

BAB II

LANDASAN TEORI

© Hak cipta milik IBI KIKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian

A. Tinjauan Pustaka

1. Informasi

Menurut Fithrie Soufitri., (2023:4), “Informasi adalah sekumpulan data fakta yang diorganisasi atau diolah dengan cara tertentu sehingga mempunyai arti bagi penerima. Data yang telah diolah menjadi sesuatu yang berguna bagi si penerima maksudnya yaitu dapat memberikan keterangan atau pengetahuan.”

Menurut Erni Widarti et al. (2024:17) “Informasi adalah kumpulan informasi atau kenyataan yang diolah dengan cara khusus sehingga mempunyai atau mempunyai tujuan bagi penerima informasi tersebut.”

Menurut Kenneth C.Laudon dan Jane P. Laudon (2022:46) definisi informasi yaitu

:
“By information we mean data that have been shaped into a form that is meaningful and useful to human beings.”

Informasi dimaksud dengan data yang telah terbentuk menjadi sebuah yang sangat berarti dan berguna bagi manusia.

Menurut Erni Widarti et al. (2024:19) manfaat dari informasi adalah :

- Meminimalkan risiko Anda mengambil keputusan yang salah.
- Memberikan gambaran yang tepat dan faktual sesuai dengan kondisi yang ada saat ini.
- Menyajikan tren atau kecenderungan tentang suatu hal di masa mendatang.

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



- d. Mengurangi potensi keresahan akibat kondisi yang tidak pasti.
- e. Menjadi dasar dalam menentukan solusi dari sebuah permasalahan.
- f. Mengefektifkan alur kerja dalam sebuah organisasi.
- g. Mempertahankan citra positif perusahaan dan meningkatkan kepercayaan.
- h. Membantu menyusun materi promosi agar lebih jelas dan terarah.
- i. Menjadi dasar pertanggungjawaban atas sebuah keputusan atau tindakan yang sudah dilakukan.
- j. Memberikan bukti valid (bukan kesan, opini, atau isu) kepada pihak lain.

Menurut Erni Widarti et al. (2024:20) Jenis – jenis Informasi secara umum adalah:

- a. Faktual: jenis Informasi yang sesuai dengan fakta yang benar-benar terjadi
- b. Opini: jenis Informasi berupa pendapat seseorang tentang suatu hal
- c. Deskripsi: jenis Informasi dalam bentuk penjelasan rinci dan detail tentang sesuatu

C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

2. Sistem Informasi

Menurut Kenneth C.Laudon dan Jane P. Laudon (2022:46) definisi sistem informasi yaitu :

“An Information System can be defined technically as a set of interrelated components that collect (or retrieve), process, store and distribute information to support decision making and control in an organization.”

Sistem Informasi bisa didefinisikan secara teknis sebagai sekumpulan komponen yang saling terkait yang mengumpulkan (atau mendapatkan kembali), memproses,

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



menyimpan, dan mendistribusikan informasi untuk mendukung pengambilan keputusan dan pengendalian dalam suatu organisasi.

C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian

Menurut Fithrie Soufitri., (2023:6), “Sistem Informasi adalah integrasi dari komponen-komponen yang telah dianalisa dan diproses sehingga menghasilkan informasi yang diperlukan untuk dapat membantu manajer dalam pengambilan suatu keputusan.”

Data

Menurut Rina Fitriana et al. (2022:15), “Data merupakan suatu kumpulan atau catatan fakta-fakta untuk memberikan gambaran yang luas terkait dengan suatu keadaan.”

Menurut Annita Sari et al. (2023:97), “Data merupakan segala fakta dan angka yang dapat dijadikan bahan untuk menyusun suatu informasi.”

Menurut Annita Sari et al. (2023:98) Jenis data dilihat dari derajat datanya, terbagi dua, yaitu:

a. Data Primer

Data primer adalah data yang langsung diperoleh dari sumber data pertama di lokasi penelitian atau objek penelitian. Data primer ini, disebut juga data asli atau data baru. Data yang diperoleh secara langsung dari masyarakat baik yang dilakukan melalui wawancara, observasi dan alat lainnya, merupakan data primer. Data primer yang bersifat polos, apa adanya dan masih mentah memerlukan analisa lebih lanjut.

b. Data Sekunder

Data sekunder adalah data yang diperoleh dari sumber kedua atau sumber sekunder dari data yang dibutuhkan. Data yang diperoleh secara tidak langsung, misalnya

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



melalui dokumen laporan- laporan peneliti terdahulu. Data sekunder disebut juga data tersedia. Data ini biasanya digunakan untuk melengkapi data primer. Bahan kepustakaan yang dapat dipergunakan dalam penelitian tidak hanya berupa teori-teori yang telah matang, siap untuk dipakai, tetapi dapat pula berupa hasil-hasil penelitian yang masih memerlukan pengujian kebenarannya.

Menurut Annita Sari et al. (2023:99) Data dibedakan berdasarkan waktu pengumpulannya dibagi menjadi dua, yaitu :

a. Data Berkala (*Time Series*)

Data Berkala (*Time Series*) adalah data yang terkumpul dari waktu ke waktu untuk memberikan gambaran perkembangan suatu kegiatan atau keadaan.

b. Data Kerat Lintang (*Cross Section*)

Data Kerat Lintang (*Cross Section*) adalah data yang terkumpul pada suatu waktu tertentu memberikan gambaran perkembangan suatu kegiatan atau keadaan pada waktu itu.

Menurut Annita Sari et al. (2023:101) Jenis data dilihat dari kemungkinan pengukurannya, yaitu :

a. Data Kualitatif

Data kualitatif adalah data yang mengacu pada data kualitas objek penelitian, yaitu ukuran data berupa non angka yang merupakan satuan kualitas (misalnya, istimewa, baik, buruk, tinggi, rendah, sedang), atau juga berupa serangkaian

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



informasi verbal dan non-verbal yang disampaikan informan kepada peneliti untuk menjelaskan perilaku atau peristiwa yang sedang menjadi fokus perhatian.

b. Data Kuantitatif

Data kuantitatif adalah data dalam wujud angka yang merupakan satuan ukuran kuantitatif tertentu dari objek yang diteliti (misalnya, frekuensi, volume, berat dan sebagainya). Data kuantitatif merupakan hasil konversi dari data yang bersifat kualitatif ke dalam angka-angka kuantitatif. Dalam penelitian kuantitatif dikenal ada dua jenis data, yaitu: (1) data nominal, yaitu data yang memiliki ciri nominal atau data yang hanya dapat digolongkan secara terpisah berdasarkan kategori. Pada jenis data ini tidak ada penjenjangan yang ada hanya pemilahan berdasarkan kategori; dan (2) data kontinum, yaitu data yang bersifat bertingkat atau berjenjang.

4. Data Mining

Menurut Yahya (2022:1), “*Data Mining* adalah prosedur menambang dan menggali pengetahuan dari data.”

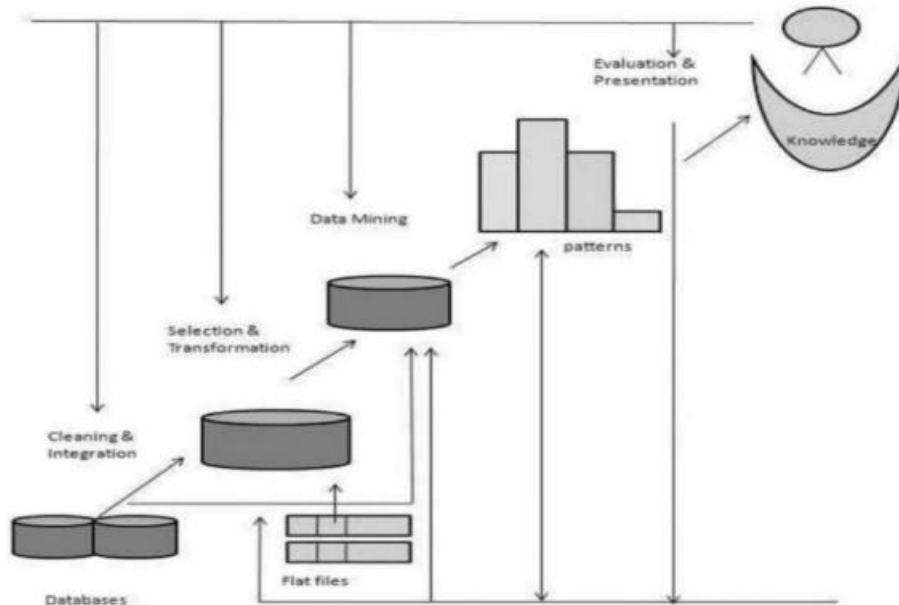
Menurut Pan Ning-Tan (2019:4) definisi Data Mining yaitu :

“*Data mining is the process of automatically discovering useful information in large data repositories.*”

Data mining adalah proses menemukan informasi yang berguna secara otomatis dalam repositori data yang besar.



data mining merupakan bagian dari Knowledge Discovery in Database (KDD), yang prosesnya mengubah data mentah menjadi informasi yang berguna. Berikut dibawah adalah gambar yang menjelaskan proses tersebut.



Gambar 2. 1
Proses Data Mining/Proses KDD
Sumber : Yahya (2022:24)

Proses *Data Mining* menurut Yahya (2022:23) :

a. Pembersihan data

Langkah ini melibatkan penghapusan data yang berisik dan memperbaiki ketidak konsistenan dalam data. Pembersihan data dilakukan sebagai langkah *preprocessing data* sambil menyiapkan data untuk *data warehouse*.

b. Integrasi data

Data dikumpulkan dari berbagai sumber yang berbeda dan digabungkan ke dalam satu sumber yang sama, seperti gudang data.

c. Pemilihan data

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian

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



Dalam langkah ini, data yang relevan dengan tugas analisis diambil dari database.

d. Transformasi data

Dalam langkah ini, data diubah ke dalam bentuk yang sesuai untuk penambangan dengan melakukan operasi ringkasan atau agregasi.

e. Data mining

Pada langkah ini, metode cerdas diterapkan untuk mengekstrak pola data.

f. Evaluasi pola

Pada langkah ini, pola data dievaluasi.

g. Presentasi pengetahuan

Pada langkah ini, pengetahuan direpresentasikan.

Menurut Pan Ning-Tan (2019 : 43), terdapat beberapa isu terkait data yang penting untuk keberhasilan data mining:

a. Tipe Data

Kumpulan data bervariasi dalam beberapa aspek. Misalnya, atribut yang digunakan untuk menggambarkan objek data dapat berupa tipe yang berbeda—kuantitatif atau kualitatif—dan kumpulan data sering memiliki karakteristik khusus; misalnya, beberapa dataset mengandung deret waktu atau objek yang memiliki hubungan eksplisit satu sama lain. Tipe data ini menentukan alat dan teknik yang dapat digunakan untuk menganalisisnya. Penelitian baru dalam data mining seringkali dipacu oleh kebutuhan untuk menyesuaikan diri dengan area aplikasi baru dan jenis data baru mereka.

b. Kualitas Data

C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

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

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

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

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



Data sering kali tidak sempurna. Meskipun sebagian besar teknik data mining dapat mentolerir beberapa tingkat ketidaksempurnaan, fokus pada pemahaman dan peningkatan kualitas data biasanya meningkatkan hasil analisis. Isu kualitas data yang sering perlu diatasi meliputi adanya noise dan outlier, data yang hilang, tidak konsisten, atau duplikat, serta data yang bias atau tidak representatif terhadap fenomena atau populasi yang seharusnya dijelaskan oleh data tersebut.

c. *Data Preprocessing*

Sering kali, data mentah perlu diproses agar lebih cocok untuk analisis. Satu tujuan mungkin adalah meningkatkan kualitas data, sementara tujuan lainnya berfokus pada modifikasi data agar lebih sesuai dengan teknik atau alat data mining tertentu. Misalnya, atribut kontinu, seperti panjang, terkadang perlu diubah menjadi atribut dengan kategori diskrit, seperti pendek, sedang, atau panjang, agar teknik tertentu dapat diterapkan. Selain itu, jumlah atribut dalam kumpulan data sering dikurangi karena banyak teknik lebih efektif ketika data memiliki jumlah atribut yang relatif kecil.

Menurut Pan Ning-Tan (2019 : 9) Tugas data mining umumnya dibagi menjadi dua kategori utama, yaitu:

a. *Predictive Task*

Tujuan dari tugas ini adalah untuk memprediksi nilai dari atribut tertentu berdasarkan nilai atribut lainnya. Atribut yang diprediksi biasanya dikenal sebagai variabel target atau variabel dependen, sedangkan atribut yang

digunakan untuk membuat prediksi disebut sebagai variabel penjelas atau variabel independen.

b. Descriptive Task

Di sini, tujuannya adalah untuk menemukan pola (korelasi, tren, kluster, trajektori, dan anomali) yang merangkum hubungan dasar dalam data. Tugas data mining deskriptif sering bersifat eksploratif dan sering kali memerlukan teknik pemrosesan lanjutan untuk memvalidasi dan menjelaskan hasilnya.

Data mining menggunakan berbagai konsep, seperti pengambilan sampel, estimasi, dan pengujian hipotesis yang berasal dari statistik, serta algoritma pencarian, teknik pemodelan, dan teori pembelajaran yang diambil dari kecerdasan buatan, pengenalan pola, dan pembelajaran mesin. Selain itu, data mining juga cepat mengadopsi ide-ide dari bidang lain, seperti optimisasi, komputasi evolusioner, teori informasi, pengolahan sinyal, visualisasi, dan pengambilan informasi, untuk mengatasi tantangan dalam menganalisis big data. Berikut adalah gambar yang menunjukkan hubungan antar bidang dalam data mining.

C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

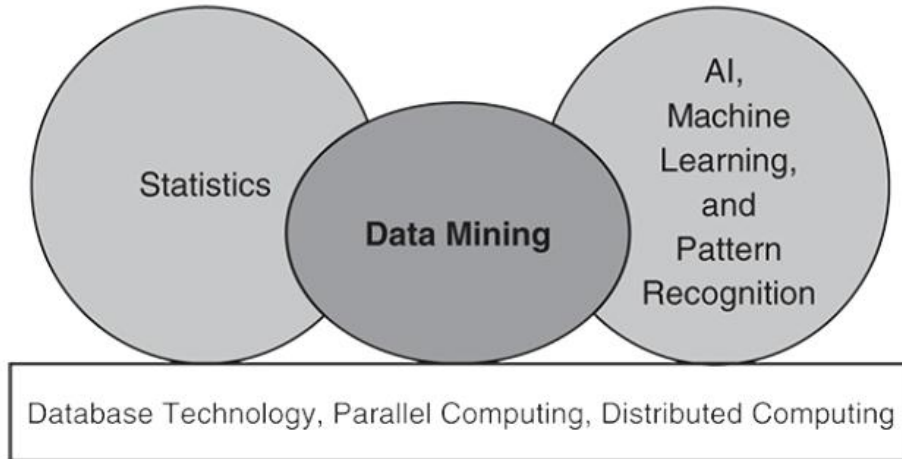
Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.





Gambar 2. 2
Hubungan Antar Bidang Data Mining

Sumber : Pan Ning-Tan (2019 : 7)

Menurut Pan Ning Tan (2019:9) data mining memiliki beberapa metode

pengolahan yaitu :

a. Prediksi (*Predictive*)

Mengacu pada tugas membangun model untuk variabel target berdasarkan variabel penjelas. Terdapat dua jenis tugas pemodelan prediktif: klasifikasi, yang digunakan untuk variabel target diskrit, dan regresi, yang digunakan untuk variabel target kontinu.

b. Asosiasi (*Association*)

Analisis asosiasi digunakan untuk menemukan pola yang menggambarkan fitur-fitur yang sangat terkait dalam data. Pola yang ditemukan biasanya diwakili dalam bentuk aturan implikasi atau subset fitur. Karena ukuran ruang pencariannya yang eksponensial, tujuan analisis asosiasi adalah untuk mengekstrak pola-pola yang paling menarik dengan cara yang efisien.

c. Klastering (*Clustering*)

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

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

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Analisis kluster bertujuan untuk menemukan kelompok pengamatan yang saling terkait, sehingga pengamatan yang berada dalam kluster yang sama lebih mirip satu sama lain dibandingkan dengan pengamatan yang berada di kluster lain. Klustering telah digunakan untuk mengelompokkan pelanggan yang saling terkait, menemukan area di lautan yang memiliki dampak signifikan terhadap iklim Bumi, serta untuk mengompresi data.

d. *Anomaly Detection*

Anomaly Detection adalah tugas mengidentifikasi pengamatan yang memiliki karakteristik yang sangat berbeda dari sisa data. Tujuan dari algoritma *Anomaly Detection* adalah untuk menemukan anomali yang sebenarnya dan menghindari penandaan objek normal sebagai anomali.

4. Text Mining

Menurut Rina Fitriana et al. (2022:363), “Text mining berurusan dengan data tidak terstruktur yang ditemukan dalam dokumen, email, media sosial, atau website.”

Menurut Prastyadi Wibawa Rahayu et al. (2024:163), “Dalam kasus penambahan teks, bagaimanapun, sumber data adalah kumpulan dokumen, dan pola yang menarik ditemukan bukan di antara catatan database yang di formalkan tetapi dalam data tekstual yang tidak terstruktur dalam koleksi dokumen”

Menurut Prastyadi Wibawa Rahayu et al. (2024:163) tujuan dari Text Mining dapat dibagi menjadi dua :

- a. Pengkategorisasian data teks (*Text Categorization*) Dalam pengkategorisasian, text mining dipergunakan sebagai alat untuk menemukan kategori yang sesuai dengan kelas yang telah ditentukan (Supervised Learning)



- b. Pengelompokan data teks (*Text Clustering*) Pada pengelompokan, text mining berfungsi sebagai alat untuk mengelompokkan data teks berdasarkan kesamaan karakteristik, dan clustering dapat digunakan untuk

(C) Hak cipta milik IBIKKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian

Menurut Prastyadi Wibawa Rahayu et al. (2024:163) Text Mining memiliki beberapa teknik yaitu :

- a. Klasifikasi Text

Teknik Klasifikasi Teks melibatkan penggunaan algoritma seperti Naïve bayes, Super Vecor Machine (SVM), atau Deep Learning untuk mengklasifikasikan teks kedalam kategori atau label yang relevan, misalnya klasifikasi teks dapat digunakan untuk memprediksi apakah sbuah email adalah spam atau bukan.

- b. Pengelompokan Text (*Clustering*)

Teknik pengelompokan teks melibatkan pengelompokan dokumen yang memiliki kesamaan berdasarkan pola atau karakteristik yang ada dalam teks. Metode seperti K-Means atau Hierarchical Clustering dapat digunakan untuk mengelompokkan dokumen yang serupa. Misalnya pengelompokan teks dapat digunakan untuk mengelompokkan berita-berita berdasarkan topik yang sama.

- c. Analisis Sentimen

Teknik analisis sentimen digunakan untuk menganalisis sentiment atau sikap opini dalam teks, apakah bersifat positif, negatif atau netral. Metode seperti analisis sentiment lexicon, atau Machine Learning dapat digunakan untuk menganalisis sentiment dalam taks, misalnya untuk melihat respon penggunaan terhadap produk atau layanan.

- d. Ekstraksi Informasi

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



Teknik ekstraksi informasi bertujuan untuk menemukan dan mengestrak informasi tertentu dari teks, seperti entitas nama, tanggal, lokasi, atau angka. Metode seperti Named Entity Recognition (NER) atau Regular Expression (RegEx) dapat di gunakan untuk mengestrak informasi tersebut. Contohnya, ekstraksi informasi dapat digunakan untuk mengestrak nama-nama produk dari ulasan pelanggan.

e. Pemrosesan Bahasa Alami (NLP)

Teknik pemrosesan bahasa alami melibatkan penerapan teknik dan model yang dirancang untuk memahami, menginterpretasi, dan menghasilkan teks yang lebih terstruktur dan bermakna. Pemrosesan bahasa alami dapat mencakup pemahaman teks, analisis sintaks, analisis semantik, atau generasi teks / penerapan NLP dapat di gunakan dalam berbagi tugas seperti chatbot, pemahaman pertanyaan atau terjemahan otomatis.

6. Analisis Sentimen

Pengertian Analisis Sentimen menurut Bing Liu (2020:1) adalah

“Sentiment analysis, also called opinion mining, is the field of study that analyzes people’s opinions, sentiments, appraisals, attitudes, and emotions toward entities and their attributes expressed in written text.”

Analisis sentimen, disebut juga penambangan opini, adalah bidang studi yang menganalisis opini, sentimen, penilaian, sikap, dan emosi masyarakat terhadap entitas dan atributnya yang diungkapkan dalam teks tertulis.

C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

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



Menurut Asrumi et al. (2023:1), “Analisis sentimen, juga disebut makna opini, adalah bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi masyarakat terhadap entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, topik, dan atributnya.”

Menurut Bing Liu (2020:1) ruang lingkup analisis sentimen

“*Many related names and slightly different tasks – for example, sentiment analysis, opinion mining, opinion analysis, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, and review mining – are now all under the umbrella of sentiment analysis.*”

Banyak nama terkait dan tugas yang sedikit berbeda – misalnya, analisis sentimen, penambangan opini, analisis opini, ekstraksi opini, penambangan sentimen, analisis subjektivitas, analisis pengaruh, analisis emosi, dan penambangan tinjauan – kini semuanya berada di bawah payung analisis sentimen.

7. Machine Learning

Menurut Imam Cholissodin et al. (2020:162) , “*Machine Learning (ML)* atau Mesin Pembelajaran adalah cabang dari AI yang fokus belajar dari data (learn from data), yaitu fokus pada pengembangan sistem yang mampu belajar secara “mandiri” tanpa harus berulang kali diprogram manusia.”

Menurut Budi Raharjo (2020:1), “Pembelajaran mesin adalah tentang mengekstraksi pengetahuan dari data.”

Pengertian *Machine Learning* menurut Aurelien Geron (2022:20) adalah

“*Machine learning is the science (and art) of programming computers so they can learn from data.*”

Machine learning adalah ilmu (dan seni) dalam memprogram komputer agar dapat belajar dari data

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

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

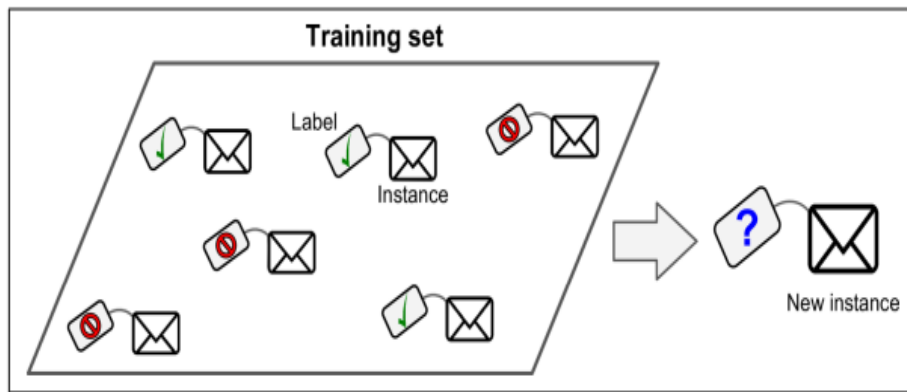


Menurut Aurelien Geron (2022:20) ada beberapa kriteria dalam *Machine Learning*,

seperti berikut :

a. Supervised Learning

Dalam Supervised learning, set pelatihan yang disediakan untuk algoritma mencakup solusi yang diinginkan, yang disebut label. Berikut ilustrasi yang dapat menjelaskan Supervised learning.



Gambar 2.3
Contoh Supervised learning
Sumber : Aurelien Geron (2022:33)

Tugas *supervised learning* yang umum adalah klasifikasi. Filter spam adalah contoh yang baik untuk ini: filter tersebut dilatih dengan banyak contoh email beserta kelasnya (spam atau bukan spam), dan harus belajar bagaimana mengklasifikasikan email baru. Tugas umum lainnya adalah memprediksi nilai numerik target, seperti harga mobil, berdasarkan serangkaian fitur (jarak tempuh, usia, merek, dll.). Tugas semacam ini disebut regresi. Untuk melatih sistem, Anda perlu memberikannya banyak contoh mobil, termasuk fitur-fitur dan targetnya (yaitu, harga mereka).

b. Unsupervised Learning

Hak Cipta Dilindungi Undang-Undang

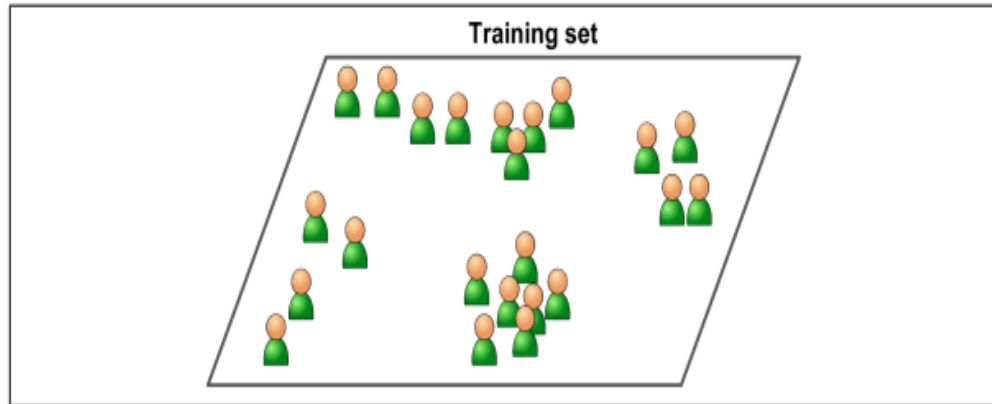
© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian

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



Dalam *unsupervised learning*, seperti yang mungkin Anda duga, data pelatihan tidak berlabel. Sistem mencoba belajar tanpa guru. Berikut ilustrasinya pada gambar 2. 4



Gambar 2. 4
Ilustrasi Unsupervised learning
Sumber : Aurelien Geron (2022:37)

Sebagai contoh, bayangkan terdapat banyak data mengenai pengunjung suatu blog. Dengan data ini, dapat digunakan algoritma *clustering* untuk mendeteksi kelompok-kelompok pengunjung yang memiliki kesamaan. Dalam proses ini, algoritma tidak diberikan informasi mengenai kelompok mana pengunjung blog itu dapat dimasukkan; algoritma tersebut akan menemukan pola hubungan antar pengunjung secara mandiri. Sebagai ilustrasi, algoritma mungkin akan menemukan bahwa 40% pengunjung merupakan remaja yang menyukai komik dan biasanya membaca blog setelah sekolah, sedangkan 20% lainnya adalah orang dewasa yang menyukai fiksi ilmiah dan mengakses blog saat akhir pekan. Jika algoritma *hierarchical clustering* digunakan, kelompok-kelompok tersebut juga bisa terbagi lagi menjadi subkelompok yang lebih kecil. Hasil ini dapat bermanfaat

C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

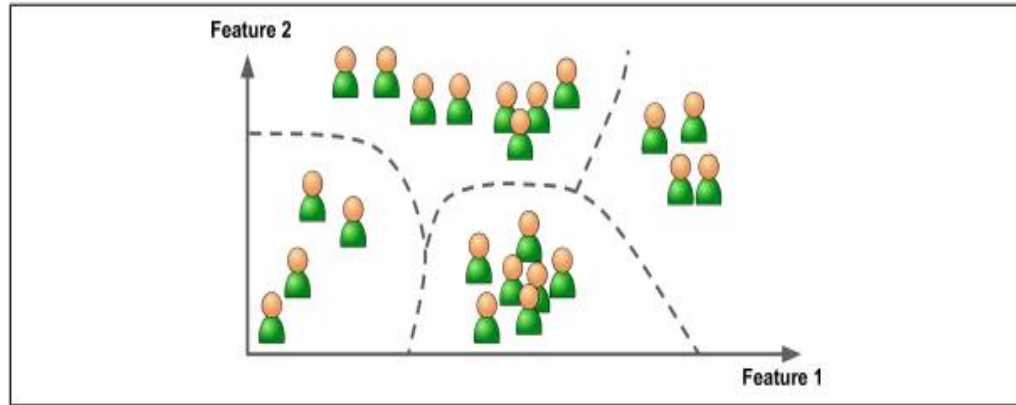
Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian

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



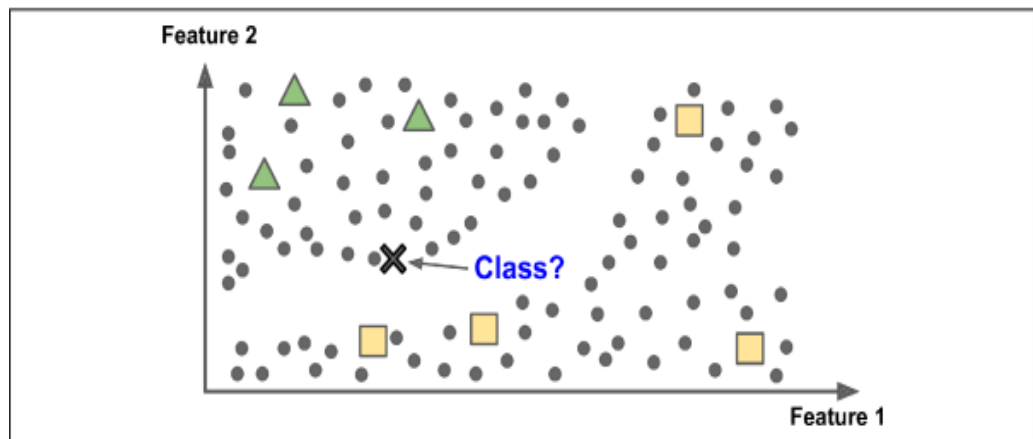
untuk menargetkan konten yang lebih spesifik bagi tiap kelompok pengunjung,
Berikut contoh gambar nya.



Gambar 2. 5
Contoh Unsupervised learning
Sumber : Aurelien Geron (2022:41)

c. Semi-supervised learning

Karena pelabelan data biasanya memakan waktu dan biaya yang cukup besar, sering kali terdapat banyak data yang tidak berlabel dan hanya sedikit data yang berlabel. Beberapa algoritma dapat menangani data yang sebagian berlabel. Pendekatan ini disebut *semi-supervised learning*. Berikut contoh gambar yang dapat menjelaskan *Semi-supervised learning*.



Gambar 2. 6
Contoh Semi-supervised learning
Sumber : Aurelien Geron (2022:45)

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian

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



Pada pembelajaran semi-terawasi dengan dua kelas (segitiga dan kotak), contoh-contoh tanpa label (lingkaran) berperan dalam membantu menentukan bahwa instance baru (tanda silang) lebih cocok untuk diklasifikasikan ke dalam kelas segitiga daripada kelas kotak, meskipun instance tersebut lebih dekat ke kotak yang berlabel.

Google Photos adalah contoh yang baik untuk ini. Setelah mengunggah semua foto keluarga ke layanan tersebut, sistem secara otomatis mengenali bahwa orang yang sama, A, muncul di foto 1, 5, dan 11, sedangkan orang lain, B, muncul di foto 2, 5, dan 7. Ini adalah bagian tak berlabel dari algoritma (clustering). Sekarang, yang dibutuhkan sistem hanyalah informasi tentang siapa orang-orang tersebut. Cukup tambahkan satu label untuk setiap orang, dan sistem akan dapat memberi nama pada setiap orang di semua foto, yang berguna untuk pencarian foto.

d. *Self-supervised learning*

Pendekatan lain dalam *Machine Learning* melibatkan pembuatan dataset berlabel penuh dari dataset yang sepenuhnya tidak berlabel. Setelah seluruh dataset diberi label, algoritma *Supervised learning* apa pun dapat digunakan. Pendekatan ini disebut *self-supervised learning*.

Sebagai contoh dan ilustrasi untuk menjelaskan *self-supervised learning*, jika terdapat dataset besar berisi gambar tanpa label, sebagian kecil dari setiap gambar dapat secara acak ditutupi, kemudian model dilatih untuk memulihkan gambar aslinya. Selama pelatihan, gambar yang ditutupi digunakan sebagai input untuk model, dan gambar aslinya digunakan sebagai label.

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



8. Twitter

C Hak Cipta milik BI KKG (Institut Bisnis dan Informatika Kwik Kian Gie) Twitter adalah layanan untuk teman, keluarga, dan rekan kerja agar dapat berkomunikasi dan tetap terhubung melalui pertukaran pesan singkat. Pengguna bisa mengirimkan postingan, yang dapat berisi foto, video, tautan, dan teks. Postingan tersebut akan muncul di halaman profil, dikirimkan ke pengikut (*followers*) dan dapat dicari dengan fitur pencarian di Twitter.

Twitter memiliki beberapa fitur, berikut fiturnya :

a. *Post*

Post adalah pesan yang ingin dipublikasikan di Twitter dan dapat berisi foto, video, tautan, dan teks. Klik atau ketuk tombol post untuk mempublikasikan pembaruan ke profil.

b. *Repost*

Repost adalah pengulangan dari sebuah postingan. Fitur Repost di twitter membantu pengguna dan orang lain untuk membagikan postingan tersebut dengan cepat kepada semua pengikut. Pengguna dapat melakukan Repost pada postingan sendiri atau postingan orang lain.

Terkadang, orang menulis "RT" di awal sebuah postingan untuk menunjukkan bahwa mereka membagikan ulang konten milik orang lain. Ini bukanlah perintah atau fitur resmi di Twitter, tetapi menandakan bahwa mereka mengutip postingan dari pengguna lain.

c. *Replies*

Replies adalah balasan terhadap postingan orang lain. Replies dapat dilakukan dengan mengklik atau mengetuk ikon reply dari sebuah postingan.

Hak Cipta Dilindungi Undang-Undang

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

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

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Ketika dua orang saling membalas, hanya orang-orang terkait, seperti yang mengikuti akun yang membalas atau yang terlibat dalam percakapan, yang akan melihat balasan tersebut diberanda.



Hak cipta milik © IKKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Python

Pengertian Python menurut Al Sweigart (2020:30) adalah

“Python is a programming language (with syntax rules for writing what is considered valid Python code) and the Python interpreter software that reads source code (written in the Python language) and performs its instructions.”

Python adalah bahasa pemrograman (dengan aturan untuk menulis sintaks yang benar dalam Python) dan perangkat lunak *interpreter* Python yang membaca kode sumber (yang ditulis dalam bahasa Python) dan menjalankan perintah-perintahnya.

Pengertian Python menurut Gowrishankar S. dan Veena A. (2019:8) adalah

“Python is a multi-paradigm programming language having full support for Object-oriented programming and Structured programming and there are a number of language features which support Functional programming.”

Python adalah bahasa pemrograman multi-paradigma yang sepenuhnya mendukung pemrograman berorientasi objek (OOP) dan pemrograman terstruktur, serta memiliki sejumlah fitur bahasa yang mendukung pemrograman fungsional.

Menurut Eric Matthes (2019:36) alasan mengapa menggunakan bahasa pemrograman Python adalah

“Python is an incredibly efficient language: your programs will do more in fewer lines of code than many other languages would require. Python’s syntax will also help

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



you write “clean” code. Your code will be easy to read, easy to debug, and easy to extend and build upon compared to other languages.”

C Hak Cipta Milik IBIKKGG (Institut Bisnis dan Informatika Kwik Kian Gie)
Python adalah bahasa yang sangat efisien: program Anda akan melakukan lebih banyak hal dengan lebih sedikit baris kode dibandingkan dengan bahasa lain. Sintaks Python juga akan membantu Anda menulis kode yang "bersih". Kode Anda akan lebih mudah dibaca, lebih mudah untuk diperbaiki, dan lebih mudah untuk dikembangkan dan dibangun dibandingkan dengan bahasa lain.

Menurut Harris Wang (2023:19) Python sangat kuat karena ada 2 alasan :

- a. Python memiliki banyak konstruksi bahasa dan pernyataan yang kuat untuk merepresentasikan data dan operasi pada berbagai jenis data.
- b. bahwa ada ribuan *third-party packages/libraries/modules* yang tersedia untuk berbagai kebutuhan pemrograman, selain pustaka standar besar yang disertakan dalam setiap rilis Python.

10. Orange Data Mining

Orange Data Mining adalah perangkat lunak *open-source* yang digunakan untuk menganalisis data dan *Machine Learning*, memungkinkan pengguna untuk menggabungkan berbagai komponen atau "widget" dalam alur kerja yang terstruktur. Setiap widget memiliki fungsi tertentu, seperti mengambil data, membersihkan data (pra-pemrosesan), memvisualisasikan data, membangun model, atau mengevaluasi hasil analisis. Dengan menggabungkan widget-widget ini, pengguna dapat membangun alur kerja yang kompleks dengan cara yang mudah.

Berikut beberapa fitur yang dapat dilakukan pada Orange Data Mining :

Hak Cipta Dilindungi Undang-Undang

1. Dilarang menyalin atau seluruhnya atau sebagian tanpa izin tertulis dari pencipta atau sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKGG.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKGG.



a. ***Interactive Data Visualization***

Orange Data Mining memanfaatkan visualisasi data untuk membantu menemukan pola tersembunyi dan memberikan pemahaman lebih dalam tentang prosedur analisis data. Beberapa visualisasi yang tersedia termasuk scatter plot, box plot, dan histogram, serta visualisasi model tertentu seperti dendrogram, silhouette plot, dan pohon keputusan. Berbagai jenis visualisasi lainnya juga tersedia melalui add-on, seperti visualisasi jaringan, word cloud, peta geografis, dan banyak lagi.

b. ***Visual Programming***

Orange adalah alat *data mining* yang baik untuk pemula maupun data scientist ahli. Berkat antarmuka penggunaannya, pengguna dapat fokus pada analisis data daripada menghabiskan waktu untuk *coding* yang melelahkan, sehingga pembuatan pipeline analitik data yang kompleks menjadi lebih mudah.

c. ***Extensions***

Dengan memanfaatkan Add-on khusus, dapat melakukan pemrosesan bahasa alami dan penambangan teks, menganalisis jaringan, melakukan penambangan aturan asosiasi, atau menangani masalah keadilan dalam pembelajaran mesin. Beberapa di antaranya juga menangani data khusus, seperti deret waktu, set data kelangsungan hidup, spektra, atau ekspresi gen.

11. Penelitian

Menurut Umesh Kumar dan D.P. Kothari (2022:1) penelitian adalah

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

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

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

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



“Research is defined as a systematic, controlled, empirical and critical investigation of hypothetical propositions about the presumed relationship about various phenomena. Research is a systematic investigation to find answers to a problem.”

Penelitian didefinisikan sebagai investigasi yang sistematis, terkontrol, empiris, dan kritis terhadap proposisi hipotesis mengenai hubungan yang diduga antara berbagai fenomena. Penelitian adalah investigasi yang sistematis untuk menemukan jawaban atas suatu masalah.

12. Penelitian Kuantitatif

Menurut Umesh Kumar dan D.P. Kothari (2022:6) penelitian kuantitatif adalah

“Quantitative research is variables based whereas qualitative research is attributes based. Quantitative research is based on measurement or quantification of the phenomenon under study. That is, it is data based and hence more objective and popular.”

Penelitian kuantitatif berbasis variabel, sedangkan penelitian kualitatif berbasis atribut. Penelitian kuantitatif didasarkan pada pengukuran atau kuantifikasi fenomena yang sedang diteliti. Artinya, penelitian ini berbasis data sehingga lebih objektif dan lebih populer.

B. Penelitian Terdahulu

Analisis Sentimen Data Twitter Terkait ChatGPT Menggunakan

Orange Data Mining

Kesimpulan yang bisa diambil dari penelitian yang dilakukan oleh Tri Yuli Pahtoni dan Handaru Jati, yaitu :

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



ChatGPT menjadi salah satu model yang dirilis oleh OpenAI. ChatGPT adalah model chatbot yang dilatih untuk dapat memahami dan menghasilkan teks dalam berbagai bahasa. Kecerdasan chatGPT tersebut menimbulkan kekhawatiran di antara beberapa pendidik. Publik saat ini cenderung menggunakan media sosial dalam menyampaikan opini terhadap sesuatu. Twitter menjadi salah satu media sosial yang sering digunakan dalam menyampaikan opini tersebut. Kumpulan tweet dari publik tersebut menunjukkan perasaan orang terhadap suatu peristiwa. Hal ini dapat menjadi kumpulan data yang dapat digunakan untuk mengukur sentimen publik terhadap rilisnya chatGPT.

Metode analisis sentimen disajikan dan dijelaskan secara rinci dengan menggunakan perangkat orange data mining. Hasilnya menunjukkan bahwa sebagian besar masyarakat mengungkapkan kebahagiaan terhadap rilisnya chatGPT, menempatkan emosional terkejut pada kategori tertinggi kedua berdasarkan publik dengan bahasa Inggris. Berbeda hasil dengan analisis sentimen berbahasa Indonesia, sebagian besar terkejut atas rilisnya chatGPT oleh OpenAI. Hasil analisis sentimen menunjukkan bahwa chatGPT mendapatkan reaksi netral oleh publik, artinya perilis chatGPT tidak terdapat penolakan oleh publik.

2. Analisis Sentimen Terhadap Penggunaan Chatgpt Berdasarkan Twitter Menggunakan Algoritma Naive Bayes

Kesimpulan yang bisa diambil dari penelitian yang dilakukan oleh Dwi Transiska, Dimas Febriawan, dan Firman Noor Hasan, yaitu :

Penelitian terkait analisis sentimen terhadap penggunaan ChatGPT berdasarkan Twitter menggunakan algoritma Naive Bayes ini dengan data awalnya berjumlah 300 data dengan topik ChatGPT pada rentan waktu pengambilan dari tanggal 13 juni 2023 sampai

Hak Cipta Dilindungi Undang-Undang

Hak Cipta Milik IBI BKG (Institut Bisnis dan Informatika Kwik Kian Gie)

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

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

b. Pengutipan tidak merugikan kepentingan yang wajar IBI BKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBI BKG.

Institut Bisnis dan Informatika Kwik Kian



30 juni 2023, dan diolah dengan beberapa fitur pada proses Preprocessing yang meliputi Transform Cases, tokenize, Filter token, dan Filter stopword yang ada pada Rapidminer.

Penelitian ini juga menggunakan rasio perbandingan untuk data latih dan data ujinya sebesar 80:20. Hasil dari prediksi ini juga didapatkan 158 data positif dan 60 data negatif.

Itu artinya kehadiran ChatGPT menjadi hal yang mendapat respon positif di kalangan netizen twitter. Dalam segi implementasi algoritma Naïve Bayes, peneliti mendapatkan hasil precision sebesar 89,19%, recall sebesar 95,18% dan nilai accuracy sebesar 96,55%.

Nilai accuracy tersebut membuktikan bahwa algoritma Naïve Bayes merupakan algoritma yang memiliki nilai akurasi yang paling baik bahkan mendekati sempurna.

Hasil akurasi pada penerapan metode ini berdasarkan dari tweet user Twitter yang menggunakan ChatGPT sebagai kebutuhan sehari-hari, jika di lihat berdasarkan pelabelan dan akurasi, sentimen positif lebih banyak dari pada sentimen negatif. Hal ini dikarenakan mayoritas pengguna ChatGPT merasa terbantu dengan hadirnya chatbot yang mampu meringankan pekerjaan manusia sehari-hari.

3. Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naive Bayes

Kesimpulan yang bisa diambil dari penelitian yang dilakukan oleh Yuma Akbar dan Tri Sugiharto, yaitu :

Hasil analisis sentimen telah ditampilkan dalam pengujian, dimana didapat hasil dari pemodelan analisis decision tree sebanyak 754 tweets yang diolah = 43 (negatif), 711 (positif). Bobot kata yang tidak terdeteksi atau diurutan pertama ada 31, kata 'chatgpt' diurutan kedua yakni 17, dan seterusnya. Selanjutnya, hasil dari pemodelan analisis naïve bayes sebanyak 754 tweets yang diolah = 436 (negatif), 318 (positif). Bobot kata yang

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI BKKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian

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

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

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



tidak terdeteksi atau diurutan pertama ada 31, kata 'chatgpt' diurutan kedua yakni 17, dan seterusnya. Perbandingan kedua algoritma mendapat hasil prediksi positif(class precision) = 74.88%, prediksi negatif(class precision) = 100.00%, class recall(true positif) = 100.00%, class recall(true negatif) = 30.14%, yang mana didapat average : accuracy = 77.33%, precision = 100.00%, recall = 30.18%. Kesimpulan dari hasil tersebut adalah positif, artinya respon atau sentimen yang diambil dari sosial media twitter khususnya di Indonesia menunjukkan feedback baik terhadap ChatGPT.

4 Analisis Sentimen Pengguna X terhadap Chatgpt dengan Algoritme Naive Bayes

Kesimpulan yang bisa diambil dari penelitian yang dilakukan oleh Bayu Purbayanto dan Teguh Nurhadi Suharsono, yaitu :

Berdasarkan hasil penelitian, dari 4.861 data hasil crawlingdidapatkan data sebanyak 2.884. Setelah dilakukan preprocessingdata denganpelabelanmenggunakan Lexicon based,dihasilkan1.342 data positif, 456 data negatif,dan 1.084 data netral.

Algoritma Naïve Bayes cukup efektif untuk digunakan dalam analisis sentimen dengan akurasi sebesar 87,175043% dan dengan sentimen positif sebanyak 1.543 data,

sentimen negatif 318 data, dan sentimen netral sebanyak 1.023. Dengan 53,5% sentimen positif yang didapatkan, maka dapat dilihat bahwa pengguna X terbantu dan menerima penggunaan Chatgpt dalam menyelesaikan permasalahan yang dimilikinya.

Namun, tentunya sentimen negatif yang diberikan tidak dapat diabaikan. OpenAI dapat melakukan pengembangan dalam memberikan kualitas informasi yang lebih akurat untuk para pengguna Chatgpt agar dapat menjadi model yang lebih sempurna dan dapat memberikan manfaat yang lebih signifikan.



5. Analisis Sentimen Terhadap Aplikasi Chatgpt Pada Twitter Menggunakan

© Algoritma Naïve Bayes

Kesimpulan yang bisa diambil dari penelitian yang dilakukan oleh Muhamad Ilmar Rifaldi, Yudhi Raymond Ramadhan, dan Irsan Jaelani, yaitu :

Dari hasil penelitian dengan menggunakan data sebanyak 301 tweet pengklasifikasian data dengan metode Naïve Bayes dan pengujian menggunakan Confussion Matrix didapatkan hasil akurasi sebesar 80%. Dengan nilai Precission sebesar 80,95% dan nilai Recall sebesar 89,47%. Dapat dinilai bahwa analisis sentimen dengan algoritma Naïve bayes mampu mengklasifikasikan sentimen pada twitter dengan baik dengan memperoleh akurasi sebesar 80%. Juga dapat disimpulkan bahwa sentimen masyarakat terhadap aplikasi ChatGPT pada media sosial Twitter cenderung positif dengan jumlah presentasi tweet sebanyak 74,09%. Hasil penelitian ini dapat menjadi acuan bagi pemerintah khususnya bagi kementrian informasi dan komunikasi yang bertanggung jawab untuk memfilter aplikasi yang beredar pada masyarakat, bahwasanya aplikasi ChatGPT ini cenderung diterima oleh masyarakat Indonesia melihat lebih banyaknya sentimen positif dari masyarakat pada Twitter. Sehingga tidak perlu dilakukan pelarangan atau pemblokiran terhadap ChatGPT.

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian

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