan akurasi yang tidak signifikan. Perbandingan tersebut data dilihat pada Tabel 4.22.

Tabel 4.22 Perbandingan Hasil Pengujian Model

Metode	Ukuran Evaluasi	Pengujian	
		Training Validasi	Testing Data Baru
KNN	Akurasi	0.942	0.916
	Precision	0.952	0.875
	Recall	0.934	0.965
	F1-Score	0.939	0.918
KNN + IG	Akurasi	0.975	0.95
	Precision	1.0	0.933
	Recall	0.95	0.966
	F1-Score	0.973	0.95

Setelah melakukan pengujian terhadap kedua model dengan menggunakan data baru, hasil performa model dapat dilihat seperti pada Gambar 4.5.



Gambar 4.5 Hasil Pengujian Model Terbaik

Pada Gambar 4.7 dapat dilihat bahwa implementasi seleksi fitur *Information Gain* berpengaruh terhadap performa metode *K-Nearest Neighbor* dalam klasifikasi berita *hoax*. Terdapat peningkatan pada setiap performa evaluasi model. Akurasi

yang dihasilkan tanpa seleksi fitur adalah sebesar 91.6%, sedangkan akurasi dengan menggunakan seleksi fitur adalah sebesar 95%. Akurasi yang tinggi menunjukkan bahwa model dapat dengan baik mengklasifikasikan berita fakta ke dalam kelas fakta, dan berita *hoax* ke dalam kelas *hoax*, tanpa memperhatikan perbandingan ketepatan prediksi *hoax* terhadap kesalahan prediksi. Sehingga, perlu adanya ukuran evaluasi lain yang memperhitungkan kesalahan prediksi, yaitu *precision*, *recall*, dan *F1-Score*.

Nilai *precision* yang dihasilkan tanpa seleksi fitur adalah sebesar 87.5%, sedangkan *precision* dengan menggunakan seleksi fitur adalah sebesar 93.3%. Seleksi fitur *Information Gain* meningkatkan nilai *precision* dengan cukup signifikan. Nilai *precision* yang tinggi menunjukkan bahwa model dapat dengan baik mengklasifikasikan berita *hoax* ke dalam kelas *hoax*, dan sedikit kesalahan prediksi berita fakta ke dalam kelas *hoax*. Nilai *recall* yang dihasilkan tanpa seleksi fitur adalah sebesar 96.5%, sedangkan *recall* dengan menggunakan seleksi fitur adalah sebesar 96.6%. Tidak terlihat peningkatan yang signifikan dari nilai *recall*, karena nilai *recall* sudah cukup tinggi sebelum mengimplementasikan seleksi fitur. Nilai *recall* yang tinggi menunjukkan bahwa model dapat dengan baik mengklasifikasikan berita *hoax* ke dalam kelas *hoax*, dan sedikit kesalahan prediksi berita *hoax* ke dalam kelas fakta. Nilai *F1-Score* yang dihasilkan tanpa seleksi fitur adalah sebesar 91.8%, sedangkan *F1-Score* dengan menggunakan seleksi fitur adalah sebesar 95%. Terdapat sedikit peningkatan dari nilai *F1-Score*. Nilai *F1-Score* digunakan untuk melihat keseimbangan antara *precision* dan *recall*, sehingga nilai ini bisa digunakan juga sebagai ukuran evaluasi sebuah model selain menggunakan akurasi.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, ditarik kesimpulan bahwa implementasi seleksi fitur *Information Gain* dapat meningkatkan performa *precision*, *recall*, *F1-Score*, dan akurasi dari metode *K-Nearest Neighbor* dalam mengklasifikasikan berita *hoax*. Setelah dilakukan validasi model dengan *10-Fold Cross Validation*, nilai $k=5$ dipilih menjadi model terbaik pada eksperimen metode *K-Nearest Neighbor*. Pada pengujian data baru, model *K-Nearest Neighbor* tanpa seleksi fitur dengan nilai $k=5$ menghasilkan performa *precision*, *recall*, *F1-Score*, dan akurasi secara berturut-turut yaitu 87.5%, 96.5%, 91.8%, dan 91.6%. Pada eksperimen seleksi fitur *Information Gain*, kombinasi *threshold* 0.5% dengan parameter nilai $k=3$ adalah kombinasi yang dipilih menjadi model terbaik. *Threshold* ini menyeleksi sekitar 25 fitur dengan nilai *Information Gain* tertinggi. Pada pengujian data baru, model *K-Nearest Neighbor* dengan seleksi fitur *Information Gain* menghasilkan performa *precision*, *recall*, *F1-Score*, dan akurasi secara berturut-turut yaitu 93.3%, 96.6%, 95%, dan 95%.

5.2 Saran

Berdasarkan penelitian yang sudah dilakukan, saran-saran yang dapat penulis sampaikan untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Metode seleksi fitur dapat diganti dengan metode lainnya untuk menghasilkan performa *precision*, *recall*, *F1-Score*, dan akurasi yang lebih baik.
2. *Dataset* dapat diganti dengan *dataset* lainnya, karena *dataset* lain dapat menghasilkan performa yang berbeda.
3. Menambah variasi eksperimen nilai k untuk mendapatkan performa yang lebih baik dan beragam.
4. Menambah variasi eksperimen *threshold value* untuk mendapatkan performa yang lebih baik dan beragam.

DAFTAR PUSTAKA

- Aggarwal, C. C., & Aggarwal, C. C. (2018). Machine Learning for Text: An Introduction. In *Machine Learning for Text*. https://doi.org/10.1007/978-3-319-73531-3_1
- Aggarwal, C. C., & Zhai, C. (2012). Mining Text Data. In *Springer Science+Business Media, LLC* (Vol. 15, Issue 1). <https://doi.org/10.1111/j.1751-1097.1972.tb06217.x>
- Assiroj, P., Hidayanto, A.N., Prabowo, H., Warnars, H.L.H.S., et al., 2018. Hoax news detection on social media: a survey. In: 2018 Indonesian Association for Pattern Recognition International Conference (INAPR), IEEE, pp. 186–191.
- Boehmke, B., & Greenwell, B. (2019). *Hands-On Machine Learning with R*. Chapman and Hall/CRC.
- Das, M., Kamalanathan, S., & Alphonse, P. (2021). A Comparative Study on TF-IDF feature weighting method and its analysis using unstructured dataset. *CEUR Workshop Proceedings*, 2870, 98–107.
- Diaz Gene. (2016, 11 Oktober). Stopwords Indonesian (ID). Diakses pada 22 Februari 2022, dari <https://github.com/stopwords-iso/stopwords-id>.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann Publishers Inc.
- Kementerian Komunikasi dan Informatika. (2022, 18 Maret). Penanganan Sebaran Konten Hoaks Covid-19 Jumat (18/03/2022). Diakses pada 26 Maret 2022, dari <https://www.kominfo.go.id/content/detail/40626/penanganan-sebaran-konten-hoaks-covid-19-jumat-18032022/0/infografis>
- Largerone, C., Moulin, C., & Géry, M. (2011). Entropy based feature selection for text categorization. *ACM Symposium on Applied Computing*, 924–928.
- Listiowarni, I., & Puspa Dewi, N. (2020). Pemanfaatan Klasifikasi Soal Biologi Cognitive Domain Bloom's Taxonomy Menggunakan KNN Chi-Square Sebagai Penyusunan Naskah Soal. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 11(2), 186–197. <https://doi.org/10.31849/digitalzone.v11i2.4798>
- Mubaroq, I. M., & Setiawan, E. B. (2020). *The Effect of Information Gain Feature Selection for Hoax Identification in Twitter Using Classification Method Support Vector*. 5(September), 107–118. <https://doi.org/10.21108/indoic.2020.5.2.499>
- Paramitha, A. A., Indriati, & Sari, Y. A. (2020). Analisis Sentimen Terhadap Ulasan Pengguna MRT Jakarta Menggunakan Information Gain dan Modified K-Nearest Neighbor. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(4), 1125–1132.
- Sagita, R., Enri, U., & Primajaya, A. (2020). Klasifikasi Berita Clickbait Menggunakan K-Nearest Neighbor (KNN). *JOINS (Journal of Information System)*, 5(2), 230–239. <https://doi.org/10.33633/joins.v5i2.3705>
- Salsabila, N. A., Winatmoko, Y. A., Septiandri, A. A., & Jamal, A. (2019). Colloquial Indonesian Lexicon. *2018 International Conference on Asian Language Processing (IALP)*. <https://doi.org/10.1109/IALP.2018.8629151>
- Saputra, V. W., Sari, Y. A., & Widodo, A. W. (2019). Klasifikasi Jenis Makanan

menggunakan Neighbor Weighted K-Nearest Neighbor dengan Seleksi Fitur Information Gain. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(Mei 2019), 4749–4758.

Zong, C., Xia, R., & Zhang, J. (2021). *Text Data Mining*. Springer.

Zuliarso, E., Anwar, M. T., Hadiono, K., & Chasanah, I. (2020). Detecting Hoaxes in Indonesian News Using TF/TDM and K Nearest Neighbor. *IOP Conference Series: Materials Science and Engineering*, 835(1).

LAMPIRAN

Lampiran 1. Contoh Prediksi Data *Testing* Berita

Data Testing 1		Aktual	Prediksi
<p>Jakarta - Pemerintah melaporkan kasus harian positif Corona hari ini menembus 55.209 kasus. Jawa Barat (Jabar) menjadi wilayah yang melaporkan kasus tertinggi dalam 24 jam terakhir. Data mengenai kasus Corona ini dipublikasikan Humas BNPB, Sabtu (12/2/2022). Data kasus COVID-19 diperbarui setiap hari per pukul 12.00 WIB. Dengan penambahan 55.209, total kasus Corona di Indonesia sejak Maret 2020 hingga hari ini berjumlah 4.763.252. Selain itu, pemerintah melaporkan jumlah pasien sembuh dari Corona pada hari ini sebanyak 32.570 orang. Dengan demikian, total pasien sembuh dari Corona di Indonesia menjadi 4.282.847. Ada pula penambahan kasus kematian akibat Corona pada hari ini sebanyak 107 kasus, sehingga total kasus kematian akibat Corona menjadi 145.065. Sementara itu, berdasarkan sebaran kasus COVID-19, Jawa Barat masih melaporkan tambahan terbanyak, yakni 14.106. Kasus tersebut terdiri dari transmisi lokal sebanyak 14.060 dan pelaku perjalanan luar negeri sebanyak 46. Kemudian diikuti DKI Jakarta dengan 12.417 kasus Corona. Kasus tersebut terdiri atas 12.126 transmisi lokal dan 291 dari pelaku perjalanan luar negeri. Baca juga: Anies: 9.000 Lebih Anak Jakarta Ortunya Meninggal karena COVID-19 Berikut ini data sebaran kasus baru Corona hari ini: 1. Aceh: 462. Sumatera Utara: 9553. Sumatera Barat: 4004. Riau: 3515. Jambi: 896. Sumatera Selatan: 6397. Bengkulu: 1398. Lampung: 4889. Bangka Belitung: 15010. Kepulauan Riau: 21711. DKI Jakarta: 12.41712. Jawa Barat: 14.10613. Jawa Tengah: 3.10014. DI Yogyakarta: 1.06515. Jawa Timur: 5.88016. Banten: 7.28317. Bali: 2.32318. Nusa Tenggara Barat: 48319. Nusa Tenggara Timur: 15420. Kalimantan Barat: 35921. Kalimantan Tengah: 23522. Kalimantan Selatan: 61223. Kalimantan Timur: 77124. Kalimantan Utara: 3725. Sulawesi Utara: 42826. Sulawesi Tengah: 14927. Sulawesi Selatan: 81428. Sulawesi Tenggara: 19529. Gorontalo: 3930. Sulawesi Barat: 2731. Maluku: 31932. Maluku Utara: 2233. Papua: 61534. Papua Barat: 302</p> <p>Baca juga: Positif Corona RI 12 Februari Tambah 55.209, Dekati Puncak Gelombang Kedua Sebaran kasus sembuh Corona hari ini: 1. Aceh: 52. Sumatera Utara: 293. Sumatera Barat: 314. Riau: 385. Jambi: 136. Sumatera Selatan: 307. Bengkulu: 08. Lampung: 249. Bangka Belitung: 3710. Kepulauan Riau: 1011. DKI Jakarta: 23.89012. Jawa Barat: 1.81713. Jawa Tengah: 72114. DI Yogyakarta: 2515. Jawa Timur: 2.96616. Banten: 1.28217. Bali: 1.12818. Nusa Tenggara Barat: 3719. Nusa Tenggara Timur: 5620. Kalimantan Barat: 6721. Kalimantan Tengah: 1722. Kalimantan Selatan: 6123. Kalimantan Timur: 9524. Kalimantan Utara: 125. Sulawesi Utara: 3126. Sulawesi Tengah: 2027. Sulawesi Selatan: 3428. Sulawesi Tenggara: 829. Gorontalo: 030. Sulawesi Barat: 231. Maluku: 4832. Maluku Utara: 033. Papua: 034. Papua Barat: 47</p> <p>(aud/aud) sebaran kasus corona data sebaran kasus corona update kasus corona corona covid-19</p>		Fakta	Fakta
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga	
Jakarta - Pemerintah melaporkan kasus harian positif Corona hari ini sebanyak 40.489 kasus. DKI Jakarta menjadi wilayah yang melaporkan kasus tertinggi dalam 24 jam terakhir. Data mengenai kasus Corona ini dipublikasikan	0.04482	Fakta	

<p>Humas BNPB, Jumat (11/2/2022). Data kasus COVID-19 diperbarui setiap hari per pukul 12.00 WIB. Dengan penambahan 40.489, total kasus Corona di Indonesia sejak Maret 2020 hingga hari ini berjumlah 4.708.043. Selain itu, pemerintah melaporkan jumlah pasien sembuh dari Corona pada hari ini sebanyak 15.767 orang. Dengan demikian, total pasien sembuh dari Corona di Indonesia menjadi 4.250.277. Ada pula penambahan kasus kematian akibat Corona pada hari ini sebanyak 100 kasus, sehingga total kasus kematian akibat Corona menjadi 144.958. Baca juga: Kasus Corona RI 11 Februari Tambah 40.498, Pasien Meninggal Tembus 100 Sementara itu, berdasarkan sebaran kasus COVID-19, DKI Jakarta masih melaporkan tambahan terbanyak, yakni 10.707. Kasus tersebut terdiri dari transmisi lokal sebanyak 10.329 dan pelaku perjalanan luar negeri sebanyak 378. Kemudian diikuti Jawa Barat dengan 8.945 kasus Corona. Kasus tersebut terdiri atas 8.925 transmisi lokal dan 20 dari pelaku perjalanan luar negeri. Berikut ini data sebaran kasus baru Corona hari ini: 1. Aceh: 342. Sumatera Utara: 6123. Sumatera Barat: 1714. Riau: 3345. Jambi: 546. Sumatera Selatan: 4457. Bengkulu: 518. Lampung: 3179. Bangka Belitung: 12310. Kepulauan Riau: 13011. DKI Jakarta: 10.70712. Jawa Barat: 8.94513. Jawa Tengah: 2.65114. DI Yogyakarta: 61015. Jawa Timur: 4.50616. Banten: 5.21817. Bali: 1.83418. Nusa Tenggara Barat: 41219. Nusa Tenggara Timur: 9720. Kalimantan Barat: 22121. Kalimantan Tengah: 17222. Kalimantan Selatan: 54423. Kalimantan Timur: 35824. Kalimantan Utara: 3225. Sulawesi Utara: 37826. Sulawesi Tengah: 4727. Sulawesi Selatan: 43828. Sulawesi Tenggara: 13329. Gorontalo: 1530. Sulawesi Barat: 1131. Maluku: 23432. Maluku Utara: 1033. Papua: 48034. Papua Barat: 165</p> <p>Baca juga: Kasus Konser Musik Penuh Penonton di Makassar Naik Penyidikan Sebaran kasus sembuh Corona hari ini: 1. Aceh: 32. Sumatera Utara: 613. Sumatera Barat: 214. Riau: 335. Jambi: 86. Sumatera Selatan: 377. Bengkulu: 48. Lampung: 159. Bangka Belitung: 1910. Kepulauan Riau: 011. DKI Jakarta: 7.40012. Jawa Barat: 1.59013. Jawa Tengah: 69514. DI Yogyakarta: 4715. Jawa Timur: 2.86916. Banten: 1.91717. Bali: 54018. Nusa Tenggara Barat: 2719. Nusa Tenggara Timur: 1720. Kalimantan Barat: 14921. Kalimantan Tengah: 722. Kalimantan Selatan: 8123. Kalimantan Timur: 9724. Kalimantan Utara: 225. Sulawesi Utara: 226. Sulawesi Tengah: 727. Sulawesi Selatan: 3328. Sulawesi Tenggara: 329. Gorontalo: 130. Sulawesi Barat: 031. Maluku: 5432. Maluku Utara: 133. Papua: 034. Papua Barat: 27</p> <p>Simak Video 'Update Corona RI 11 Februari 2022: Bertambah 40.489 Kasus': [Gambas: Video 20detik](eva/idn)covid-19 kasus harian covid-19 kasus harian corona sebaran corona</p>		
<p>Jakarta - Sebanyak 25 RT di Jakarta Barat masuk zona merah Corona. Dari 25 RT tersebut, zona merah terbanyak di Kecamatan Kembangan. "(Terbanyak) di Kembangan Utara ada 3 RT, di Kembangan Selatan ada 3 RT, di Meruya Utara ada 3 RT dan di Srengseng ada 1 RT," ujar Plt Wakil Wali</p>	0.21772	Fakta

<p>Kota Jakarta Barat, Iin Mutmainnah kepada wartawan, Jumat (11/2/2022). Dari data yang diterima detikcom, tercatat 25 RT masuk dalam zona merah. Data tersebut berdasarkan pada 6 Februari 2022. Baca juga: Kasus Aktif COVID-19 di Jakut Tembus 15.310, Terbanyak di Priok. Sebanyak 25 RT tersebut tersebar di 13 kelurahan yang ada di Jakarta Barat. Diantaranya, Kalideres, Pegadungan, Cengkareng Timur, Cengkareng Barat, Kapuk, Duri Kosambi, Kembangan Utara, Kembangan Selatan, Meruya Utara, Srengseng, Kota Bambu Selatan, Kemanggisan dan Palmerah. Dari keseluruhan, tercatat 221 jiwa terkonfirmasi positif COVID. Sedangkan jumlah rumah yang terpapar ada sekitar 183 rumah. Berikut sebaran 25 RT zona merah di Jakarta Barat: 1. Kelurahan Kalideres: terdapat 1 RT zona merah, yaitu RT 06/ RW 172. Kelurahan Pegadungan: terdapat 1 RT zona merah, yaitu RT 02/ RW 193. Kelurahan Cengkareng Timur: terdapat 3 RT zona merah, yaitu RT 07/ RW 14, RT 09/ RW 14 dan RT 06/ RW 144. Kelurahan Kapuk: terdapat 2 RT zona merah, yaitu RT 05/ RW 16 dan RT 05/ RW 145. Kelurahan Cengkareng Barat: terdapat 1 RT zona merah, yaitu RT 08/ RW 056. Kelurahan Duri Kosambi: terdapat 1 RT zona merah, yaitu RT 01/ RW 057. Kelurahan Kembangan Utara: terdapat 3 RT zona merah, yaitu RT 13/ RW 09, RT 14/ RW 09 dan RT 07/ RW 108. Kelurahan Kembangan Selatan: terdapat 3 RT zona merah, yaitu RT 10/ RW 01, RT 01/ RW 02 dan RT 10/ RW 039. Kelurahan Meruya Utara: terdapat 3 RT zona merah, yaitu RT 02/ RW 04, RT 23/ RW 08 dan RT 08/ RW 0510. Kelurahan Srengseng terdapat 3 RT zona merah, yaitu RT 07/ RW 06 dan RT 08/ RW 0711. Kelurahan Kota Bambu Selatan: terdapat 2 RT zona merah, yaitu rusun di RW 01 dan RT 11/ RW 0612. Kelurahan Kemanggisan: terdapat 1 RT zona merah, yaitu RT 11/ RW 0113. Kelurahan Palmerah: terdapat 2 RT zona merah, yaitu RT 05/ RW 12 dan RT 07/ RW 12. Baca juga: Wagub: 14 RT di Jakarta Zona Merah, 216 Zona Oranye. Sebanyak 25 RT di Jakarta Barat masuk zona merah kasus COVID-19. Status zona merah ini pun memaksa Pemkot Jakarta Barat menerapkan micro-lockdown di 25 RT tersebut. "Menerapkan micro-lockdown terhadap 25 RT zona merah sesuai dengan data corona.jakarta.go.id," ujar Plt Wakil Wali Kota Jakarta Barat, Iin Mutmainnah, saat dihubungi, Kamis (10/2). Pelaksanaan micro-lockdown akan dilakukan hingga 14 Februari mendatang. Tenggat tersebut sesuai dengan ketentuan PPKM level 3. Baca juga: 25 RT di Jakbar Masuk Zona Merah Corona, Micro-lockdown Diterapkan. Simak video 'Satgas Klaim Stok Oksigen Aman di Tengah Gelombang 3 Corona': [Gambas: Video 20 detik] (ain/zap)jabodetabek jakarta barat zona merah corona covid-19 jakarta</p>		
<p>Jakarta - Camat Kembangan Joko menyebut RT yang menjadi zona merah Corona atau COVID-19 di wilayahnya telah berkurang. Dia menyebut ada empat RT yang sebelumnya zona merah, kini sudah menjadi zona oranye penyebaran Corona. "Kecamatan Kembangan sesuai publish data minggu ini dari Sudin Kesehatan Jakarta Barat adalah</p>	0.27012	Fakta

<p>RT dengan zona merah terbanyak karena data awalnya jumlah RT zona merahnya berjumlah 11 kemudian setelah diverifikasi ternyata hanya tujuh RT," ujar Joko kepada wartawan, Jumat (11/2/2022).Baca juga: 25 RT di Jakbar Zona Merah Corona, Terbanyak di Kecamatan Kembangan</p> <p>Dia mengatakan terdapat warga isolasi mandiri (isoman) yang sudah dinyatakan negatif Corona pada Rabu (10/2). Dia mengatakan berkurangnya jumlah warga yang menjalani isolasi mandiri membuat zona merah berkurang. "Jadi setelah dicek lapangan oleh Tim Satgas Kelurahan masing-masing, ternyata ada beberapa warga isoman yang sudah sembuh dan sudah negatif, sehingga tidak masuk kriteria zona merah lagi tetapi menjadi zona oranye," jelasnya.Dia menyebut empat RT tersebut sudah bebas dari micro-lockdown. Aktivitas warga juga telah kembali normal.Empat RT yang telah terbebas dari zona merah terdiri dari RT 14 RW 09 dan RT 07 RW 10 di Kelurahan Kembangan Utara. Kemudian, RT 01 RW 02 di Kembangan Selatan dan RT 07 RW 06 di Kelurahan Srengseng.Sedangkan tujuh RT yang masih menjadi zona merah Corona terdiri dari tiga RT di Meruya Utara, dua RT di Kembangan Selatan, satu RT di Kembangan Utara, dan satu RT di Srengseng.Baca juga: Ada 14 Kasus COVID di Apartemen Belmont Jakbar, Micro-lockdown Diperluas"Jumlah keseluruhan RT di Kembangan terdiri dari 619 RT dengan status 254 RT masih menjadi zona hijau, 251 RT masuk zona kuning, 107 RT masuk zona oranye dan tujuh RT di antaranya masih zona merah," ucapnya.Berikut sebaran tujuh RT zona merah di Kembangan, Jakarta Barat:1. Kelurahan Kembangan Utara: RT 13 RW 092. Kelurahan Kembangan Selatan: RT 11 RW 01 dan RT 10 RW 033. Kelurahan Meruya Utara: RT 02 RW 04, RT 23 RW 08 dan RT 08 RW 054. Kelurahan Srengseng: RT 08 RW 07Baca juga: Pusat Isoter Masjid Hasyim Asy'ari Jakbar Mulai Diisi Pasien COVID(ain/haf)jakarta barat jabodetabek covid-19 virus corona jakarta</p>		
---	--	--

Data Testing 2		Aktual	Prediksi
<p>Beredar postingan video di Twitter oleh akun @ThePr0diga1Son memposting sebuah video yang memperlihatkan sebanyak 10.000 warga turun ke jalan untuk melakukan protes. Postingan @ThePr0diga1Son juga disertai narasi bahwa warga melakukannya dalam rangka memprotes tirani Covid-19 dan pemaksaan vaksinasi.Postingan @ThePr0diga1Son beredar di tengah peristiwa unjuk rasa ribuan warga Belanda untuk menentang pembatasan kegiatan karena Covid-19 dan protes atas mandatisasi vaksin Covid-19. Protes tersebut berlangsung di tengah naiknya kasus Covid-19 di Belanda.</p>		<i>Hoax</i>	<i>Hoax</i>
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga	
<p>Beredar sebuah narasi oleh akun Facebook KarmilaSari UmmAhmad IbnuSholeh yang mengatakan bahwa campuran daun pepaya dan jahe dapat menjadi obat varian Omicron. Narasi tersebut juga mencantumkan tata cara membuat</p>	0.11876	<i>Hoax</i>	

campuran daun pepaya dan jahe agar dapat dijadikan sebagai obat varian Omicron.		
Pesan berantai yang mengajak warga untuk menyetop mengirim berita tentang Covid-19 melalui media sosial, beredar di aplikasi pesan Whatsapp, 13 Juli 2021. Pesan berantai ini beredar di tengah kondisi lonjakan jumlah pasien Covid-19 di Indonesia. Berita tentang Covid-19 dianggap menurunkan imun sehingga warga mudah terpapar penyakit. Larangan mengunggah berita Covid-19 di medsos, diklaim dilakukan oleh sejumlah negara seperti Timor Leste, Brunei Darussalam, Singapura, Malaysia, Australia, termasuk Cina. Berikut ini isi pesan berantai tersebut: Supaya Covid tdk berkembang, kita STOP kirim berita ttg Covid. Seperti yg dilakukan oleh Timor Leste, Brunei Darussalam, Singapura, Malaysia, Australia, dan beberapa negara di Asia lainnya, termasuk Cina. Negara ² tersebut melarang warga negaranya berkirim berita tentang Covid-19 melalui MEDSOSMARI MULAI KITA TIRU DAN LAKUKAN STOP BERITA COVID. Yang wajib kita jaga adlh: IMAN-IMUN-AMIN dan patuhi Protokol kesehatan. Abaikan berita dan jangan sebar berita Covid yg bikin resah, semakin kita resah, semakin mudah terpapar penyakit, apapun penyakitnya. Kita galang persatuan melawan Covid dengan cara tersebut. Kasihanilah bagi yg imunnya lemah akan menambah, stress...itu salah satu sebab mudahnya terkena penyakit.	0.13403	Hoax
Sejumlah tangkapan layar dan video dengan narasi 3 ambulans di Ngemplak, Sukoharjo, ugal-ugalan menabrak mobil, beredar di media sosial dalam sepekan terakhir. Unggahan ini beredar di tengah lonjakan kasus Covid-19 di Indonesia dan penerapan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) Darurat di Jawa-Bali. Tangkapan layar dan video yang diunggah pada 7 Juli 2021 di Facebook ini diklaim bahwa ambulans tersebut kosong dan diminta berputar-putar dengan upah Rp 300 ribu untuk menakuti warga. "Ternyata mobil ambulans yang sering muter-muter di suatu wilayah itu kosong, tidak ada pasien/jenazah, dengan maksud tujuan menakuti atau membuat warga panik, dan percaya kalo banyak korban berjatuh akibat Covid," tulis narasi itu. Sementara dalam video yang beredar, terdapat suara seorang pria yang mengatakan, "Ambulans gak ono penumpang (gak ada penumpangnya) do ugal-ugalan mlakune (jalannya ugal-ugalan). Ambulans telu, sing siji lungu (ambulans tiga, yang satu pergi). Ning Ngemplak Gang 3, tiga beruntun (ambulans) jadinya empat," Tangkapan layar video yang diklaim sebagai ambulans kosong ugal-ugalan dan berputar-putar untuk menakuti warga agar percaya Covid-19.	0.13497	Hoax

Data Testing 3	Aktual	Prediksi
Jakarta - Sebanyak 25 RT di Jakarta Barat masuk zona merah kasus COVID-19. Status zona merah ini pun memaksa Pemkot Jakarta Barat menerapkan micro-lockdown di 25 RT tersebut. "Menerapkan micro-lockdown terhadap 25 RT zona merah sesuai dengan data	Fakta	Fakta

<p>corona.jakarta.go.id," ujar Plt Wakil Wali Kota Jakarta Barat, Iin Mutmainnah, saat dihubungi, Kamis (10/2/2022).Baca juga: Sebaran 40.618 Corona RI 10 Februari: DKI Tembus 11.090, Jabar 9.403Iin mengimbau masyarakat agar mengantisipasi penyebaran virus COVID-19. Dengan cara tertib melaksanakan 6M dan 3T. "Kita juga turunkan Satgas COVID ini berkaitan dengan perda tentang COVID. Jadi nanti tim Satgas COVID dari jajaran terkait, Satpol PP bersama jajaran wilayah itu bersama-sama untuk melakukan tertib prokes," kata Iin.Pelaksanaan micro-lockdown akan dilakukan hingga 14 Februari mendatang. Tenggat tersebut sesuai dengan ketentuan PPKM level 3."Adapun kebutuhan dari warga isoman karena sebagian besar warga isomannya tidak berada di rumah, maksudnya ada yang di apart, juga tersebar di beberapa rumah di lingkungan RT tersebut," jelas Iin."Jadi prinsipnya Pak Lurah sudah komunikasi dengan RT dengan RW terkait kebutuhan warga isoman, bisa dilakukan RT atau RW. Jika membutuhkan bantuan kelurahan nanti akan dikoordinasikan dengan sudin terkait," sambungnya.Diberitakan sebelumnya, Wakapolres Metro Jakarta Barat AKBP Bismo Teguh Prakoso mengatakan zona merah di wilayahnya tak sampai 37 titik. Bismo mengatakan sejumlah pasien sudah sembuh dari COVID-19.Baca juga: Vaksin 1 dan 2 Sinovac Boosternya Apa? Ketahui Lagi Jenis-jenisnya" Ini sesuai aplikasi itu ada 37 zona merah. Tapi sudah divalidasi oleh para Kapolsek kemudian dengan para lurahnya, pukesmas dan rekan-rekan Konramil itu ternyata ada yang sudah lepas dari zona merah dan sudah sembuh. Sehingga berkurang dari jumlah sesuai aplikasi itu. Jadi fakta di lapangan kurang dari 37 demikian," jelas Bismo, kepada wartawan, Selasa (8/2/2022).Dia menyebut titik zona merah berada di wilayah Kembangan, Kalideres, Cengkareng, dan Palmerah. Bismo mengatakan aplikasi tersebut membantu memantau kasus COVID di wilayah Jakarta Barat.Sebaran data kasus aktif COVID-19 di Jakarta Barat berdasarkan situs corona.jakarta.go.id pada Selasa (10/2/2022):Tegal Alur: 445 jiwaKembangan Utara: 447 jiwaDuri Kosambi: 466 jiwaMeruya Utara: 473 jiwaKedoya Utara: 480 jiwaKalideres: 522 jiwaCengkareng Barat: 602 jiwaPalmerah: 614 jiwaCengkareng Timur: 617 jiwaKebon Jeruk: 649 jiwaDuri Kepa: 670 jiwaKapuk: 705 jiwaSimak juga 'Jakpus Episentrum Covid-19 Jabodetabek, Satgas Sentil Perkantoran':[Gambas:Video 20detik](ain/dwia)pemkot jakarta barat jakarta barat covid-19 micro-lockdown micro lockdown jabodetabek</p>		
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga
<p>Jakarta - Sebanyak 25 RT di Jakarta Barat masuk zona merah Corona. Dari 25 RT tersebut, zona merah terbanyak di Kecamatan Kembangan."(Terbanyak) di Kembangan Utara ada 3 RT, di Kembangan Selatan ada 3 RT, di Meruya Utara ada 3 RT dan di Srengseng ada 1 RT," ujar Plt Wakil Wali Kota Jakarta Barat, Iin Mutmainnah kepada wartawan, Jumat (11/2/2022).Dari data yang diterima detikcom, tercatat 25 RT masuk dalam zona merah. Data tersebut berdasarkan pada 6 Februari 2022. Baca juga: Kasus Aktif COVID-19 di Jakut Tembus 15.310, Terbanyak di PriokSebanyak 25 RT tersebut tersebar di 13 kelurahan yang ada di Jakarta Barat. Diantaranya, Kalideres, Pegadungan, Cengkareng Timur, Cengkareng Barat, Kapuk, Duri Kosambi, Kembangan Utara, Kembangan Selatan, Meruya Utara, Srengseng, Kota Bambu Selatan, Kemanggisan dan</p>	0.21864	Fakta

<p>Palmerah. Dari keseluruhan, tercatat 221 jiwa terkonfirmasi positif COVID. Sedangkan jumlah rumah yang terpapar ada sekitar 183 rumah. Berikut sebaran 25 RT zona merah di Jakarta Barat: 1. Kelurahan Kalideres: terdapat 1 RT zona merah, yaitu RT 06/ RW 172. Kelurahan Pegadungan: terdapat 1 RT zona merah, yaitu RT 02/ RW 193. Kelurahan Cengkareng Timur: terdapat 3 RT zona merah, yaitu RT 07/ RW 14, RT 09/ RW 14 dan RT 06/ RW 144. Kelurahan Kapuk: terdapat 2 RT zona merah, yaitu RT 05/ RW 16 dan RT 05/ RW 145. Kelurahan Cengkareng Barat: terdapat 1 RT zona merah, yaitu RT 08/ RW 056. Kelurahan Duri Kosambi: terdapat 1 RT zona merah, yaitu RT 01/ RW 057. Kelurahan Kembangan Utara: terdapat 3 RT zona merah, yaitu RT 13/ RW 09, RT 14/ RW 09 dan RT 07/ RW 108. Kelurahan Kembangan Selatan: terdapat 3 RT zona merah, yaitu RT 10/ RW 01, RT 01/ RW 02 dan RT 10/ RW 039. Kelurahan Meruya Utara: terdapat 3 RT zona merah, yaitu RT 02/ RW 04, RT 23/ RW 08 dan RT 08/ RW 0510. Kelurahan Srengseng terdapat 3 RT zona merah, yaitu RT 07/ RW 06 dan RT 08/ RW 0711. Kelurahan Kota Bambu Selatan: terdapat 2 RT zona merah, yaitu rusun di RW 01 dan RT 11/ RW 0612. Kelurahan Kemanggisan: terdapat 1 RT zona merah, yaitu RT 11/ RW 0113. Kelurahan Palmerah: terdapat 2 RT zona merah, yaitu RT 05/ RW 12 dan RT 07/ RW 12. Baca juga: Wagub: 14 RT di Jakarta Zona Merah, 216 Zona Oranye. Sebanyak 25 RT di Jakarta Barat masuk zona merah kasus COVID-19. Status zona merah ini pun memaksa Pemkot Jakarta Barat menerapkan micro-lockdown di 25 RT tersebut. "Menerapkan micro-lockdown terhadap 25 RT zona merah sesuai dengan data corona.jakarta.go.id," ujar Plt Wakil Wali Kota Jakarta Barat, Iin Mutmainnah, saat dihubungi, Kamis (10/2). Pelaksanaan micro-lockdown akan dilakukan hingga 14 Februari mendatang. Tenggat tersebut sesuai dengan ketentuan PPKM level 3. Baca juga: 25 RT di Jakbar Masuk Zona Merah Corona, Micro-lockdown Diterapkan. Simak video 'Satgas Klaim Stok Oksigen Aman di Tengah Gelombang 3 Corona': [Gambas: Video 20detik] (ain/zap) jabodetabek jakarta barat zona merah corona covid-19 jakarta</p>		
<p>Jakarta - Pemerintah melaporkan kasus harian positif Corona hari ini sebanyak 40.489 kasus. DKI Jakarta menjadi wilayah yang melaporkan kasus tertinggi dalam 24 jam terakhir. Data mengenai kasus Corona ini dipublikasikan Humas BNPB, Jumat (11/2/2022). Data kasus COVID-19 diperbarui setiap hari per pukul 12.00 WIB. Dengan penambahan 40.489, total kasus Corona di Indonesia sejak Maret 2020 hingga hari ini berjumlah 4.708.043. Selain itu, pemerintah melaporkan jumlah pasien sembuh dari Corona pada hari ini sebanyak 15.767 orang. Dengan demikian, total pasien sembuh dari Corona di Indonesia menjadi 4.250.277. Ada pula penambahan kasus kematian akibat Corona pada hari ini sebanyak 100 kasus, sehingga total kasus kematian akibat Corona menjadi 144.958. Baca juga: Kasus Corona RI 11 Februari Tambah 40.498, Pasien</p>	0.22998	Fakta

<p>Meninggal Tembus 100 Sementara itu, berdasarkan sebaran kasus COVID-19, DKI Jakarta masih melaporkan tambahan terbanyak, yakni 10.707. Kasus tersebut terdiri dari transmisi lokal sebanyak 10.329 dan pelaku perjalanan luar negeri sebanyak 378. Kemudian diikuti Jawa Barat dengan 8.945 kasus Corona. Kasus tersebut terdiri atas 8.925 transmisi lokal dan 20 dari pelaku perjalanan luar negeri. Berikut ini data sebaran kasus baru Corona hari ini: 1. Aceh: 342. Sumatera Utara: 6123. Sumatera Barat: 1714. Riau: 3345. Jambi: 546. Sumatera Selatan: 4457. Bengkulu: 518. Lampung: 3179. Bangka Belitung: 12310. Kepulauan Riau: 13011. DKI Jakarta: 10.70712. Jawa Barat: 8.94513. Jawa Tengah: 2.65114. DI Yogyakarta: 61015. Jawa Timur: 4.50616. Banten: 5.21817. Bali: 1.83418. Nusa Tenggara Barat: 41219. Nusa Tenggara Timur: 9720. Kalimantan Barat: 22121. Kalimantan Tengah: 17222. Kalimantan Selatan: 54423. Kalimantan Timur: 35824. Kalimantan Utara: 3225. Sulawesi Utara: 37826. Sulawesi Tengah: 4727. Sulawesi Selatan: 43828. Sulawesi Tenggara: 13329. Gorontalo: 1530. Sulawesi Barat: 1131. Maluku: 23432. Maluku Utara: 1033. Papua: 48034. Papua Barat: 165</p> <p>Baca juga: Kasus Konser Musik Penuh Penonton di Makassar Naik Penyidikan Sebaran kasus sembuh Corona hari ini: 1. Aceh: 32. Sumatera Utara: 613. Sumatera Barat: 214. Riau: 335. Jambi: 86. Sumatera Selatan: 377. Bengkulu: 48. Lampung: 159. Bangka Belitung: 1910. Kepulauan Riau: 011. DKI Jakarta: 7.40012. Jawa Barat: 1.59013. Jawa Tengah: 69514. DI Yogyakarta: 4715. Jawa Timur: 2.86916. Banten: 1.91717. Bali: 54018. Nusa Tenggara Barat: 2719. Nusa Tenggara Timur: 1720. Kalimantan Barat: 14921. Kalimantan Tengah: 722. Kalimantan Selatan: 8123. Kalimantan Timur: 9724. Kalimantan Utara: 225. Sulawesi Utara: 226. Sulawesi Tengah: 727. Sulawesi Selatan: 3328. Sulawesi Tenggara: 329. Gorontalo: 130. Sulawesi Barat: 031. Maluku: 5432. Maluku Utara: 133. Papua: 034. Papua Barat: 27</p> <p>Simak Video 'Update Corona RI 11 Februari 2022: Bertambah 40.489 Kasus': [Gambas: Video 20detik] (eva/idn) covid-19 kasus harian covid-19 kasus harian corona sebaran corona</p>		
<p>Jakarta - Suku Dinas Kesehatan Jakarta Barat melaporkan bed occupancy rate (BOR) atau keterisian tempat tidur di rumah sakit rujukan COVID-19 sudah mencapai 75 persen. Sedangkan BOR ICU 61%. Angka ini merujuk data jumlah tempat tidur dan pasien ICU di 23 RS rujukan COVID-19 Jakarta Barat. Data itu tercatat per Kamis (10/2) kemarin. "Keterisian isolasi mencapai 75% sedangkan keterisian ICU sebesar 61%," ujar Kasi Pencegahan dan Pengendalian Penyakit Sudinkes Jakarta Barat, dr Arum Ambarsari, saat dimintai konfirmasi, Jumat (11/2/2022). Baca juga: 2 RT di Tanjung Duren Jakbar Masuk Zona Oranye, Vaksin Digencarkan Arum mengatakan total ada 23 RS rujukan COVID di Jakarta. Dari total kapasitas 970 tempat tidur, kini sudah terisi 731. Sedangkan untuk bed ICU disediakan 144. Selain itu, Arum menyebutkan saat ini Masjid KH Hasyim Asy'ari mulai menerima pasien COVID-</p>	0.23683	Fakta

19 untuk isolasi. Sudah ada 11 pasien yang diisolasi di masjid tersebut. Diberitakan sebelumnya, tempat isolasi terpusat pasien Corona atau COVID-19 Masjid KH Hasyim Asy'ari, Jalan Daan Mogot, Jakarta Barat (Jakbar), mulai diisi pasien Corona. Ada dua pasien yang mulai menjalani isolasi di lokasi ini. "Jadi hari ini ada dua pasien, satu pasien sejak Selasa, ini satu baru masuk lagi," ujar Kepala UPT Masjid Hasyim Asy'ari Dikki Syafrin saat dihubungi, Kamis (10/2). Dikki menyebut pasien yang baru masuk berasal dari Puskesmas Kembangan, Jakarta Barat. Keduanya merupakan orang tanpa gejala (OTG). Baca juga: Zona Merah di Kembangan Jakbar Berkurang Jadi 7 RT (ain/idn)jakarta barat dinas kesehatan corona keterisian tempat tidur jabodetabek jakarta		
--	--	--

Data Testing 4		Aktual	Prediksi
Video pendek yang berisi klaim bahwa anak-anak kebal terhadap virus Corona penyebab Covid-19, SARS-CoV-2, beredar di Instagram. Menurut perempuan dalam video itu, yang mengklaim dirinya sebagai peneliti, anak-anak yang meninggal dalam setahun terakhir juga tidak terkait dengan Covid-19. Video tersebut juga mempromosikan agar publik menolak vaksin Covid-19 karena dianggap sebagai genosida. "Masihkah vaksin harus dipaksakan untuk masyarakat? Harus berapa banyak anak-anak mati karena vaksin dan membuat pemerintah sadar bahwa vaksin harus dihentikan?" demikian teks yang tertulis dalam video itu. Sementara perempuan dalam video tersebut berkata, "Virus ini tidak berpengaruh terhadap anak-anak. Mereka kebal terhadap virus ini. Dua ratus lebih anak meninggal dalam setahun, dan tidak ada hubungannya dengan Covid-19. Hanya karena 200 lebih anak meninggal, kalian ingin menyuntikkan vaksin kepada anak-anak yang lain? Akuilah bahwa ini merupakan sebuah pembunuhan besar-besaran, genosida." Akun ini membagikan video tersebut pada 30 Mei 2021. Akun itu menulis, "Masihkah kita mau dipermainkan, dijadikan kelinci percobaan, manggut-manggut aja disuruh ini-itu. Kita ini manusia berakal, bukan kawanan hewan ternak." Gambar tangkapan layar unggahan di Instagram yang berisi video dengan klaim keliru terkait penularan Covid-19 terhadap anak-anak.		<i>Hoax</i>	<i>Hoax</i>
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga	
Video pendek yang berisi klaim bahwa vaksinasi Covid-19 hanya percobaan beredar di Instagram. Menurut klaim yang dilontarkan oleh seorang pria itu, vaksinasi Covid-19 yang dilakukan di seluruh dunia saat ini sebenarnya hanyalah sebuah uji klinis karena vaksin Covid-19 yang digunakan saat ini mengantongi izin penggunaan darurat saja. Berikut pernyataan pria tersebut: "Kenapa vaksinnya tidak disetujui oleh FDA (Badan Pengawas Obat dan Makanan Amerika Serikat)? Vaksin butuh bertahun-tahun untuk dikembangkan. Dan butuh paling tidak 2-3 tahun untuk uji klinis. Jadi, yang sekarang ini, 'vaksinasi' akan jadi uji klinis sampai Januari 2023. Semua yang mendapatnya sekarang ada dalam uji klinis, bukan dalam masa penggunaan obat yang telah disetujui. Yang kita punya adalah hak penggunaan darurat. Hak penggunaan darurat dapat disetujui untuk 'vaksin' dalam	0.04154	<i>Hoax</i>	

<p>kedaruratan kesehatan publik. Begitulah, ini sebuah percobaan."Akun ini mengunggah video itu pada 3 April 2021. Akun tersebut juga menulis, "Kenapa vaksin tidak diwajibkan di Amerika dan negara Eropa lain? Karena belum ada satu pun vaksin yang lolos uji klinis, dan hanya punya Izin Penggunaan Darurat (Emergency Use Authorization). Lalu, kenapa di negara antah-berantah seolah-olah 'diwajibkan'?"Gambar tangkapan layar unggahan di Instagram yang berisi klaim sesat terkait vaksinasi Covid-19.</p>		
<p>Video berdurasi 9 menit yang berisi pernyataan dari aktor Russell Brand beredar di Instagram. Menurut klaim yang menyertai video ini, Brand mengomentari tentang bocornya ribuan email milik Direktur Institut Nasional Alergi dan Penyakit Menular Amerika Serikat Anthony Fauci. Menurut klaim itu, Brand menyatakan bahwa, dalam ribuan email Fauci, ditemukan indikasi bahwa mantan kepala penasihat medis Gedung Putih itu terlibat dalam pembiayaan terciptanya Covid-19 di Laboratorium Wuhan, Cina. Brand juga menyinggung komunikasi intens antara Fauci dan Mark Zuckerberg, pendiri Facebook. Dalam komunikasi itu, dibahas tentang konten apa saja yang harus disensor di Facebook, Instagram, dan YouTube. Akun ini mengunggah video beserta klaim tersebut pada 7 Juni 2021. Di akhir keterangannya, akun itu menulis, "Siapa yang sebenarnya yang waras atau tak waras dalam menyikapi plandemi? Mereka yang selalu disensor namun akhirnya terbukti benar?" Hingga kini, video tersebut telah ditonton 2.283 kali. Gambar tangkapan layar video milik aktor Russel Brand yang berisi klaim keliru terkait email Direktur Institut Nasional Alergi dan Penyakit Menular Amerika Serikat Anthony Fauci.</p>	0.05026	<i>Hoax</i>
<p>Pesan berantai yang berisi klaim bahwa, usai vaksinasi Covid-19, tubuh justru lebih rentan terinfeksi virus Corona beredar Facebook. Karena itu, menurut pesan tersebut, setelah disuntik vaksin Covid-19, penerima vaksin dianjurkan untuk tidak banyak beraktivitas secara berat dan tidak pergi keluar rumah. Akun ini membagikan pesan berantai itu pada 12 Maret 2021. Menurut pesan tersebut, usai vaksinasi Covid-19, imunitas tubuh belum terbentuk dengan sempurna. Antibodi baru terbentuk secara sempurna dua pekan setelah vaksinasi dosis kedua. "Ini ada beberapa lansia di Surabaya yang kena Covid-19 setelah divaksin. Enggak mau istirahat. Karena merasa sudah aman, lalu keluyuran keluar," demikian yang tertulis dalam pesan itu. Pesan berantai ini pun menyinggung bahwa vaksin Covid-19 dosis kedua harus diberikan 21-28 hari setelah vaksinasi dosis pertama. "Vaksin kedua makan waktu kira-kira 14-21 hari baru jadi. Jadi, hitung-hitung dari vaksin dosis pertama ke vaksin dosis kedua sampai kekebalan terbangun itu harus menunggu sekitar dua bulan. Baru 85-92 persen kebal Covid-19." Gambar tangkapan layar pesan berantai yang beredar di Facebook yang berisi klaim menyesatkan terkait vaksinasi Covid-19.</p>	0.06074	<i>Hoax</i>

Data Testing 5		Aktual	Prediksi
<p>Jakarta - Satgas Penanganan COVID-19 mengungkapkan provinsi di Pulau Jawa dan Bali mendominasi jumlah kasus nasional selama 3 minggu. Satgas menyebut kenaikan kasus Corona di Provinsi Banten menjadi yang tercepat dengan 620 kali lipat dalam 6 minggu. "Data menunjukkan provinsi di Jawa dan Bali konsisten mendominasi jumlah kasus nasional setidaknya selama 3 minggu. Per tanggal 6 Februari, kenaikan kasus mingguan nasional masih disumbangkan provinsi Jawa-Bali," kata Juru Bicara Satgas Penanganan COVID-19 Wiku Adisasmito dalam konferensi pers virtual di kanal YouTube Setpres, Kamis (10/2/2022). Wiku memaparkan DKI Jakarta menjadi provinsi penyumbang terbanyak untuk kasus Corona nasional. Dia menyebut kasus COVID-19 di DKI naik 138 kali lipat. "Pertama DKI Jakarta menyumbangkan 42 persen kasus nasional. Kasus di DKI Jakarta naik 138 kali lipat dalam 6 minggu," ucap Wiku. Baca juga: Satgas: Laju Penularan Corona di Jakpus Paling Tinggi di Jabodetabek Berikut data persentase sumbangan masing-masing provinsi di Jawa-Bali untuk kasus nasional: 1. DKI Jakarta sumbang 42 persen kasus nasional, naik 138 kali lipat dalam 6 minggu. 2. Jawa Barat sumbang 14,31 persen kasus nasional, naik 336 kali lipat dalam 6 minggu. 3. Banten sumbang 14,31 persen kasus nasional, naik 620 kali lipat dalam 6 minggu. 4. Jawa Timur sumbang 5 persen kasus nasional, naik 83 kali lipat dalam 6 minggu. 5. Bali sumbang 5 persen kasus nasional, naik 392 kali lipat dalam 6 minggu. 6. Jawa Tengah sumbang 3 persen kasus nasional, naik 67 kali lipat dalam 6 minggu. 7. DI Yogyakarta sumbang 1 persen kasus nasional, naik 51 kali lipat dalam 6 minggu. (zak/lir) kenaikan kasus corona covid-19 satgas penanganan covid-19 jawa bali banten</p>		Fakta	Fakta
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga	
<p>Jakarta - Indonesia saat ini mengalami lonjakan kasus Corona. Satgas COVID-19 mengatakan bahwa ketersediaan oksigen di beberapa daerah saat ini masih mencukupi. "Sampai saat ini ketersediaan oksigen dalam bentuk oxygen concentrator dan oxygen generator masih mencukupi atau dapat memenuhi pelayanan lebih dari 12 sampai dengan 48 jam," kata Juru Bicara Satgas COVID-19 Wiku Adisasmito dalam jumpa pers melalui kanal YouTube BPNB, Kamis (10/2/2022). Baca juga: Menyelam Cari Aki Tercebur ke Laut, Pekerja Galangan Kapal di Banten Hilang Wiku menambahkan bahwa ketersediaan oksigen di DKI Jakarta masih mencukupi. Begitu pula ketersediaan oksigen di wilayah Banten. "Sebagai contoh di DKI Jakarta menurut dari per 6 Februari 2022, terdapat 1.541 oxygen concentrator dan 2 oxygen generator. Sementara di Banten ada 389 oxygen concentrator dan 4 oxygen generator," jelasnya. Di Jawa Barat, Wiku menjelaskan bahwa lebih dari seribu oxygen concentrator masih tersedia. Dia juga memastikan ketersediaan oksigen di daerah lainnya yang mengalami lonjakan kasus masih mencukupi. "Sedangkan di Jawa Barat terdapat 1.545 oxygen concentrator dan 18 oxygen generator. Kecukupan yang sama juga telah dipastikan pada daerah lainnya, terutama di daerah dengan kenaikan kasus yang tergolong tinggi," sebut Wiku. Baca juga: Isoter Rusunawa Kemiri Boyolali Dijadikan RS Darurat COVID-19 Lebih lanjut, Wiku mengatakan bahwa</p>	0.06560	Fakta	

<p>pemerintah pusat akan terus berkomunikasi dengan daerah dalam penanganan virus Corona. Komunikasi itu dilakukan guna memastikan ketersediaan fasilitas kesehatan."Ke depannya pemerintah akan berkoordinasi dengan pemerintah daerah untuk melakukan pendataan secara real time dalam memenuhi tiap kebutuhan fasilitas kesehatan secara efektif," tutur dia. Baca juga: Kena Omicron Gejala Ringan? Deretan 'Obat Warung' Ini Bisa untuk Bekal Isoman(lir/zak)ketersediaan oksigen oksigen stok oksigen satgas covid-19 wiku adisasmito</p>		
<p>Jakarta - Satgas COVID-19 mengungkap rasio testing virus Corona (COVID-19) di Indonesia dalam 1 minggu terakhir. Satgas menyebut ada 22 provinsi berada di bawah rasio tes Corona nasional, di antaranya Jawa Tengah dan Jawa Timur."Per 6 Februari 2022, angka nasional untuk rasio testing per 100 penduduk masih jauh di atas standar WHO. Saat ini rasio testing nasional adalah 7 dari seribu orang, sudah melebihi standar WHO, yaitu 1 dari 1.000 orang," kata Juru Bicara Satgas COVID-19 Wiku Adisasmito dalam siaran YouTube BNPB, Kamis (10/2/2022). Baca juga: Satgas Respons Kru MotoGP di Sirkuit Mandalika yang Bayar PCR Sampai Rp 6 Juta Namun, menurut Wiku, rasio testing COVID-19 di Jawa Tengah dan Jawa Timur masih rendah. Dia menyebut rasio testing kedua provinsi itu paling rendah di Jawa dan Bali. "Meskipun demikian, masih 22 provinsi yang rasio testing-nya di bawah rasio testing nasional. Terlebih pula, 2 dari provinsi Jawa-Bali, yaitu Jawa Timur dan kemudian Jawa Tengah termasuk dalam kelompok rasio testing masih di bawah angka nasional," kata dia. "Rasio testing di Jawa Timur adalah lima orang dari seribu dites. Sementara di Jawa Tengah angkanya tergolong rendah, yaitu angkanya 2 dari seribu orang dites," lanjutnya. Baca juga: Kasus Covid-19 Dunia Turun 17 Persen, Kematian Naik 7 Persen Wiku meminta 22 kepala daerah dimaksud untuk meningkatkan jumlah tes Corona. Wiku mewanti-wanti kasus yang dilaporkan tak sesuai dengan kondisi yang ada di lapangan. "Untuk itu, sama mohon kepada pimpinan daerah di 22 provinsi tersebut, terutama Jawa Timur dan Jawa tengah, untuk kembali memasifkan testing. Jangan sampai data yang dilaporkan lebih kecil dari kondisi kasus sebenarnya dan berimbas pada penentuan kebijakan yang kurang sesuai dengan situasi real di lapangan," katanya. Baca juga: Langgar Prokes, Tiga Bocah di Bandung Dihukum 'Baca Pancasila'(lir/zak)testing corona testing covid-19 virus corona covid-19 satgas covid-19</p>	0.07742	Fakta
<p>Jakarta - Camat Kembangan Joko menyebut RT yang menjadi zona merah Corona atau COVID-19 di wilayahnya telah berkurang. Dia menyebut ada empat RT yang sebelumnya zona merah, kini sudah menjadi zona oranye penyebaran Corona. "Kecamatan Kembangan sesuai publish data minggu ini dari Sudin Kesehatan Jakarta Barat adalah RT dengan zona merah terbanyak karena data awalnya jumlah RT zona merahnya berjumlah 11 kemudian setelah diverifikasi ternyata hanya tujuh RT," ujar Joko kepada</p>	0.12084	Fakta

<p>wartawan, Jumat (11/2/2022).Baca juga: 25 RT di Jakbar Zona Merah Corona, Terbanyak di Kecamatan KembanganDia mengatakan terdapat warga isolasi mandiri (isoman) yang sudah dinyatakan negatif Corona pada Rabu (10/2). Dia mengatakan berkurangnya jumlah warga yang menjalani isolasi mandiri membuat zona merah berkurang. "Jadi setelah dicek lapangan oleh Tim Satgas Kelurahanannya masing-masing, ternyata ada beberapa warga isoman yang sudah sembuh dan sudah negatif, sehingga tidak masuk kriteria zona merah lagi tetapi menjadi zona oranye," jelasnya.Dia menyebut empat RT tersebut sudah bebas dari micro-lockdown. Aktivitas warga juga telah kembali normal.Empat RT yang telah terbebas dari zona merah terdiri dari RT 14 RW 09 dan RT 07 RW 10 di Kelurahan Kembangan Utara. Kemudian, RT 01 RW 02 di Kembangan Selatan dan RT 07 RW 06 di Kelurahan Srengseng.Sedangkan tujuh RT yang masih menjadi zona merah Corona terdiri dari tiga RT di Meruya Utara, dua RT di Kembangan Selatan, satu RT di Kembangan Utara, dan satu RT di Srengseng.Baca juga: Ada 14 Kasus COVID di Apartemen Belmont Jakbar, Micro-lockdown Diperluas"Jumlah keseluruhan RT di Kembangan terdiri dari 619 RT dengan status 254 RT masih menjadi zona hijau, 251 RT masuk zona kuning, 107 RT masuk zona oranye dan tujuh RT di antaranya masih zona merah," ucapnya.Berikut sebaran tujuh RT zona merah di Kembangan, Jakarta Barat:1. Kelurahan Kembangan Utara: RT 13 RW 092. Kelurahan Kembangan Selatan: RT 11 RW 01 dan RT 10 RW 033. Kelurahan Meruya Utara: RT 02 RW 04, RT 23 RW 08 dan RT 08 RW 054. Kelurahan Srengseng: RT 08 RW 07Baca juga: Pusat Isoter Masjid Hasyim Asy'ari Jakbar Mulai Diisi Pasien COVID(ain/haf)jakarta barat jabodetabek covid-19 virus corona jakarta</p>		
---	--	--

Data Testing 6	Aktual	Prediksi
<p>Sukabumi - Sebanyak sembilan petugas Lembaga Pemasyarakatan (Lapas) Kelas II B Kota Sukabumi terkonfirmasi positif COVID-19. Hasil itu diketahui pasca melakukan tes massal petugas lapas pada Rabu dan Kamis, 9-10 Februari 2022 kemarin."Setelah dua hari kemarin berturut, petugas Lapas Sukabumi melakukan PCR, dari 28 orang yang PCR dikonfirmasi 9 orang yang positif COVID-19," kata Kepala Lembaga Pemasyarakatan Kelas IIB Sukabumi, Christo Victor Nixon Toar kepada detikcom, Jumat (11/2/2022).Lebih lanjut, sembilan petugas itu melakukan isolasi mandiri dan work from home (WFH). Christo mengatakan, upaya tracing dilakukan kepada 11 orang yang diduga melakukan kontak erat. "(Petugas menunjukkan) gejala ringan saja seperti pilek dan mau flu, tapi itu hanya sedikit orang. Kebanyakan sehat dan bugar. Tadi pagi sejumlah 11 orang saya perintahkan untuk dilakukan PCR di Labkesda Sukabumi, hasil dari tracing positif petugas," ujarnya.Baca juga: Kapolda Jabar: Knalpot Bising Memecah Konsentrasi Pengguna JalanBelum dapat diketahui apakah sembilan petugas tersebut terinfeksi varian Omicron atau tidak. Adapun untuk warga binaan, kata dia, akan dilakuka tes Antigen jika memiliki keluhan sakit."Sejauh ini belum ada rencana Tes PCR lagi, kami hanya melakukan tes antigen bagi</p>	Fakta	Fakta

<p>Warga Binaan yang mengeluhkan sakit saja," kata dia. Sejak meningkatnya kasus COVID-19 di Indonesia pada akhir Januari, pihaknya telah mengambil kebijakan untuk membatasi kegiatan di dalam Lapas. "Seluruh layanan Pembinaan kami tutup kecuali layanan Penitipan Makanan dari keluarga WBP," pungkasnya. Baca juga: Patung Jokowi Motoran Bakal Jadi Ikon Wisata Baru Mandalika Simak Video "Tips Asupan yang Disarankan WHO untuk Percepatan Pemulihan Covid-19": [Gambas: Video 20detik] (yum/bbn) lapas kelas ii b sukabumi petugas lapas positif covid-19 virus corona covid-19 kota sukabumi birojabar jawa barat</p>			
Tetangga Terdekat (k=3)	Jarak <i>Euclidean</i>	Label Tetangga	
<p>Gorontalo - Gubernur Gorontalo Rusli Habibie mengumumkan dirinya positif COVID-19 varian Omicron. Rusli kini menjalani isolasi mandiri (Isoman) di rumahnya. "Beliau (Gubernur Gorontalo) saat ini sementara Isoman," kata Kepala Dinas Komunikasi, Informatika, dan Statistik Masran Rauf, Jumat (11/2/2022). Humas Pemprov Gorontalo turut menyebarkan video Gubernur Gorontalo yang mengumumkan dirinya positif Omicron. Dalam video itu, tampak Rusli menggunakan pakaian olahraga. Rusli mengatakan, sebelum dinyatakan positif Omicron, dia pada Minggu (6/2) sempat berkunjung ke Jakarta. Saat tiba di Bandara Gorontalo, dia melakukan tes antigen dan dinyatakan negatif. Setelah itu, Rusli bertolak ke Kabupaten Boalemo untuk menghadiri rapat Forkopimda dan sejumlah agenda pemerintahan lainnya. Beberapa kali dilakukan rapid antigen hasilnya tetap negatif. Rusli kembali ke Gorontalo pada Selasa sore. Ia mulai merasakan demam dan meminta dites PCR. Hasil dari Laboratorium BPOM Gorontalo menyatakan dia positif Omicron. "Saya dinyatakan oleh Kepala BPOM Gorontalo (lab test PCR) positif COVID-19 varian Omicron," kata Rusli dalam video yang diunggah Humas Gorontalo. "Saya hari pertama batuk kering berdahak, kedua panas tinggi hampir 40 (derajat Celsius). Alhamdulillah demam sudah berangsur pulih, hari ini sudah sehat dan saya berolahraga," lanjutnya. Rusli mengingatkan warganya untuk lebih waspada karena seorang gubernur dengan protokol kesehatan yang ketat dan telah disuntik vaksin COVID sebanyak 3 kali dan dalam pengawasan dokter pun bisa terpapar COVID-19. Baca juga: Sambil Renang, Gubernur Gorontalo Umumkan Diri dan Istri Positif COVID Sementara itu, tim dokter juga sudah melacak orang-orang yang kontak erat dengan Rusli, termasuk sang istri Idah Syahidah, ajudan, humas, dan orang-orang yang bertugas di dekatnya. Hasilnya dinyatakan negatif. (nvl/nvl) gubernur gorontalo gubernur gorontalo positif omicron rusli habibie gubernur gorontalo rusli habibie gorontalo omicron covid-19 biromakassar</p>	0.10085	Fakta	
<p>Jakarta - Sebanyak 14 pegawai Pengadilan Negeri (PN) Tangerang positif COVID-19. Akibatnya, Pengadilan Negeri Tangerang ditutup sementara selama satu hari kerja. Penutupan dilakukan sebagai langkah pencegahan penyebaran COVID-19. Pelayanan di PN Tangerang ditutup pada Jumat (11/2) dan buka kembali pada Senin</p>	0.10541	Fakta	

(14/2)."Betul (14 pegawai di PN Tangerang positif COVID-19). Jumat kemarin tutup, tapi Senin aktivitas berjalan seperti biasa," kata pejabat Humas PN Tangerang, Arif Budi Cahyono, saat dihubungi detikcom, Sabtu (12/2/2022). Baca juga: 1.232 Kasus Baru COVID di Kabupaten Bogor, Terbanyak Cibinong Arif mengatakan 14 pegawai di PN Tangerang positif COVID-19. Ke-14 pegawai itu terdiri atas hakim, panitera, dan staf."Ada hakim, panitera pengganti, pegawai," ujar Arif. Lebih lanjut Arif menjelaskan, ke-14 pegawai diketahui positif COVID-19 setelah melakukan swab antigen di Pengadilan Negeri Tangerang, Kamis (10/2/2022)."Hari Kamis diadakan swab antigen," jelas Arif. Baca juga: Siswa Terpapar COVID-19, SD di Parepare Lockdown 10 Hari Simak Video 'PR Indonesia Jika Ingin 'Berdamai' dengan Covid-19 Seperti Swedia': [Gambas: Video 20detik] (aud/aud) jabodetabek tangerang pengadilan negeri tangerang pn tangerang		
London - Pangeran Charles dinyatakan positif COVID-19 untuk kedua kalinya. Putra tertua Ratu Elizabeth II itu kini menjalani isolasi. "Pangeran Wales telah dites positif Covid-19 dan sekarang menjalani isolasi," kata sebuah pesan di halaman Twitter resmi sang pangeran, seperti dilansir AFP, Kamis (10/2/2022). Charles seharusnya menghadiri peresmian patung di Winchester, Inggris selatan, pada Kamis waktu setempat. Akan tetapi, kehadirannya yang dibatalkan pada menit terakhir itu membuatnya kecewa. "(Pangeran Charles) sangat kecewa", kata Clarence House. Baca juga: Pangeran William & Kate Middleton Akan Mewarisi Kastil Terbesar di Dunia Pangeran Charles baru-baru ini bertemu dengan Ratu Elizabeth. Akan tetapi ratu yang berusia 95 tahun itu saat ini tidak menunjukkan gejala apa pun. Hal itu diungkapkan oleh sumber kerajaan yang tidak disebutkan namanya kepada kantor berita Asosiasi Pers domestik. Charles juga bertemu dengan beberapa orang di sebuah resepsi di British Museum pada Rabu waktu setempat. Dia juga sempat bertemu dengan Menteri Keuangan Rishi Sunak dan Menteri Dalam Negeri Priti Patel. Ratu Elizabeth diketahui memiliki masalah kesehatan baru-baru ini. Kondisi itu memaksanya mundur dari tugas resmi atas saran medis. Kembali lagi ke Pangeran Charles, pria yang berusia 73 tahun itu juga dinyatakan positif pada Maret 2020 lalu. Dia menghabiskan sekitar satu minggu melakukan isolasi di perkebunan Queen's Balmoral di Skotlandia. Baca juga: Ratu Elizabeth Ungkap Harapan agar Camilla Dikenal sebagai Permaisuri (lir/zak) pangeran charles pangeran charles positif corona covid-19 ratu elizabeth ii inggris	0.11872	Fakta

Data Testing 7	Aktual	Prediksi
Video yang diklaim menunjukkan antrian warga Malaysia yang ingin menginap di Stadion Shah Alam karena positif Covid-19 beredar di WhatsApp sejak 17 Mei 2021. "Warga Malaysia antri nginep di stadion shah alam... semua positif covid 19... ngerii x..." demikian narasi yang menyertai video itu. Dalam video tersebut, tampak antrian panjang ratusan orang yang memasuki sebuah gedung stadion. Banyak di antara mereka	Hoax	Fakta

yang membawa tas ransel, bahkan koper. Terlihat pula sejumlah petugas kesehatan yang memakai pakaian hazmat lengkap dengan masker dan faceshield. Beberapa gambar tangkapan layar video yang disebar dengan klaim keliru. Video ini bukan video antrian warga Malaysia yang ingin menginap di Stadion Shah Alam karena positif Covid-19.			
Tetangga Terdekat (k=3)	Jarak <i>Euclidean</i>	Label Tetangga	
Beredar sebuah informasi di internet yang mengklaim bahwa obat Viagra dapat menyembuhkan Covid-19. Informasi tersebut menyebar ketika ada seorang perawat bernama Monica Almeida yang menderita asma dan positif Covid-19 mengalami koma selama 28 hari. Dilansir dari dailymail.co.uk, obat disfungsi ereksi tersebut memungkinkan aliran darah mengalir lebih besar ke seluruh area tubuh hingga merelaksasi dinding pembuluh darah dan menyebabkan kondisinya membaik.	0.13195	<i>Hoax</i>	
Gorontalo - Gubernur Gorontalo Rusli Habibie mengumumkan dirinya positif COVID-19 varian Omicron. Rusli kini menjalani isolasi mandiri (Isoman) di rumahnya. "Beliau (Gubernur Gorontalo) saat ini sementara Isoman," kata Kepala Dinas Komunikasi, Informatika, dan Statistik Masran Rauf, Jumat (11/2/2022). Humas Pemprov Gorontalo turut menyebarkan video Gubernur Gorontalo yang mengumumkan dirinya positif Omicron. Dalam video itu, tampak Rusli menggunakan pakaian olahraga. Rusli mengatakan, sebelum dinyatakan positif Omicron, dia pada Minggu (6/2) sempat berkunjung ke Jakarta. Saat tiba di Bandara Gorontalo, dia melakukan tes antigen dan dinyatakan negatif. Setelah itu, Rusli bertolak ke Kabupaten Boalemo untuk menghadiri rapat Forkopimda dan sejumlah agenda pemerintahan lainnya. Beberapa kali dilakukan rapid antigen hasilnya tetap negatif. Rusli kembali ke Gorontalo pada Selasa sore. Ia mulai merasakan demam dan meminta dites PCR. Hasil dari Laboratorium BPOM Gorontalo menyatakan dia positif Omicron. "Saya dinyatakan oleh Kepala BPOM Gorontalo (lab test PCR) positif COVID-19 varian Omicron," kata Rusli dalam video yang diunggah Humas Gorontalo. "Saya hari pertama batuk kering berdahak, kedua panas tinggi hampir 40 (derajat Celsius). Alhamdulillah demam sudah berangsur pulih, hari ini sudah sehat dan saya berolahraga," lanjutnya. Rusli mengingatkan warganya untuk lebih waspada karena seorang gubernur dengan protokol kesehatan yang ketat dan telah disuntik vaksin COVID sebanyak 3 kali dan dalam pengawasan dokter pun bisa terpapar COVID-19. Baca juga: Sambil Renang, Gubernur Gorontalo Umumkan Diri dan Istri Positif COVID. Sementara itu, tim dokter juga sudah melacak orang-orang yang kontak erat dengan Rusli, termasuk sang istri Idah Syahidah, ajudan, humas, dan orang-orang yang bertugas di dekatnya. Hasilnya dinyatakan negatif. (nvl/nvl) gubernur gorontalo gubernur gorontalo positif omicron rusli habibie gubernur gorontalo rusli habibie gorontalo omicron covid-19 biromakassar	0.13977	Fakta	
Depok - Jamaludin (36) kaget saat menerima pesan WhatsApp (WA) dari Kementerian Kesehatan (Kemenkes).	0.15611	Fakta	

<p>Dia bertambah kaget setelah membaca isi pesannya. Jamaludin 'divonis' positif COVID-19. Padahal dia tidak pernah menjalani tes usap PCR. Tak percaya dengan WA Kemenkes itu, dia lalu membuka aplikasi PeduliLindungi. Dia makin kaget karena statusnya menunjukkan warna hitam yang artinya positif COVID-19. "Saya mengecek aplikasi PeduliLindungi dan mendapatkan hasil PCR positif dari laboratorium RS Brawijaya Depok. Sebelumnya tidak pernah melakukan PCR swab sekalipun di RS Brawijaya Depok," kata Jamaludin saat dihubungi, Jumat (11/2/2022). Status positif Corona itu diterimanya sejak Rabu (9/2) lalu. Berbagai upaya dilakukannya untuk bisa membetulkan status tersebut. Baca juga: RS di Depok Kirim Permohonan Maaf ke Warga Tangel soal 'Vonis' COVID "Saya nggak bisa ke mana-mana selama dua hari. Dari jam 03.00 pagi sampai malam masih telepon 119 menanyakan gimana. Terus telepon ke pihak RS," kata Jamal. Jamaludin mengaku menghubungi 119 extension 9 untuk menanyakan mekanisme koreksi status positif COVID-19. Ia lalu diarahkan untuk menghubungi pihak RS. "Saya menghubungi RS Brawijaya Depok melalui call centre namun belum berhasil dan saya kembali menghubungi IGD RS Brawijaya Depok," sambungnya. Tahapan panjang dilaluinya untuk membetulkan statusnya di PeduliLindungi. Setelah menghubungi pihak RS, Jamaludin diberi info bahwa perubahan status sedang diurus. Status PeduliLindungi Kembali Hijau Status Jamaludin di PeduliLindungi akhirnya kembali berwarna hijau yang artinya negatif COVID-19. Dia mengatakan status itu telah kembali per Jumat (10/2) malam. Namun, ia masih meminta permintaan maaf secara tertulis dari pihak RS. "Saya minta sama mereka ada surat permintaan maaf dari rumah sakit secara tertulis kepada saya. Ada bentuk tanggung jawab dan mereka mengakui kalau itu kelalaian dari rumah sakit," katanya. Baca juga: Kagetnya Warga Tangel 'Divonis' Positif COVID-19, Padahal Tak Tes PCR Jamaludin mengakui sempat berobat ke RS Brawijaya Depok. Namun, itu dilakukan pada tahun lalu dan bukan untuk pengecekan COVID-19. Bagaimana duduk perkara kasus ini? Simak di halaman berikutnya. Saksikan juga 'Heboh Wanita Protes Dapat Hasil Positif Covid Sebelum Tes PCR': [Gambas: Video 20detik] "Sudah lama banget tahun lalu kali ya. Saya pernah berobat sekali di sana, waktu itu masuk IGD karena sudah lewat jam praktik," kata Jamaludin. Riwayat Jamaludin berobat ke RS Brawijaya ini diduga menjadi faktor yang tak terlepas dari kesalahan input data hasil tes COVID-19. Duduk Perkara Status 'Positif Corona' Pihak RS Brawijaya menjelaskan masalah yang dihadapi Jamaludin. Pihak RS mengakui terjadi kesalahan input data COVID-19 kepada Jamaludin. "Jadi memang kami akui ada salah penginput-an data karena namanya sama, tanggal lahirnya sama. Kalau kami di rumah sakit kan pencocokan datanya berdasarkan nama, jenis kelamin, dan tanggal lahir," kata Supervisor on Duty (SOD) Marketing RS Brawijaya Depok,</p>		
--	--	--

<p>Wahyuana Kumala, saat dimintai konfirmasi. Dia mengatakan petugas di laboratorium keliru karena ada nama yang sama namun berbeda alamat. "Di laboratorium (ada) anak baru di situ dia input. Ditanya namanya siapa langsung klik, ternyata tidak memerhatikan alamat," ujarnya. Baca juga: Akui Salah, RS Brawijaya Depok Minta Maaf 'Vonis' Warga Tangsel Positif COVID Pihak RS Minta Maaf Dia mengatakan warga bernama Jamaludin yang positif Corona itu tinggal di Bojongsari, Depok bukan di Ciputat, Tangsel. Wahyuana meminta maaf kepada pihak terkait yang telah dirugikan. "Kami akui ada kesalahan dan kami sudah meminta maaf," tutur Wahyuana. Pihak RS juga akan mengirimkan surat permohonan maaf kepada Jamaludin terkait salah input hasil tes COVID-19. Surat itu akan dikirim ke rumah Jamaludin di Tangerang Selatan (Tangsel). "Sudah kami proses hari ini, kami rencana kirim ke rumah beliau untuk permohonan maaf. Kami sedang proses," katanya. (jbr/idn)round-up jabodetabek pedulilindungi kemenkes depok covid-19</p>		
--	--	--

Data Testing 8		Aktual	Prediksi
<p>Unggahan yang berisi klaim bahwa anosmia bukan gejala khas virus dan bisa diobati dengan mecobalamin beredar di Instagram pada 17 Februari 2021. Klaim itu dilengkapi dengan foto obat mecobalamin. "Anosmia Bukan Gejala khas Virus. Memangnya dari dulu gak pernah ada yg merasakan gejala ini?? Jangan mau di takut2i otak dengkul! Minum aja Mecobalamin. 5 cap setiap 1 jam sampai diare ringan," demikian narasi di bagian awal unggahan tersebut. Kemudian, unggahan itu menyinggung soal Covid-19. "Covid19 cuma kurang vitamin dosis tinggi dan mineral elektrolit. Tidak ada virus ganas! Obat penyebab bergejala berat dan kematian!" Gambar tangkapan layar unggahan di Instagram yang berisi klaim keliru terkait anosmia dan obat mecobalamin.</p>		<i>Hoax</i>	<i>Hoax</i>
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga	
<p>Pesan berantai yang berjudul "Tiga Kejanggalan dan Ketidak Wajaran terhadap Isu Corona yang Terjadi di Negeri +62" beredar di Facebook pada 24 Juni 2021. Dalam pesan berantai itu, terdapat beberapa klaim yang dilontarkan. Pertama, Cina disebut mengakui bahwa virus Corona penyebab Covid-19 bukan virus yang membahayakan, melainkan hanya virus flu biasa. Kedua, menurut pesan berantai tersebut, semua virus dan bakteri tidak bisa berkembang biak di tempat yang bersih dan suci. Sementara klaim ketiga adalah, jika Covid-19 dikategorikan sebagai pandemi, kasus kematian seharusnya ada di mana-mana. "Korban yang mati pada bergelimpangan di rumah-rumah, apartemen di pasar-pasar dan di tempat mereka berada." Gambar tangkapan layar pesan berantai di Facebook yang berisi sejumlah klaim keliru terkait Covid-19.</p>	0.06879	<i>Hoax</i>	
<p>Video berdurasi 9 menit yang berisi pernyataan dari aktor Russell Brand beredar di Instagram. Menurut klaim yang menyertai video ini, Brand mengomentari tentang bocornya ribuan email milik Direktorat Institut Nasional Alergi dan</p>	0.11788	<i>Hoax</i>	

Penyakit Menular Amerika Serikat Anthony Fauci. Menurut klaim itu, Brand menyatakan bahwa, dalam ribuan email Fauci, ditemukan indikasi bahwa mantan kepala penasihat medis Gedung Putih itu terlibat dalam pembiayaan terciptanya Covid-19 di Laboratorium Wuhan, Cina. Brand juga menyinggung komunikasi intens antara Fauci dan Mark Zuckerberg, pendiri Facebook. Dalam komunikasi itu, dibahas tentang konten apa saja yang harus disensor di Facebook, Instagram, dan YouTube. Akun ini mengunggah video beserta klaim tersebut pada 7 Juni 2021. Di akhir keterangannya, akun itu menulis, "Siapa yang sebenarnya yang waras atau tak waras dalam menyikapi pandemi? Mereka yang selalu disensor namun akhirnya terbukti benar?" Hingga kini, video tersebut telah ditonton 2.283 kali. Gambar tangkapan layar video milik aktor Russel Brand yang berisi klaim keliru terkait email Direktur Institut Nasional Alergi dan Penyakit Menular Amerika Serikat Anthony Fauci.		
Pesan berantai yang diklaim berasal dari Ikatan Dokter Indonesia (IDI) beredar di Facebook. Pesan berantai itu berisi sejumlah klaim yang meragukan pandemi Covid-19. "Tidak ada pandemi, tidak ada Covid-19, dan tidak ada virus yang beterbangan yang mematikan. Semua itu adalah bentuk pengelabuan dan pembodohan global," demikian narasi di awal pesan berantai tersebut. Akun ini membagikan pesan berantai itu pada 21 Maret 2021. Beberapa klaim yang terdapat dalam pesan berantai tersebut antara lain: Gambar tangkapan layar pesan berantai yang beredar di Facebook yang berisi klaim-klaim keliru dan menyesatkan seputar pandemi dan Covid-19.	0.12908	<i>Hoax</i>

Data Testing 9	Aktual	Prediksi
Serang - Jumlah pasien terpapar virus Corona atau COVID-19 di Banten dalam sepekan ini terus meningkat. Meski begitu, tempat tidur pasien penanganan virus dan ruang ICU diklaim masih tersedia. "Dari hasil pemetaan yang ada, evaluasi kondisi tempat tidur dan ICU, insyaallah di posisi aman," kata Wakil Gubernur Banten Andik Hazrumy di Serang, Banten, Kamis (10/2/2022). Jika di tingkat nasional kenaikan per harinya 900 persen, begitu pun terjadi di Banten. Bahkan, menurut Andika, sampai berkali-kali lipat dibandingkan varian sebelum Omicron. Baca juga: Kajati Banten Telah Kantongi Tersangka Dugaan Korupsi Komputer UNBK Rp 25 MPada Rabu (9/2) kemarin, Dinas Kesehatan Banten mencatat penambahan kasus pasien terjangkit mencapai 6.026 per hari. Semua wilayah di Banten, kecuali Kabupaten Serang, berada di zona oranye. Ia menambahkan, varian Omicron ini penyebarannya sangat mengkhawatirkan dan cepat. Jika satu lingkungan kantor ada satu pasien yang terpapar, bisa dipastikan seluruhnya bisa terjangkit. "Yang kena satu, se-kantor kena, satu kena bisa tujuh (orang)," ujar Andik. Seluruh tempat tidur termasuk ICU di Banten sendiri ada 2.950 tempat. Jumlah ini sudah ditambah bila dibandingkan dengan tahun 2020 sebanyak 2.000 tempat tidur. Pusat pemerintahan Banten di KP3B saat ini sudah menerapkan Work From Home (WFH). Kapasitas setiap kantor hanya 25 persen kecuali dinas yang esensial yaitu Dinas Kesehatan, Rumah Sakit, Satpol PP, Dinas Perhubungan dan Badan Penanggulangan Bencana Daerah atau	Fakta	Fakta

BPBD. Baca juga: Menyelam Cari Aki Tercebur ke Laut, Pekerja Galangan Kapal di Banten Hilang (bri/bbn)covid-19 virus corona banten birojabar			
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga	
<p>Jakarta - Virus Corona (COVID-19) varian Omicron membuat kasus global meningkat drastis. Dalam waktu sebulan jumlah kasus meningkat 100 juta kasus. Dilansir dari kantor berita Reuters, Kamis (10/2/2022), varian Omicron, yang mendominasi lonjakan kasus infeksi virus Corona di seluruh dunia, menyumbang hampir semua kasus baru yang dilaporkan setiap hari. Per Rabu (9/2), jumlah kasus Omicron di dunia melampaui 400 juta. Baca juga: Raja Salman-Biden Teleponan Bahas Iran-Yaman, Corona Dunia Tembus 400 Juta Menurut penghitungan Reuters, butuh lebih dari sebulan untuk jumlah kasus COVID-19 mencapai 400 juta dari 300 juta. Kondisi itu lebih cepat dibandingkan pertambahan dari 200 juta kasus menuju 300 juta kasus. Butuh waktu lima bulan untuk penambahan dari 200 juta kasus ke 300 juta kasus. Baca juga: Kasus Covid-19 Dunia Turun 17 Persen, Kematian Naik 7 Persen Pandemi ini telah menewaskan lebih dari 6 juta orang di seluruh dunia. Kematian telah meningkat 70 persen dalam lima minggu terakhir berdasarkan rata-rata tujuh hari. Menurut analisa Reuters, kasus mulai menurun di banyak negara, rata-rata lebih dari 2 juta kasus masih dilaporkan setiap hari. Baca juga: Mungkinkah Omicron Picu Long COVID? Ini Penjelasan Pakar WHO Lima negara teratas yang melaporkan jumlah kasus terbanyak dalam rata-rata tujuh hari - Amerika Serikat, Prancis, Jerman, Rusia, dan Brasil - menyumbang sekitar 37 persen dari semua kasus baru yang dilaporkan di seluruh dunia, menurut analisis Reuters. Setengah Juta Orang Meninggal Sejak Omicron Muncul Organisasi Kesehatan Dunia (WHO) mengungkapkan sejak omicron ditemukan dan dinyatakan variant of concern (VoC), setengah juta orang meninggal karena. Omicron ditetapkan WHO sebagai VoC pada November 2021. Manajer Insiden WHO, Abdi Mahamud, mengatakan ada 130 juta kasus dan 500 ribu kematian di dunia akibat varian Omicron. Menurutnya, Omicron dengan cepat mengambil alih Delta sebagai varian yang dominan di dunia karena lebih mudah menular. Baca juga: PM Inggris Berniat Hapus Aturan Isolasi Mandiri Penderita Corona! "Di zaman vaksin yang efektif, setengah juta orang meninggal, itu benar-benar sesuatu," kata Mahamud dalam interaksi langsung di saluran media sosial WHO yang dikutip dari NDTV. "Sementara semua orang mengatakan Omicron lebih ringan, (mereka) melewatkan fakta bahwa setengah juta orang telah meninggal sejak ini (varian Omicron) terdeteksi. Ini sangat tragis," tegasnya. Pada kesempatan yang sama, pimpinan teknis COVID-19 WHO Maria Van Kerkhove mengatakan kemunculan varian Omicron ini sangat mengejutkan. Pasalnya, jumlah kasus yang terjadi jauh lebih tinggi daripada yang diketahui. "Itu menyebabkan puncak-puncak kasus, yang sebelumnya terlihat hampir datar," beber</p>	0.07148	Fakta	

<p>dia. "Kita masih berada di tengah pandemi ini. Semoga kita semakin mendekati akhir. Banyak negara belum melewati puncak Omicron mereka," lanjut Maria. Baca juga: WHO: Nyaris Setengah Juta Orang Meninggal Sejak Omicron Muncul Amerika Terbanyak Amerika Serikat memimpin dunia dalam jumlah kasus COVID-19 terbanyak yang dilaporkan setiap hari. Satu juta kasus baru dilaporkan di negara itu setiap tiga hari. Namun, kasus dan rawat inap di negara itu melambat dari puncaknya pada Januari tahun ini, menurut analisis Reuters. Negara itu telah mencatat lebih dari 900.000 kematian terkait dengan COVID-19. Di Prancis, angka rata-rata tujuh hari kasus baru telah mencapai lebih dari 210.000 kasus per hari. Bertambah sekitar satu juta kasus baru setiap lima hari. Total kumulatif untuk kasus COVID yang dikonfirmasi di Prancis sejak awal pandemi melewati 20 juta pada Kamis pekan lalu. Kondisi di Indonesia Pada Kamis (10/2), kasus COVID-19 di Indonesia bertambah 40.618 kasus. Sehingga, total kasus di Indonesia berjumlah 4.667.554 kasus. Sebanyak 288.186 merupakan kasus aktif. Pemerintah juga melaporkan sebanyak 18.182 pasien sembuh dari Corona dalam 24 jam terakhir. Adapun total kasus sembuh Corona di Indonesia hingga hari ini sebanyak 4.234.510. Selain itu, kasus kematian COVID-19 di Indonesia bertambah 74 kasus. Sehingga, total kasus kematian pasien Corona di RI sebanyak 144.858 kasus. Baca juga: WHO: Subvarian Omicron BA.2 Lebih Menular dari Omicron Aslinya (aik/aik) round-up virus corona omicron varian omicron who</p>		
<p>Banjarmasin - Kapolri Jenderal Listyo Sigit Prabowo hari ini meninjau langsung akselerasi vaksinasi serentak di Gedung Sultan Suriansyah, Banjarmasin, Kalimantan Selatan (Kalsel). Dalam kesempatan ini, Sigit juga memantau secara virtual vaksinasi COVID-19 di 4.988 tempat vaksinasi di 34 provinsi di Indonesia. Sigit mengingatkan target vaksinasi pada hari ini untuk di Kalsel, sebanyak 21 ribu dosis. Ia pun mengapresiasi jajaran Forkopimda Kalsel yang telah bersinergi dan berusaha keras melakukan akselerasi percepatan vaksinasi untuk masyarakat. "Tentunya dalam kesempatan ini saya ucapkan terima kasih, apresiasi atas kerja keras kerja sama dari TNI-Polri dan seluruh pemda, baik provinsi, kota madya, maupun kabupaten, yang terus bekerja sama membantu percepatan program akselerasi vaksinasi nasional," kata Sigit, Kamis (10/2/2022). Baca juga: Kapolri Minta Forkopimda Babel Perkuat Pencegahan Lonjakan COVID-19 Sigit menuturkan angka harian COVID-19 mengalami kenaikan. Tercatat kemarin angka positif COVID-19 sebanyak 46 ribuan atau meningkat 10 ribu dibanding hari sebelumnya. Bahkan angka tersebut hampir mendekati puncak angka positif harian COVID-19 di Indonesia pada Juli 2021. Kapolri Jenderal Listyo Sigit meninjau langsung vaksinasi serentak di wilayah Kalsel. (Dok. Istimewa) Namun Sigit menegaskan terjadinya lonjakan kasus COVID-19 harus dihadapi dengan upaya yang maksimal dari seluruh pihak dalam melakukan penanganan dan pengendalian. Segala strategi harus</p>	0.09007	Fakta

<p>dilakukan untuk menekan angka harian. Terpenting menurut Sigit, masyarakat harus tetap tenang dan tidak panik dalam menghadapi penyebaran COVID-19."Tentunya ini menjadi kewaspadaan bagi kita semua untuk kemudian bersiap-siap, walaupun dalam kesempatan ini kita ingatkan masyarakat tak perlu panik. Namun upaya menghadapi lonjakan ini kita harus lakukan dengan maksimal," ujar Sigit. Mantan Kabareskrim Polri ini lantas bicara soal upaya menghadapi lonjakan kasus COVID-19. Salah satunya adalah strategi untuk mempercepat akselerasi vaksinasi yang wilayahnya belum 100 persen. Dari data yang dia terima, untuk vaksinasi dosis pertama secara nasional sudah mencapai 89 persen. Sehingga, diharapkan secepatnya bisa mencapai 100 persen. Sigit juga meminta masyarakat yang sudah melaksanakan vaksinasi sebanyak dua kali, dan sudah memasuki masa enam bulan, melaksanakan vaksinasi ketiga atau booster. "Karena ini sangat penting sebab kecenderungan setelah enam bulan tingkat imunitas menurun. Jadi mau tak mau harus dilaksanakan vaksin untuk kemudian meningkatkan imunitas," ujar Sigit. Baca juga: Kapolri Paparkan Strategi untuk Antisipasi Lonjakan COVID-19 Lebih lanjut, mantan Kapolda Banten ini mengatakan vaksinasi sangat penting untuk menghindari fatalitas jika terpapar COVID-19. Berdasarkan data, rata-rata memang yang sudah divaksinasi dua kali atau booster bisa terkena Omicron, namun kecenderungannya tanpa gejala atau gejala ringan. Kapolri Jenderal Listyo Sigit meninjau langsung vaksinasi serentak di wilayah Kalsel. (Dok. Istimewa) "Utamanya yang belum vaksin atau belum lengkap khususnya lansia atau komorbid memang ada beberapa yang mengalami fatalitas atau meninggal. Tolong ingatkan keluarga, tetangga atau masyarakat yang memiliki usia lansia segera vaksinasi," kata Sigit. Selain vaksinasi, Sigit tidak henti-hentinya mengingatkan masyarakat untuk berdisiplin terhadap protokol kesehatan. Ia ingin pemakaian masker kembali menjadi kebiasaan rutin di masa pandemi COVID-19 yang kembali melonjak. "Salah satu penularan yang bisa terjadi manakala masyarakat di tempat kerumunan membuka masker," ujar Sigit. Dalam kesempatan ini, Sigit juga mengimbau Forkopimda agar menyiapkan tempat isolasi terpusat (isoter), manakala ada masyarakat yang terpapar COVID-19, namun tidak memungkinkan untuk melakukan isolasi mandiri (isoman) di rumah. Di tempat isoter petugas medis akan lebih mudah mengawasi dan memberi pelayanan kesehatan ke masyarakat yang terpapar COVID-19. "Di sana (isoter) disiapkan dokter, disiapkan obatnya setiap hari dicek. Ini akan bisa menjadi jauh lebih baik dibanding melaksanakan secara mandiri karena kita agak susah mengontrolnya dan risiko tak terjadi penularan tentunya akan sangat sulit diawasi," papar Sigit. Lebih lanjut Sigit menyatakan jajarannya di seluruh Indonesia kembali mengaktifkan pelaksanaan PPKM mikro. Ini dilakukan untuk mengawasi masyarakat di wilayahnya yang sedang melaksanakan isolasi mandiri. "PPKM mikro memiliki tugas tambahan mengecek wilayahnya, khususnya di wilayah yang</p>		
--	--	--

diperlukan ada yang isoman diawasi secara ketat. Cek apakah sudah dapat obat atau belum. Kemudian dikontrol agar kita menjaga laju varian Omicron ini kita kendalikan," ujar Sigit. Selain itu, Sigit memberikan pengarahan kepada seluruh jajaran yang menggelar kegiatan vaksinasi serentak di Indonesia melalui sambungan virtual. Ia menekankan kesiapan dan pencegahan terkait lonjakan kasus pertumbuhan COVID-19. (hri/fjp) kapolri kapolri jenderal listyo sigit prabowo akselerasi vaksinasi covid-19 covid-19		
Narasi tentang seorang dokter Malaysia Dr. Chai Koh Meow yang meninggal karena vaksin booster mRNA, beredar di Facebook, 23 November 2021. Narasi ini mengambil sumber dari situs berbahasa Mandarin, sinchew.com. Unggahan itu menulis bahwa Chai Koh Meow meninggal setelah menerima vaksin booster Pfizer, untuk melengkapi vaksin pertamanya yang menggunakan Sinovac. "Seorang dokter Malaysia berusia 58 tahun, Dr. Chai Koh Meow, deputi direktur Departemen Kesehatan Malaysia, menerima suntikan vaksin booster Covid buatan Pfizer hari Selasa yang lalu, sebagai tambahan ke atas vaksin Sinovac yang diterimanya terdahulu, meninggal dunia setelah mengalami gejala-gejala tubuh menjadi tidak nyaman seperti demam (colds) dan rasa sakit (soreness)," demikian isi narasi tersebut. Tangkapan layar unggahan dengan klaim dokter Malaysia Dr. Chai Koh Meow meninggal setelah mendapat vaksin booster Pfizer	0.09524	Hoax

Data Testing 10		Aktual	Prediksi
Pesan berantai yang berisi klaim bahwa obat yang bernama Ivermectin dapat menyembuhkan sekaligus mencegah Covid-19 beredar di grup-grup percakapan WhatsApp. Klaim ini dibagikan bersama tautan sebuah artikel berita yang menyebut bahwa obat tersebut akan dibagikan di Kudus, Jawa Tengah. Beberapa pekan terakhir, kasus Covid-19 di Kudus tengah melonjak, di mana varian Delta menjadi varian virus Corona yang mendominasi penularan di sana. "Utk warga Jkt, skrg Ivermectin 12 mg utk Covid sdh tersedia bebas di Apotik Jkt. Jadi utk warga Jkt, andai ada saudara/teman yg Positif Covid, disarankan utk segera minum Ivermectin 12 mg, produksi PT. Harsen Farma Indonesia," demikian narasi dalam pesan berantai itu. "Beberapa studi menunjukkan aktifitas pencegahan covid. Dosis pencegahan/profilaksis tanggal 1 dan tanggal 3 satu tablet 12 mg tiap bulan." Gambar tangkapan layar pesan berantai di WhatsApp yang berisi klaim keliru terkait Ivermectin.		Hoax	Hoax
Tetangga Terdekat (k=3)	Jarak Euclidean	Label Tetangga	
Video yang memperlihatkan sebuah proses vaksinasi beredar di Twitter. Dalam video yang diambil dari atas lokasi vaksinasi itu, terlihat bahwa petugas tidak menyuntikkan vaksin ke lengan penerima, melainkan ke bagian bajunya. Video itu pun diklaim sebagai bukti kebohongan vaksinasi Covid-19 oleh Yahudi. Dalam video itu, memang terdapat sebuah teks yang ditulis dalam huruf Ibrani. Akun ini membagikan video tersebut pada 24 Maret 2021 dengan narasi, "Tu liat kelakuan Yahudi ngebohongin dunia se olah2 di vaccine padahal tdk, dia ga tau ada cctv yg lg ngerekam.."	0.22597	Hoax	

Hingga artikel ini dimuat, cuitan itu telah di-retweet sebanyak 393 kali dan disukai hingga 682 kali. Gambar tangkapan layar cuitan di Twitter yang berisi klaim sesat terkait video proses vaksinasi Covid-19 di Israel yang diunggahnya.		
Pesan berantai yang berjudul "Genosida Umat Islam Indonesia dengan Vaksin Covid-19" beredar di Facebook. Pesan ini berisi sejumlah klaim yang meragukan pandemi Covid-19 di dunia, termasuk di Indonesia. Di bagian awal pesan itu, disebutkan bahwa Inggris dan para ahli medis Eropa telah menyatakan Covid-19 tidak lagi dikategorikan sebagai pandemi maupun penyakit berisiko tinggi. Kemudian, terdapat klaim-klaim lainnya, seperti bahwa orang-orang di Cina dan negara-negara Eropa tidak ada yang mau memakai masker lagi karena seluruh warganya sudah sehat, bahwa Malaysia, Brunei Darussalam, Filipina, Thailand, dan Vietnam menolak vaksin, bahwa rumah sakit adalah tempat genosida umat Islam, dan bahwa melepas masker saat berbicara wajib karena, kalau tidak, akan menelan racun karbon dioksida atau CO ₂ . Akun ini membagikan pesan berantai tersebut pada 26 Juni 2021. Hingga artikel ini dimuat, unggahan akun itu telah mendapatkan 99 reaksi dan 57 komentar serta dibagikan lebih dari 200 kali. Gambar tangkapan layar pesan berantai yang beredar di Facebook yang berisi klaim-klaim keliru terkait Covid-19 dan vaksin.	0.23142	<i>Hoax</i>
Pesan berantai yang berisi klaim bahwa, usai vaksinasi Covid-19, tubuh justru lebih rentan terinfeksi virus Corona beredar Facebook. Karena itu, menurut pesan tersebut, setelah disuntik vaksin Covid-19, penerima vaksin dianjurkan untuk tidak banyak beraktivitas secara berat dan tidak pergi keluar rumah. Akun ini membagikan pesan berantai itu pada 12 Maret 2021. Menurut pesan tersebut, usai vaksinasi Covid-19, imunitas tubuh belum terbentuk dengan sempurna. Antibodi baru terbentuk secara sempurna dua pekan setelah vaksinasi dosis kedua. "Ini ada beberapa lansia di Surabaya yang kena Covid-19 setelah divaksin. Enggak mau istirahat. Karena merasa sudah aman, lalu keluyuran keluar," demikian yang tertulis dalam pesan itu. Pesan berantai ini pun menyinggung bahwa vaksin Covid-19 dosis kedua harus diberikan 21-28 hari setelah vaksinasi dosis pertama. "Vaksin kedua makan waktu kira-kira 14-21 hari baru jadi. Jadi, hitung-hitung dari vaksin dosis pertama ke vaksin dosis kedua sampai kekebalan terbangun itu harus menunggu sekitar dua bulan. Baru 85-92 persen kebal Covid-19." Gambar tangkapan layar pesan berantai yang beredar di Facebook yang berisi klaim menyesatkan terkait vaksinasi Covid-19.	0.23144	<i>Hoax</i>