BAB 1

PENDAHULUAN

1.1 Latar Belakang

Ragam bahasa yang digunakan oleh penutur tidak hanya mencerminkan maksud komunikasi, tetapi juga memperlihatkan emosi, sikap, dan nilai-nilai sosial. Salah satu bentuk ragam tersebut adalah kosakata kasar dan tidak kasar yang kerap muncul dalam interaksi, termasuk dalam komunikasi di media sosial. Kosakata kasar umumnya digunakan dalam konteks yang penuh emosi dan mengandung makna negatif seperti hinaan atau makian, sementara kosakata tidak kasar cenderung digunakan secara sopan dan mengikuti norma kesantunan. Pada platform seperti *Instagram*, keberadaan kosakata kasar dalam kolom komentar sering kali tidak terkontrol dan dapat menimbulkan masalah sosial seperti perundungan atau ujaran kebencian, sehingga dibutuhkan sistem yang mampu memilah kedua jenis kosakata ini secara otomatis dan efisien [1].

Pertumbuhan media sosial yang begitu cepat khususnya *Instagram* telah mempermudah pengguna dalam membagikan konten berupa gambar maupun tulisan. Namun, di sisi lain, hal ini juga memunculkan potensi maraknya penyebaran bahasa yang tidak pantas dan tindakan perundungan secara daring (*cyberbullying*). Menurut laporan terkini, *Instagram* termasuk dalam daftar platform dengan tingkat kasus *cyberbullying* yang tinggi, dimana tindakan agresif secara verbal dapat menimbulkan dampak psikologis yang serius, seperti tekanan mental, stres berkepanjangan, gangguan kecemasan, bahkan sampai pada munculnya keinginan untuk mengakhiri hidup [2].

Membangun sistem klasifikasi kosakata kasar dan tidak kasar menjadi hal yang penting karena saat ini jumlah komentar di media sosial seperti *Instagram* sangat banyak dan terus bertambah setiap waktu. Jika pengawasan dilakukan secara manual, tentu akan membutuhkan banyak waktu dan tenaga. Dengan bantuan sistem klasifikasi otomatis, komentar yang mengandung kata-kata kasar dapat dikenali dan disaring dengan lebih cepat dan efisien. Langkah ini bertujuan untuk menjaga media sosial tetap menjadi ruang yang aman, nyaman, dan bebas dari ujaran kebencian, perundungan, serta mendorong terciptanya komunikasi yang lebih sehat di dunia digital. Adapun tokoh pengguna *Instagram* yang dipilih sebagai sumber pengumpulan data dalam penelitian ini akun dari adalah Bapak Gibran Rakabuming Raka. Pemilihan Bapak Gibran Rakabuming Raka sebagai subjek dalam penelitian ini didasarkan pada status beliau sebagai figur publik yang aktif menggunakan media sosial, khususnya *Instagram*, serta posisinya yang penting sebagai Wakil Presiden Indonesia pada saat penelitian berlangsung. Sebagai sosok politisi muda yang juga memiliki latar belakang sebagai pengusaha dan merupakan putra dari Presiden Joko Widodo, akun *Instagram* beliau kerap menjadi pusat perhatian publik dan menerima beragam tanggapan dari warganet. Berdasarkan hasil pengamatan, setiap postingan Bapak Gibran Rakabuming Raka menerima sekitar 100 hingga 500 komentar per jam, tergantung pada tema postingan yang diunggah. Tanggapan tersebut mencakup berbagai bentuk ekspresi, mulai dari dukungan, kritik, hingga komentar yang mengandung kosakata kasar. Tingginya interaksi dan keberagaman komentar menjadikan akun ini sebagai sumber data yang tepat dan representatif dalam penelitian klasifikasi

kosakata kasar dan tidak kasar. Di samping itu, penelitian ini juga dapat memberikan wawasan mengenai pola komunikasi masyarakat Indonesia di ruang digital terhadap figur politik muda dalam konteks media sosial.

Machine learning saat ini memainkan peran yang sangat vital dalam berbagai bidang pemrosesan bahasa alami, khususnya dalam klasifikasi teks. Salah satu penerapannya yang berkembang pesat adalah dalam mendeteksi ujaran kasar, yang memerlukan analisis pola bahasa secara mendalam dan akurat. Metode ini memungkinkan sistem untuk mengenali pola-pola tertentu dalam teks yang mengindikasikan perilaku agresif atau tidak pantas. Sebagai ilustrasi, Rahayu dan tim peneliti (2023) melakukan studi dengan menerapkan algoritma *Random Forest* dalam upaya mengidentifikasi gejala depresi dan kecemasan pada konten *tweet*. Hasil dari penelitian tersebut menunjukkan bahwa pendekatan *Random Forest* mampu memberikan hasil yang sangat baik, yaitu mencapai akurasi sebesar 95,7%. Keberhasilan ini menjadi bukti bahwa model *machine learning*, khususnya *Random Forest*, sangat potensial dalam melakukan klasifikasi teks yang kompleks, termasuk dalam konteks analisis psikologis dan deteksi ujaran kasar di media sosial [3].

Random Forest dipilih karena kemampuannya yang baik dalam mengelola data yang tidak seimbang serta menangani fitur-fitur kompleks seperti representasi teks dalam pemrosesan bahasa alami. Algoritma ini bekerja dengan cara membentuk sejumlah pohon keputusan secara paralel, lalu menggabungkan hasilnya untuk menghasilkan prediksi akhir yang lebih akurat, sehingga mampu meminimalkan risiko *overfitting* dan meningkatkan kemampuan generalisasi

model. Dalam konteks Bahasa Indonesia, penelitian oleh Arsya dan Elsera (2023) menerapkan *Random Forest* untuk mendeteksi sarkasme dan memperoleh akurasi sebesar 53%, sedangkan studi lain oleh Santoso dan rekan-rekan (2023) dalam klasifikasi komentar di *Instagram* menunjukkan hasil yang lebih tinggi dengan tingkat akurasi mencapai 84% [4].

Walaupun penelitian mengenai *cyberbullying* di platform *Instagram* sudah cukup banyak dilakukan, kajian yang secara khusus menyoroti klasifikasi kosakata kasar masih tergolong minim. Penelitian oleh Wijaya & Widyaningrum (2024) serta Ramayasa dan kolega (2023) memang telah menyinggung topik yang berkaitan, namun penggunaan kata-kata kasar yang berdiri sendiri, terlepas dari konteks sentimen positif maupun negatif, belum dianalisis secara mendalam. Selain itu, dimensi emosional dan aspek pragmatik dari ujaran kasar juga belum menjadi fokus utama dalam penelitian-penelitian tersebut [5].

Berdasarkan uraian-uraian pada paragraf sebelumnya, maka penelitian ini mengambil judul "Klasifikasi kosakata kasar dan tidak kasar pada komentar *Instagram* menggunakan metode *Random Forest* berbasis *Machine Learning*".

1.2 Rumusan Masalah

Bagaimana merancang dan membangun sebuah sistem klasifikasi kosakata kasar dan tidak kasar pada komentar *Instagram* menggunakan metode *Random Forest* berbasis *Machine Learning*?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk merancang dan membangun sebuah sistem klasifikasi kosakata kasar dan tidak kasar pada komentar *Instagram* menggunakan metode *Random Forest* berbasis *Machine Learning*.

1.4 Batasan Penelitian

Dalam penelitian ini, terdapat beberapa batasan agar fokus penelitian tetap terarah, yaitu:

1. Data yang digunakan berupa komentar publik pada platform *Instagram* Gibran Rakabuming Raka yang ditulis dalam bahasa Indonesia.
2. Jenis komentar yang diklasifikasikan terbatas pada dua kategori, yaitu kasar dan tidak kasar.
3. Algoritma *machine learning* yang digunakan dalam klasifikasi adalah *Random Forest*.
4. Proses *preprocessing* teks meliputi pembersihan data, tokenisasi, dan normalisasi sederhana tanpa memperhatikan konteks sarkasme atau makna ganda.

1.5 Manfaat Penelitian

Penelitian ini bermanfaat untuk menghadirkan solusi dalam mengidentifikasi dan mengelompokkan berbagai jenis kosakata pada komentar *Instagram*, khususnya yang bermuatan kasar atau negatif, dengan bantuan algoritma *Random Forest* berbasis *Machine Learning*.

1.6 Sistematika Penulisan

Sistematika penulisan dari skripsi ini terdiri dari pokok-pokok permasalahan yang akan dibahas pada masing-masing yang diuraikan menjadi beberapa bagian sebagai berikut :

BAB 1 PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini berisi teori-teori yang digunakan dalam penelitian, seperti pembelajaran mesin (*machine learning*), algoritma *random forest*, *text classification* dan *natural language processing* (*NLP*), media sosial dan komentar di *instagram,* kosakata kasar di media sosial, *preprocessing* data teks bahasa indonesia, *feature extraction* dan *word embedding, phyton, Instaloader,* Sastrawi, *Scikit-learn*, *Pandas* dan *Numpy, Matplotlib* dan *Seaborn* dan penelitian terkait.

BAB 3 METODELOGI PENELITIAN

Bab ini berisi tahapan-tahapan penelitian, mulai dari identifikasi masalah, perumusan masalah, pengumpulan data, analisa sistem, perancangan sistem, implementasi sistem, pengujian sistem, kesimpulan dan saran.

BAB 4 ANALISA DAN PERANCANGAN

Bab ini berisi Klasifikasi Kosakata Kasar pada Komentar *Instagram* menggunakan metode *Random Forest* berbasis *Machine Learning*.

BAB 5 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi implementasi dan pengujian pada aplikasi yang akan dibangun.

BAB 6 PENUTUP

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan serta saran untuk penelitian atau pengembangan sistem di masa mendatang.

BAB 2
LANDASAN TEORI

2.1. Pembelajaran Mesin (*Machine Learning*)

Machine Learning adalah cabang dari kecerdasan buatan yang memungkinkan suatu sistem komputer untuk secara otomatis mempelajari pola dari data yang diamati. Proses ini dimulai dengan pengumpulan dan observasi data dalam jumlah besar, kemudian dilanjutkan dengan pembentukan model matematis atau statistik berdasarkan data tersebut. Model yang dihasilkan kemudian digunakan untuk melakukan prediksi atau mengambil keputusan dalam menyelesaikan berbagai masalah, baik yang bersifat klasifikasi, regresi, maupun pengenalan pola. Kemampuan utama dari Ma*chine Learning* terletak pada kemampuannya untuk terus belajar dan meningkatkan performa seiring bertambahnya data baru, tanpa perlu diprogram ulang secara eksplisit untuk setiap tugas tertentu [6].

Menurut Eriana & Zein (2023), *machine learning* merupakan suatu pendekatan dalam komputasi yang memanfaatkan data dan algoritma untuk meniru cara manusia belajar melalui pengalaman. Dalam proses ini, sistem komputer tidak hanya menjalankan instruksi tetap yang telah diprogram, tetapi secara bertahap dapat mengidentifikasi pola dari data yang dianalisis. Seiring waktu, akurasi dari hasil prediksi yang dilakukan oleh mesin akan meningkat karena model terus diperbarui berdasarkan data baru yang masuk. Dengan demikian, mesin memperoleh kemampuan untuk “belajar” dan membuat keputusan secara mandiri, seperti halnya manusia, tanpa harus selalu diberikan

arahan eksplisit melalui pemrograman konvensional. Pendekatan ini memungkinkan mesin menjadi lebih cerdas dan adaptif dalam menghadapi beragam situasi dan variasi data [7].

Dalam praktiknya, pembelajaran mesin (*machine learning*) terbagi ke dalam beberapa jenis, seperti *supervised learning* (dengan data berlabel), *unsupervised learning* (tanpa label), dan *reinforcement learning*. Untuk tugas seperti mengklasifikasikan komentar kasar atau konten perundungan siber di *Instagram*, pendekatan yang paling umum adalah *supervised learning*, karena data yang digunakan telah diberi label sebelumnya (misalnya: kasar atau tidak kasar). Agung Wijoyo dan rekan-rekan (2024) menguraikan bahwa *machine learning* mampu membentuk model prediktif secara otomatis dari kumpulan data tanpa instruksi eksplisit, menyerupai cara manusia belajar melalui pengalaman atau contoh [8].

2.2. Algoritma *Random Forest*

Random Forest merupakan algoritma *ensemble* yang bekerja dengan membentuk sejumlah pohon keputusan (*decision tree*) dari sampel data acak (*bootstrap*) dan subset fitur yang dipilih secara acak pula. Pendekatan ini membantu mengurangi risiko *overfitting* dan meningkatkan akurasi hasil prediksi. Karena kemampuannya dalam menangani data yang tidak terstruktur dan penuh variasi, algoritma ini sering digunakan dalam tugas klasifikasi teks di media sosial [9].

Selain penerapannya untuk klasifikasi komentar kasar, algoritma *Random Forest* juga terbukti efektif di berbagai bidang lain. Penelitian oleh Ramayasa et

al. (2023) membandingkan beberapa metode representasi teks seperti LSTM, Word2Vec, dan TF-IDF dalam model *Random Forest* pada data komentar *Instagram*. Hasilnya menunjukkan bahwa kombinasi TF-IDF dengan *Random Forest* menghasilkan akurasi tertinggi sebesar 96,25%, mengindikasikan bahwa algoritma ini sangat andal dalam mengolah fitur teks [10].

Algoritma *Random Forest* sering menunjukkan kinerja yang lebih baik dibandingkan metode klasik seperti *Naive Bayes* dan K-NN, terutama dalam tugas klasifikasi teks di media sosial. Penelitian yang dilakukan oleh Tamrizal dan Yaqin (2023) mengkaji respons masyarakat terhadap layanan BPJS Kesehatan melalui *Twitter* dengan membandingkan tiga algoritma tersebut. Dari hasil penelitian, *Random Forest* memperoleh tingkat akurasi tertinggi sebesar 87%, mengungguli *Naive Bayes* yang mencapai 80% dan K-NN sebesar 67% dalam mengelompokkan tweet menjadi kategori positif, negatif, dan netral. Temuan ini menguatkan bahwa *Random Forest* merupakan metode yang efektif dalam mengolah data teks berbahasa Indonesia yang memiliki keragaman dan kompleksitas tinggi [11].

2.3. *Text Classification dan Natural Language Processing (NLP)*

Klasifikasi teks merupakan salah satu bidang dalam *Natural Language Processing* (NLP) yang berfokus pada pengelompokan dokumen atau kalimat ke dalam kategori tertentu berdasarkan konten dan struktur bahasa yang digunakan. Proses ini mencakup sejumlah tahap awal seperti pembersihan data (*cleaning*), pemecahan teks menjadi unit-unit kecil (*tokenisasi*), normalisasi kata, serta konversi teks ke dalam bentuk numerik, seperti menggunakan TF-IDF atau teknik

word embedding. Penelitian oleh Berliansyah dan tim (2024) yang meneliti sentimen komentar pada akun Instagram Presiden Jokowi menunjukkan bahwa penggunaan algoritma *Naive Bayes* yang didukung oleh proses *preprocessing* yang tepat mampu menghasilkan akurasi hingga 84,1%, memperlihatkan peran penting NLP dalam pengolahan dan klasifikasi data media sosial [12].

Penelitian Kusni & Dewi (2024) yang berjudul *Analisis Sentimen Komentar pada Posting Instagram Akun "StandWithUs"* menggunakan kombinasi algoritma Naive Bayes dan *Decision Tree* dengan pendekatan N-Gram serta TF-IDF untuk ekstraksi fitur. Hasilnya menunjukkan akurasi lebih dari 88% dalam mengklasifikasikan komentar menjadi sentimen positif, negatif, dan netral. Studi ini juga menegaskan bahwa pemilihan fitur yang tepat berperan penting dalam meningkatkan akurasi klasifikasi pada teks media sosial [13].

Pendekatan gabungan (*hybrid*) dan ensemble kini banyak digunakan untuk meningkatkan efektivitas dalam tugas klasifikasi. Dalam penelitian berjudul *Detection of Cyberbullying using SVM, Naive Bayes, and Random Forest* oleh Monica Fachrita Ruziqiana dkk. (2023), dilakukan perbandingan antara tiga algoritma klasifikasi pada data komentar *Instagram*. Hasil penelitian menunjukkan bahwa SVM dengan penyesuaian parameter (parameter tuning) mampu meraih akurasi hingga sekitar 97,8%. Sementara itu, perpaduan antara metode klasik dan teknik ensemble terbukti memberikan lapisan prediksi tambahan yang lebih optimal [14].

2.4. Media Sosial dan Komentar di *Instagram*

Instagram kini telah berevolusi menjadi salah satu media sosial terpopuler yang memungkinkan pengguna menyebarkan konten visual sekaligus menyampaikan pendapat mereka melalui fitur komentar. Di Indonesia, *platform* ini termasuk yang paling banyak digunakan, menjadikannya lahan yang subur untuk pengumpulan data dalam riset berbasis teks. Penelitian yang dilakukan oleh Aziza dan Nur (2021) mengungkapkan bahwa berbagai bentuk bahasa gaul dan eufemisme banyak ditemukan dalam komentar Instagram, menunjukkan keberagaman sosial dan linguistik yang terjadi di dalamnya [15].

Penelitian oleh Hayati, Sirulhaq, dan Mahsun (2025) melakukan kajian forensik terhadap komentar bernada kebencian di akun Instagram milik selebritas Willie Salim. Dengan memanfaatkan analisis leksikal, semantik, dan pragmatik, studi ini mengidentifikasi bentuk ujaran seperti ejekan, penistaan agama, dan penyebaran informasi palsu. Hasil penelitian menyoroti bahwa *Instagram* merupakan salah satu media utama dalam penyebaran ujaran negatif, yang sulit diawasi secara manual sehingga diperlukan pendekatan deteksi otomatis berbasis pemrosesan bahasa alami (NLP) [16].

Di samping persoalan ujaran kasar dan *cyberbullying*, keberagaman dalam penggunaan bahasa seperti kekeliruan dalam tata bahasa maupun perbedaan dialek menandakan tingginya kompleksitas teks dalam komentar di *Instagram*. Penelitian psikolinguistik terhadap kesalahan berbahasa pada komentar *Instagram* (Agustus 2022) menunjukkan bahwa hambatan utama terletak pada aspek

sintaksis dan semantik, yang turut menjadi tantangan dalam proses otomatisasi klasifikasi komentar bermuatan kasar [17].

2.5. Kosakata di Dunia Digital

Pada era digital saat ini, media sosial seperti Instagram, menjadi wadah utama bagi masyarakat untuk berkomunikasi secara terbuka. Di dalamnya, ditemukan beragam bentuk ungkapan, mulai dari yang sopan hingga yang bernada kasar. Kosakata kasar umumnya berisi hinaan, ejekan, atau ekspresi negatif yang mencerminkan pelanggaran terhadap norma kesantunan dalam berbahasa. Sebaliknya, kosakata tidak kasar mengedepankan penggunaan bahasa yang netral dan santun demi menciptakan interaksi yang harmonis di ruang digital [18].

Selain muncul secara eksplisit, kosakata kasar di media sosial juga sering disampaikan melalui bentuk tidak langsung seperti sarkasme, yaitu ungkapan yang menyiratkan hinaan atau ejekan. Penelitian yang dilakukan oleh Lingua Franca (2022) pada platform Instagram menunjukkan bahwa sarkasme, khususnya dalam bentuk body shaming, digunakan tidak hanya sebagai bentuk sindiran, tetapi juga sebagai sarana untuk mengekspresikan pendapat sosial. Berdasarkan temuan tersebut, sekitar 50% hingga 51% komentar bersifat sarkastik ditemukan dalam jenis konten tertentu, dengan mayoritas berasal dari pengguna perempuan [19]

Di samping ujaran kasarnya netizen, penggunaan kosakata kasar di media digital juga kerap dilakukan secara terencana sebagai bagian dari strategi komunikasi yang bertujuan menarik perhatian audiens. Penelitian yang dilakukan

oleh Hidayatullah dan Rohimi (2023) terhadap kanal YouTube "Tekotok" menunjukkan bahwa ujaran kasar sering digunakan sebagai perangkat retoris untuk menyampaikan humor maupun kritik sosial. Meskipun pendekatan ini dinilai efektif dalam membangun daya tarik konten, namun berisiko mendorong normalisasi penggunaan bahasa ofensif dan memicu peningkatan perilaku agresif, terutama di kalangan pengguna remaja [20].

2.6. Dataset Bahasa Indonesia di Media Sosial

Akses terhadap dataset komentar *Instagram* berbahasa Indonesia masih sangat terbatas. Banyak penelitian lebih memilih menggunakan data dari *Twitter* karena API-nya lebih mudah diakses secara publik. Kondisi ini menimbulkan celah dalam kajian ilmiah, khususnya untuk platform *Instagram* yang memiliki karakteristik kebahasaan tersendiri. Menurut penelitian Pamungkas et al. (2023) dalam *Hate Speech Detection in Bahasa Indonesia*, sebagian besar dataset yang digunakan dalam studi berasal dari *Twitter* dan *Facebook*, sementara kontribusi dari *Instagram* masih sangat minim [21].

Bayu Wicaksono dan Nuri Cahyono (2024) menyusun sebuah dataset yang berisi sentimen komentar pada unggahan *Instagram* mengenai Program Kampus Merdeka. Dataset yang terdiri dari 1.694 entri tersebut dianalisis menggunakan algoritma *Naivee Bayes* dan *Decision Tree*. Hasil pengujian menunjukkan tingkat akurasi antara 84 hingga 86%, yang mengindikasikan bahwa dataset berukuran sedang ini cukup efektif untuk digunakan dalam pelatihan model klasifikasi teks kasar maupun analisis sentiment [22].

Musriana dkk (2023) melakukan pengumpulan data komentar dari akun *Instagram* @riaricis1795 yang memuat unsur ejekan, hinaan, serta provokasi. Kumpulan komentar ini dimanfaatkan sebagai bahan kajian dalam menganalisis fenomena ujaran kebencian, dan berpotensi digunakan lebih lanjut dalam pelabelan data untuk pengembangan model NLP yang fokus pada deteksi *hate speech* [23].

2.7. *Preprocessing* Data Teks Bahasa Indonesia

Tahapan *preprocecccing* teks merupakan langkah krusial dalam *Natural Language Processing* (NLP) yang berfungsi untuk membersihkan serta mempersiapkan data sebelum dianalisis lebih lanjut. Dalam penelitian yang dilakukan oleh Tuhpatussania (2022) berjudul Normalisasi Teks Komentar *Instagram* Masyarakat Makassar, digunakan metode *Levenshtein Distance* guna menyusun ulang ejaan dan kosakata lokal. Namun, efektivitas metode ini masih menghadapi kendala karena keragaman bentuk morfologi dalam bahasa yang digunakan [24].

Model berbahasa Indonesia seperti IndoBERT dan IndoBERTweet menunjukkan bahwa hasil klasifikasi sangat dipengaruhi oleh tahapan *preprocessing* yang diterapkan. Penelitian oleh Ulfia Khairani dan rekan-rekannya (2023) mengungkapkan bahwa mengabaikan proses *stopword removal* dan *stemming* justru menghasilkan akurasi tinggi, yakni mencapai 92,54% dalam tugas deteksi emosi pada komentar *Instagram*. Temuan ini menegaskan bahwa strategi *preprocessing* harus disesuaikan dengan karakteristik model yang digunakan agar performa tetap optimal [25].

Penghilangan kata-kata umum (*stopword*) dan penggunaan *stemming* umumnya dapat meningkatkan akurasi model, namun efektivitasnya tidak bersifat mutlak. Penelitian oleh Sherren Jessica Angelina dkk. (2021) menunjukkan bahwa kombinasi *preprocessing* lengkap berupa pembersihan data, *stemming*, dan penghapusan *stopword* memberikan hasil terbaik untuk algoritma *Naive Bayes*. Sebaliknya, pada algoritma *Decision Tree*, penghapusan *stopword* justru menyebabkan penurunan performa model [26].

2.8. *Feature Extraction* dan *Word Embedding*

Feature Extraction adalah langkah penting dalam *Natural Language Processing* (NLP) yang bertujuan mengubah teks menjadi representasi numerik agar dapat diproses oleh algoritma *machine learning*. Metode klasik seperti TF-IDF dan *Bag-of-Words* masih sering diterapkan karena kemudahan implementasinya dan kemampuannya menangkap informasi frekuensi kata. Meskipun begitu, kedua pendekatan ini belum mampu menggambarkan makna atau konteks kata secara mendalam [27].

Penggabungan teknik *word embedding* dengan algoritma berbasis statistik atau model pohon keputusan seperti *Random Forest* mampu memberikan keunggulan ganda: representasi kata yang memuat makna semantik sekaligus struktur data yang sistematis. Hal ini diterapkan oleh Fahriza Ichsani Rafif dan rekan-rekannya (2023) dalam penelitiannya, di mana mereka memanfaatkan fitur Word2Vec untuk membantu model *Random Forest* dalam menganalisis sentimen dari ulasan film berbahasa Indonesia. Hasil eksperimen menunjukkan bahwa pendekatan ini efektif dan menghasilkan tingkat akurasi yang tinggi, menguatkan

peran penting *embedding* dalam meningkatkan performa sistem klasifikasi pada teks bermuatan sentimen atau ujaran kasar di media sosial [28].

2.9. *Phyton*

Python adalah bahasa pemrograman tingkat tinggi yang sangat sesuai untuk pengembangan sistem berbasis *machine learning* dan pengolahan teks, karena memiliki sintaks yang sederhana serta didukung oleh berbagai pustaka seperti *scikit-learn*, *NLTK*, *pandas*, dan *matplotlib*. Dalam penelitian ini, *Python* dimanfaatkan secara menyeluruh, mulai dari proses pengambilan data menggunakan teknik *scraping*, *pra-pemrosesan* seperti tokenisasi dan *stemming*, hingga tahap pelatihan model *Random Forest* serta evaluasi kinerjanya. Kemampuan *Python* untuk menggabungkan seluruh alur kerja mulai dari analisis data hingga visualisasi dalam satu lingkungan pemrograman menjadikannya efisien dan mudah direproduksi. Penggunaan *Python* dalam penelitian analisis data dan pembelajaran mesin juga telah banyak diterapkan oleh kalangan akademik di Indonesia, salah satunya ditunjukkan dalam jurnal berjudul *Perbandingan Translation Library pada Python (Studi Kasus: Analisis Sentimen Penyakit Menular di Indonesia)* oleh Ratna Sari dan tim [29].

2.10. *Instaloader*

Instaloader merupakan sebuah pustaka dalam bahasa pemrograman *Python* yang digunakan untuk melakukan *web scraping* data dari *Instagram*, seperti unggahan, *caption*, komentar, hingga metadata dari akun yang bersifat publik. Dengan struktur sintaks yang sederhana, *Instaloader* memberikan kemudahan bagi peneliti untuk melakukan proses *login*, mengakses, serta menyimpan data

dalam format terorganisir seperti JSON atau file media lainnya. Penggunaan *Instaloader* sangat mendukung penelitian yang berkaitan dengan klasifikasi komentar kasar, karena mampu mengumpulkan data dalam skala besar secara cepat dan akurat. Selain itu, pustaka ini juga dapat diintegrasikan dengan tahapan pra-pemrosesan data di *Python*, sehingga mempermudah penyiapan data sebelum dianalisis lebih lanjut. Di Indonesia, pemanfaatan Instaloader dalam penelitian juga mulai meningkat, salah satunya terlihat dalam karya ilmiah berjudul *Implementasi Web Scraping pada Media Sosial Instagram* yang ditulis oleh Rio Baskara dan Fayruz Rahma [30].

2.11. Sastrawi

Sastrawi merupakan *library Python* yang secara khusus dikembangkan untuk melakukan proses *stemming* pada teks berbahasa Indonesia, yaitu proses mengubah kata yang memiliki imbuhan menjadi bentuk dasarnya, misalnya kata “menghina” diubah menjadi “hina”. Pustaka ini menggunakan algoritma Nazief-Adriani yang terkenal efektif dalam menangani struktur morfologi bahasa Indonesia, sehingga sangat bermanfaat dalam tahap awal pemrosesan teks seperti tokenisasi dan penghapusan kata-kata umum (*stopwords*). Sastrawi dikenal efisien dan mudah diterapkan dalam berbagai alur kerja *machine learning* berbasis *Python,* terutama pada proyek klasifikasi teks. Karena kemampuannya tersebut, Sastrawi telah digunakan secara luas dalam berbagai penelitian akademik, salah satunya dalam studi berjudul *Perbandingan Algoritma Stemming Porter, Sastrawi, Idris, dan Arifin & Setiono pada Dokumen Teks Bahasa Indonesia* [31].

2.12. *Scikit-learn*

Scikit-learn merupakan sebuah *library machine learning* berbasis *Python* yang banyak digunakan karena menyediakan berbagai algoritma pembelajaran seperti klasifikasi, regresi, dan klasterisasi, serta fitur tambahan untuk pra-pemrosesan data, validasi silang (*cross-validation*), dan penyetelan parameter secara otomatis melalui *grid search*. *Library* ini memiliki antarmuka yang konsisten dan mudah digunakan, seperti metode *fit()* untuk pelatihan dan *predict()* untuk prediksi, serta dukungan integrasi melalui sistem *pipeline*, yang memudahkan pengguna dalam membangun alur kerja analisis data secara efisien. Dalam penelitian klasifikasi kata kasar pada komentar *Instagram*, *Scikit-learn* berperan penting dalam mengubah data teks ke dalam bentuk numerik menggunakan metode TF-IDF, melatih model *Random Forest*, dan mengevaluasi performa model dengan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Karena kemudahan penggunaannya, dokumentasi yang lengkap, serta komunitas yang aktif, *Scikit-learn* menjadi salah satu pustaka paling direkomendasikan dalam pengembangan sistem machine learning berbasis *Python* [32].

2.13. *Pandas* dan *Numpy*

Pandas dan *NumPy* merupakan dua *library* penting dalam pemrograman *Python* yang kerap digunakan secara bersamaan dalam proses analisis data. *Pandas* menyediakan struktur data seperti *DataFrame* dan *Series* yang sangat membantu dalam mengelola dan membaca data dalam bentuk tabel, sedangkan *NumPy* mendukung pengolahan array multidimensi serta operasi numerik yang cepat dan efisien, yang menjadi fondasi berbagai metode analisis. Sinergi antara

keduanya sangat bermanfaat dalam tahap pembersihan data, perhitungan statistik seperti nilai rata-rata, median, dan deviasi standar, serta konversi data ke format yang sesuai sebelum digunakan dalam pemodelan. Selain itu, kemampuan keduanya dalam menangani data berskala kecil hingga menengah secara efisien menjadikan *Pandas* dan *NumPy* sebagai pilihan utama dalam pengolahan data untuk penelitian berbasis *machine learning* maupun eksplorasi data secara umum [33].

2.14. *Matplotlib* dan *Seaborn*

Matplotlib dan *Seaborn* merupakan dua pustaka visualisasi dalam bahasa *Python* yang kerap digunakan bersama dalam analisis data statistik dan proyek *machine learning*. *Matplotlib* menyediakan berbagai fitur dasar untuk pembuatan grafik seperti garis, batang, dan sebar (*scatter*), serta mampu menghasilkan visualisasi dua dimensi yang dapat diterapkan di platform seperti *Jupyter Notebook* maupun antarmuka GUI. Di sisi lain, *Seaborn* hadir sebagai pelengkap dengan antarmuka yang lebih sederhana dan tampilan grafis yang lebih menarik, terutama dalam menampilkan visualisasi statistik seperti *heatmap*, *boxplot*, dan distribusi data, serta mendukung integrasi langsung dengan struktur *DataFrame* dari *Pandas*. Kolaborasi keduanya memungkinkan peneliti menampilkan hasil analisis secara visual dengan cara yang lebih menarik, mudah dibaca, dan informatif, seperti dalam penggambaran penyebaran kosakata kasar maupun pengukuran performa model klasifikasi [34] .

2.15. Penelitian Terkait

Berikut adalah penelitian sebelumnya yang dijadikan acuan dan referensi dalam menyelesaikan masalah ini.

Tabel 2. 1 Penelitian Sebelumnya

No	Nama	Tahun	Judul	Hasil
1	Nona Sintia Rizki Hayati, Ahmad Sirulhaq, Mahsun	2025	Ujaran Kebencian pada Akun *Instagram* Willie Salim : Suatu Kajian Linguistik Forensik	Kesimpulan dari jurnal *"Ujaran Kebencian pada Akun Instagram Willie Salim: Suatu Kajian Linguistik Forensik"* adalah bahwa media sosial, khususnya *Instagram*, menjadi wadah subur bagi munculnya ujaran kebencian yang mencakup berbagai bentuk seperti penghinaan, pencemaran nama baik, penistaan, dan penyebaran berita bohong. Melalui pendekatan linguistik forensik dan teori tindak tutur, penelitian ini berhasil mengidentifikasi dan mengkategorikan ujaran-ujaran kebencian yang ditujukan kepada Willie Salim dalam komentar warganet. Temuan menunjukkan bahwa ujaran tersebut disampaikan secara eksplisit maupun implisit dengan muatan diskriminatif dan provokatif, yang berpotensi melanggar hukum sesuai KUHP. Penelitian ini menegaskan pentingnya peran linguistik forensik

				dalam mendukung penegakan hukum digital serta memberikan edukasi kepada masyarakat agar lebih bijak dan etis dalam berkomunikasi di ruang publik digital.
2.	Tia Siti Nurazizah	2025	*Articial Intelegent* dan *Machine Learning* pada Kehidupan Manusia	Kesimpulan dari jurnal *"Artificial Intelligence dan Machine Learning dalam Kehidupan Manusia"* adalah bahwa perkembangan teknologi AI dan ML telah membawa transformasi besar dalam berbagai aspek kehidupan manusia, mulai dari hiburan, pendidikan, kesehatan, bisnis, hingga transportasi. Teknologi ini memberikan dampak positif berupa peningkatan efisiensi, pengalaman pengguna yang personal, serta kemampuan analisis data yang lebih akurat untuk pengambilan keputusan. Namun, di balik manfaat tersebut, terdapat tantangan serius seperti ancaman terhadap privasi, hilangnya pekerjaan akibat otomatisasi, bias algoritma, dan ketergantungan yang berlebihan terhadap teknologi. Oleh karena itu, pemanfaatan AI dan ML harus dilakukan secara etis dan bertanggung jawab, dengan dukungan regulasi yang ketat serta kebijakan berkelanjutan agar

				manfaatnya dapat dirasakan tanpa mengorbankan nilai-nilai sosial dan kemanusiaan
3.	Bayu Wicaksono, Nuri Cahyono	2023	Analisis Sentimen Komentar *Instagram* Pada Program Kampus Merdeka Dengan *Algoritma Naïve Bayes* dan *Decision Tree*	Kesimpulan dari jurnal *"Analisis Sentimen Komentar Instagram pada Program Kampus Merdeka dengan Algoritma Naive Bayes dan Decision Tree"* adalah bahwa penerapan algoritma *Complement Naïve Bayes* dan *Decision Tree* pada data komentar *Instagram* mampu mengklasifikasikan sentimen pengguna secara efektif, meskipun terdapat ketidakseimbangan data (imbalance). Dengan bantuan teknik *SMOTE* untuk balancing data, akurasi model meningkat *Decision Tree* dari 84% menjadi 85% dan *Complement Naive Bayes* dari 81% menjadi 83%. Penelitian ini menunjukkan bahwa pemrosesan data yang tepat, seperti *preprocessing*, *labeling* otomatis dengan *VADER*, dan pembobotan *TF-IDF*, mampu meningkatkan performa model secara signifikan. Selain itu, *Decision Tree*

				terbukti memberikan performa yang lebih baik pada dataset yang seimbang. Temuan ini menegaskan pentingnya pemilihan algoritma yang tepat dan teknik penanganan data dalam analisis sentimen berbasis media sosial.
4.	Agung Wijoyo, Asep Yudistira Saputra, Safitri Ristiani, Sultan Rafly Sya'Ban, Mila Amalia, Randi Febriansyah	2024	Pembelajaran *Machine Learning*	Kesimpulan dari jurnal *"Pembelajaran Machine Learning"* adalah bahwa *machine learning* merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang memungkinkan komputer untuk belajar dan mengambil keputusan secara mandiri berdasarkan data tanpa perlu diprogram secara eksplisit. Algoritma *machine learning* terbagi menjadi tiga jenis utama: *supervised learning*, unsupervised learning, dan *reinforcement learning*, yang masing-masing memiliki pendekatan dan fungsi berbeda dalam pengolahan data. Proses implementasinya mencakup identifikasi masalah, pengumpulan dan persiapan data, pemilihan algoritma, hingga evaluasi dan perbaikan model. Bahasa pemrograman *Python* dipilih dalam pengembangan *machine learning* karena kemudahan sintaks,

				keberagaman *library*, dan komunitas *open source* yang luas. Dengan penerapan yang tepat, *machine learning* dapat menjadi solusi efektif dalam berbagai bidang kehidupan modern, seperti pengenalan suara, klasifikasi data, hingga rekomendasi digital.
5.	Siti Nurohanisah, Rini Astuti, Fadhil Muhammad Basysyar	2024	Deteksi Berita Palsu Menggunakan Algoritma *Random Forest*	Kesimpulan dari jurnal *"Deteksi Berita Palsu Menggunakan Algoritma Random Forest"* adalah bahwa *algoritma Random Forest* terbukti efektif dalam mengklasifikasikan berita palsu (*fake news*) dengan tingkat akurasi mencapai 87%. Penelitian ini memanfaatkan metode preprocessing data, transformasi, dan teknik *Natural Language Processing* (NLP) pada dataset berita yang diperoleh dari *kaggle.com*, untuk menyiapkan data dalam bentuk numerik yang dapat dianalisis. Evaluasi model dilakukan menggunakan *Confusion Matrix* dan *ROC Curve*, yang menunjukkan performa tinggi dengan nilai *precision*, *recall*, dan *f1-score* yang seimbang. Ciri khas dari berita palsu dan asli juga diidentifikasi melalui analisis kata kunci menggunakan *Word Cloud*.

				Hasil ini menunjukkan bahwa *Random Forest* merupakan metode yang andal dalam membedakan berita asli dan palsu, serta berpotensi besar untuk dikembangkan lebih lanjut dengan integrasi analisis sentimen, validasi dengan dataset terkini, dan penerapan antarmuka ramah pengguna guna memerangi penyebaran hoaks secara praktis di masyarakat.
6.	Heri Santoso, Raissa Amanda Putri, Sahbandi	2023	Deteksi Komentar *Cyberbullying* pada Media Sosial *Instagram* Menggunakan Algoritma *Random Forets*	Kesimpulan dari jurnal *"Deteksi Komentar Cyberbullying pada Media Sosial Instagram Menggunakan Algoritma Random Forest"* adalah bahwa penggunaan algoritma *Random Forest* terbukti efektif dalam mengklasifikasikan komentar *cyberbullying* dan *non-cyberbullying* pada platform Instagram. Dengan memanfaatkan proses *text preprocessing* yang lengkap serta *word embedding FastText* untuk representasi kata, model ini mampu mencapai akurasi terbaik sebesar 84% setelah dilakukan *hyperparameter tuning*. Penelitian ini berhasil membangun model yang tidak hanya mampu mengklasifikasikan komentar dalam dataset,

				tetapi juga dapat digunakan untuk mendeteksi komentar baru secara akurat. Meskipun hasilnya cukup baik, masih terdapat beberapa kesalahan prediksi, sehingga pengembangan lanjutan disarankan, seperti pengklasifikasian komentar ke dalam kategori yang lebih spesifik (misalnya rasisme atau seksisme) dan eksplorasi teknik embedding kata lain untuk meningkatkan performa model.
7.	Desti Lestari, Dodi Firmansyah, Ilmi Solihat	2023	Ujaran Kebencian Netizen pada Kolom Komentar di *Instagram* BEM KBM UNTIRTA Tahun 2022 (Kajian Linguistik Forensik)	Penelitian ini menemukan bahwa bentuk ujaran kebencian yang paling sering digunakan oleh netizen di kolom komentar *Instagram* BEM KBM Untirta tahun 2022 adalah berupa penghinaan dan pencemaran nama baik. Ujaran tersebut disampaikan dalam berbagai bentuk satuan bahasa, mulai dari kata, frasa, klausa, hingga kalimat, yang bermuatan makna menghina, merendahkan, dan menimbulkan ketidaknyamanan bagi penerimanya. Melalui pendekatan linguistik forensik dengan analisis semantik dan pragmatik, makna ujaran ini dipahami

				dalam konteks leksikal dan gramatikal yang bergantung pada situasi penggunaannya. Sebagian besar ujaran yang ditemukan terbukti melanggar Pasal 27 ayat (3) Undang-Undang Nomor 19 Tahun 2016 tentang Informasi dan Transaksi Elektronik. Oleh karena itu, hasil penelitian ini menegaskan pentingnya peningkatan literasi digital dan pemahaman hukum bagi masyarakat dalam berkomunikasi di media sosial.
8.	Salma Nabila, Kharisma Agustya Zahra Salsabilla, Nathania Trixie Aryanti, Vira Adhelia Andjani, Alfina Zahra Umardi, Eni Nurhayati	2023	Analisis Ujaran Kebencian dalam Kolom Komentar pada Media Sosial X, *Tik Tok*, dan *Instagram*	Hasil penelitian ini menunjukkan bahwa media sosial *TikTok* dan X merupakan platform yang paling sering menjadi tempat munculnya ujaran kebencian di kolom komentar, diikuti oleh *Instagram*, dengan bentuk ujaran yang paling umum berupa perundungan terhadap individu, terutama figur publik. X tercatat sebagai media sosial dengan tingkat penyebaran ujaran kebencian tercepat, menandakan tingginya potensi platform ini dalam menyebarluaskan konten negatif. Sebagian besar responden diketahui mengakses media sosial

				lebih dari dua jam setiap harinya, mencerminkan intensitas penggunaan yang tinggi. Oleh karena itu, penting bagi pengguna untuk memiliki kesadaran dan edukasi digital agar dapat menggunakan media sosial secara bijak dan mampu mengendalikan diri dari perilaku menyebarkan ujaran kebencian. Penelitian ini juga merekomendasikan agar studi selanjutnya mempertimbangkan faktor budaya dan lingkungan geografis demi memperoleh hasil yang lebih mendalam.
9.	Musriana, Firdaus, Istiqamah, F.Hanum	2022	Fenomena Ujaran Kebencian Warganet di Kolom Komentar Media Sosial Instagram Akun @Riaricis1795	Kesimpulan dari jurnal *"Fenomena Ujaran Kebencian Warganet di Kolom Komentar Media Sosial Instagram Akun @riaricis1795"* adalah bahwa media sosial, khususnya *Instagram*, telah menjadi ruang berkembangnya ujaran kebencian yang mencakup kategori ejekan, hinaan, dan hasutan. Penelitian ini menggunakan pendekatan deskriptif kualitatif dengan menganalisis 20 komentar negatif dari warganet yang ditujukan kepada akun @riaricis1795. Dari analisis tersebut, ditemukan bahwa ujaran kebencian paling banyak berupa ejekan,

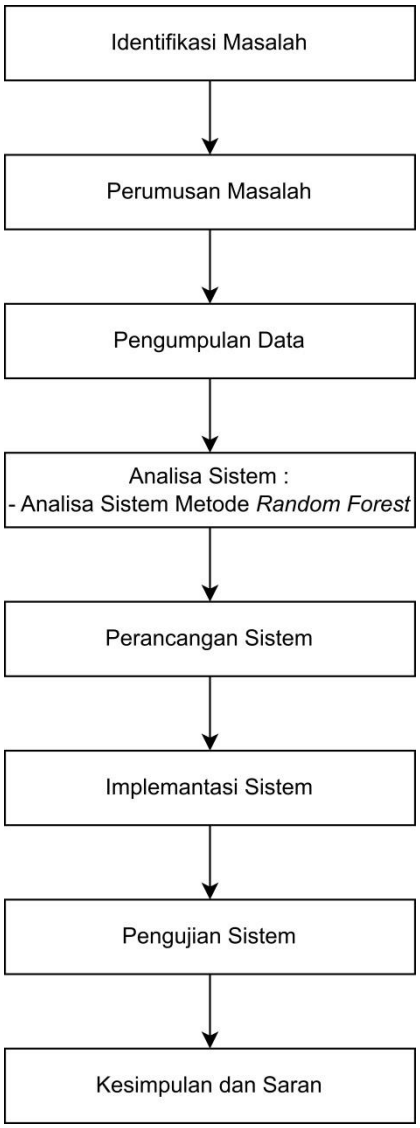
				disusul hinaan dan hasutan. Fenomena ini mencerminkan kurangnya kesadaran etika dalam berinteraksi di ruang digital dan menunjukkan bahwa ujaran kebencian bukan hanya berdampak pada individu, tetapi juga mencerminkan tantangan sosial dalam menjaga norma dan nilai komunikasi yang sehat di era digital. Penelitian ini menekankan pentingnya etika berbahasa dan literasi digital sebagai upaya pencegahan terhadap kejahatan berbahasa di media sosial.
10.	Muhammad Rafi	2020	Pengaruh *Metode Feature Extraction* Terhadap Kinerja *Random Forest* Dalam Klasifikasi Berita *Hoax* Berbahasa Indonesia	Kesimpulan dari jurnal ini adalah bahwa penggunaan metode feature extraction seperti *Word2Vec, TF-IDF*, dan *bag-of-words* secara signifikan memengaruhi kinerja model *Random Forest* dalam klasifikasi berita *hoax* berbahasa Indonesia. Penelitian menunjukkan bahwa teknik *Word2Vec*, baik skip-gram maupun *CBOW*, memberikan hasil paling optimal dengan akurasi hingga 90%, presisi 91%, recall 87%, dan f1-score 89%. Hal ini menegaskan bahwa representasi vektor yang lebih semantik dari Word2Vec dapat meningkatkan efektivitas

				model dalam mengenali berita *hoax*, dan bahwa pemilihan metode ekstraksi fitur yang tepat sangat penting dalam pengembangan sistem deteksi *hoax* berbasis *machine learning*.

BAB 3

METODOLOGI PENELITIAN

Penelitian ini dilakukan melalui serangkaian tahapan yang saling berkaitan, dengan setiap tahap dijabarkan dalam metode penelitian secara sistematis dan terstruktur, serta disajikan dalam bentuk skema yang terperinci dan mudah dipahami. Berikut tahapan-tahapan penelitian dapat dilihat pada gambar 3.1:



Gambar 3. 1 Tahapan Metodologi Penelitian

Penjelasan dari tahapan-tahapan penelitian pada Gambar 3.1 dapat dilihat pada penjelasan di bawah ini :

3.1 Identifikasi Masalah

Penelitian ini dilatarbelakangi oleh keprihatinan terhadap banyaknya komentar kasar di *Instagram*, khususnya pada akun Bapak Gibran Rakabuming Raka yang sering menjadi sasaran ujaran kebencian. Kondisi ini mencerminkan rendahnya etika bermedia sosial dan berisiko menimbulkan dampak negatif secara psikologis. Untuk itu, dikembangkanlah sistem klasifikasi otomatis yang dapat membedakan komentar kasar dan tidak kasar menggunakan metode *Random Forest* berbasis *machine learning*, yang dikenal efektif dalam mengolah data dan menghasilkan klasifikasi yang akurat. Sistem ini diharapkan mampu membantu memfilter komentar negatif dan mendorong terciptanya ruang digital yang lebih sehat.

3.2 Perumusan Masalah

Berdasarkan permasalahan yang telah didefinisikan, maka dapat dirumuskan bahwa bagaimana merancang dan membangun sebuah sistem klasifikasi kosakata kasar pada komentar *Instagram* menggunakan metode *Random Forest* berbasis *Machine Learning* secara sistematis dan objektif.

3.3 Pengumpulan Data

Data dikumpulkan dari kometar publik pada akun *Instagram* Bapak Gibran Rakabuming Raka yang memiliki tingkat interaksi tinggi. Proses pengambilan data dilakukan secara otomatis melalui teknik *web scraping* dengan

memanfaatkan bahasa pemograman *Phyton* serta pustaka seperti *BeautifulSoup* atau *Selenium.* Setelah data berhasil dikumpulkan, dilakukan tahap *preprocessing* untuk membersihkan konten dari elemen-elemen yang tidak dibutuhkan seperti emoji, simbol, dan kata-kata yang tidak relevan sebelum digunakan dalam proses pelatihan model klasifikasi.

3.4 Analisa Sistem

Tahapan selanjutnya adalah melakukan analisa metode sistem dari penelitian skripsi ini. Adapun tahapan analisa dalam penelitian ini adalah sebagai berikut :

3.4.1. Analisa Metode *Random Forest*

Penelitian ini menggunakan algoritma *Random Forest* karena kemampuannya dalam memproses data yang besar dan kompleks, seperti teks pada komentar *Instagram*. Sebagai salah satu metode *ensemble learning*, *Random Forest* membangun sejumlah pohon keputusan dan menggabungkan hasilnya melalui proses voting, yang dapat meningkatkan akurasi prediksi serta mengurangi kemungkinan *overfitting*.

Salah satu keunggulan *Random Forest* adalah kemampuannya dalam mengelola banyak fitur teks, seperti kata atau frasa dari hasil ekstraksi (misalnya *TF-IDF* atau *word embedding*), sekaligus mengidentifikasi fitur mana yang paling berpengaruh dalam proses klasifikasi. Dengan begitu, model ini tidak hanya dapat melakukan klasifikasi, tetapi juga membantu mengenali kata-kata kasar yang paling sering digunakan.

3.5 Perancangan Sistem

Perancangan sistem ini terdiri dari beberapa tahapan, dimulai dengan pengambilan data berupa komentar publik dari *Instagram* menggunakan metode *web scraping*. Data yang dikumpulkan kemudian diproses melalui tahap *preprocessing*, yang mencakup pembersihan teks, penghapusan simbol, emoji, serta kata-kata yang tidak memiliki makna penting (*stopwords*). Setelah itu, data teks dikonversi ke dalam bentuk numerik menggunakan teknik TF-IDF. Data yang telah siap kemudian dibagi menjadi dua bagian, yaitu data latih dan data uji, untuk melatih model klasifikasi menggunakan algoritma *Random Forest*. Model yang dihasilkan digunakan untuk mendeteksi komentar yang mengandung kosakata kasar dan dievaluasi berdasarkan metrik performa seperti akurasi, presisi, *recall*, dan *F1-score*.

3.6 Implementasi Sistem

Beberapa komponen pendukung yang berperan penting dalam implementasi sistem diantaranya adalah perangkat keras (*hardware*) dan perangkat lunak (*software*). Adapun spesifikasi dari perangkat keras dan perangkat lunak yang digunakan sebagai berikut:

1. Perangkat keras (*hardware*), antara lain:

   Processor : *12th Gen Intel(R) Core(TM) i5-1235U 1.30 GHz*

   Memory : 16,0 GB

   System type : *64-bit operating system, x64-based processor*

   Harddisk : 500 GB

2. Perangkat lunak (*software*), antara lain:

Sistem operasi : *Windows 11*

3.7 Pengujian Sistem

Pengujian sistem dilakukan dengan membagi data komentar *Instagram* yang telah diproses meliputi pembersihan teks, tokenisasi, dan stemming menjadi dua bagian, yaitu data latih dan data uji dengan perbandingan 80:20. Digunakannya rasio perbandingan 80:20 yaitu karena dianggap sebagai rasio seimbang dan efesien untuk membangun dan menguji model *machine learning,* selain itu rasio 80:20 digunakan untuk mencegah *overfitiing* karena jika terlalu banyak data yang digunakan untuk pelatihan misalnya 90:95%, data uji menjadi terlalu sedikit dan hasil evaluasi tidak stabil. Sebaliknya, jika terlalu sedikit data pelatihan misalnya 60% model jadi kurang belajar dan akurasi menurun. Selanjutnya data teks yang telah dibersihkan kemudian dikonversi menjadi bentuk numerik menggunakan metode TF-IDF agar dapat dianalisis oleh algoritma *machine learning*. Model *Random Forest* kemudian dilatih menggunakan data latih, dan selanjutnya diuji menggunakan data uji. Untuk mengevaluasi kemampuan model dalam membedakan komentar kasar dan tidak kasar, digunakan beberapa metrik seperti akurasi, presisi, *recall*, *F1-score*, serta *confusion matrix*. Langkah-langkah dalam evaluasi model diawali dengan menyusun *confusion matrix*, yaitu tabel yang menunjukkan perbandingan antara hasil prediksi model dan data sebenarnya, untuk mengetahui jumlah prediksi yang benar (*True Positive* dan *True Negative*) maupun prediksi yang keliru (*False Positive* dan *False Negative*). Berdasarkan tabel ini, dapat dihitung akurasi, yaitu persentase prediksi yang tepat dari seluruh data yang diuji, untuk melihat performa umum model. Setelah itu,

dihitung presisi, yaitu tingkat ketepatan model dalam mengklasifikasikan komentar sebagai kasar, dengan membandingkan jumlah komentar kasar yang terprediksi benar (TP) terhadap seluruh komentar yang diprediksi kasar (TP + FP). Kemudian dihitung *recall*, yang menunjukkan seberapa banyak komentar kasar yang berhasil dikenali oleh model, yaitu TP dibandingkan dengan seluruh komentar kasar aktual (TP + FN). Terakhir, dihitung *F1-score*, yaitu nilai rata-rata harmonis antara presisi dan *recall*, yang berguna untuk menilai keseimbangan antara ketepatan model dan kemampuannya mendeteksi komentar kasar. Pengujian sistem ini menggunakan aplikasi pendukung meliputi *Google Colab* sebagai platform utama untuk pengujian model, serta *library Python* seperti *scikit-learn*, *pandas*, dan sastrawi untuk proses klasifikasi dan analisis data teks. Selain itu, *Excel* digunakan untuk membantu tahap labeling data, dan *matplotlib* serta *seaborn* digunakan untuk memvisualisasikan hasil evaluasi model.

3.8 Kesimpulan dan Saran

Tahapan terakhir adalah menarik kesimpulan berdasarkan hasil evaluasi model. Sebagai saran, penelitian ini dapat disempurnakan dengan memperluas cakupan data dari berbagai sumber serta melakukan perbandingan kinerja dengan algoritma lain untuk memperoleh hasil yang lebih maksima