BAB III

METODOLOGI PENELITIAN

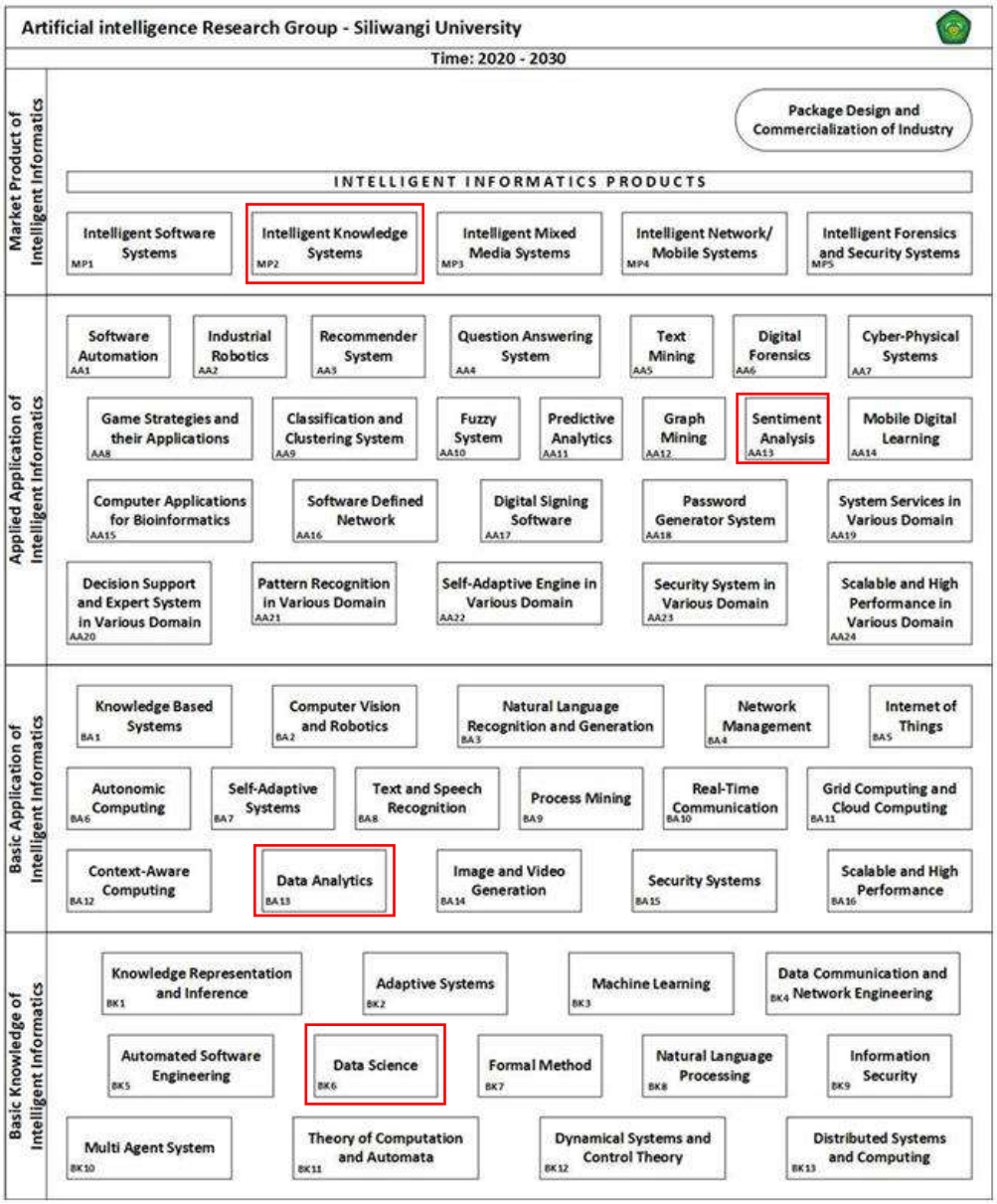
3.1 Roadmap Penelitian

Penelitian yang dilakukan merupakan bagian dari Kelompok Keahlian Informatika dan Sistem *Inteligen* (KK ISI) Universitas Siliwangi. *Roadmap* penelitian KK ISI Universitas Siliwangi yang berkolaborasi dengan *AI Research Group Siliwangi University* sebagai arah penelitian di bidang kecerdasan buatan (AI), spesifikasi dari setiap lapisan yang ditandai dengan kotak merah pada Gambar 3.1 akan digunakan dalam penelitian ini. Penelitian ini secara khusus berkaitan dengan bidang *data science*, *data analytics*, *sentiment analysis* untuk mencapai tujuan akhir berupa *intelligent knowledge system*.

Dalam konteks penelitian ini, *data science* memainkan peran penting dalam pengumpulan, pengolahan, dan pengelolaan data dari media sosial terkait sentimen publik terhadap DPR RI. Dengan pendekatan berbasis data, *data science* menjadi dasar untuk memahami pola dan tren yang mendukung pengembangan model analisis sentimen yang lebih akurat. Selanjutnya, *data analytics* digunakan untuk mengekstraksi informasi penting dari data yang telah dikumpulkan. Teknik ini memungkinkan eksplorasi, pembersihan, dan persiapan data agar dapat digunakan dalam tahapan *machine learning*. Proses ini mencakup analisis statistik dan transformasi data untuk memperoleh wawasan yang lebih mendalam terkait sentimen publik.

Sebagai inti dari penelitian ini, sentiment analysis difokuskan untuk mengidentifikasi emosi, opini, dan sentimen publik terhadap DPR RI yang muncul

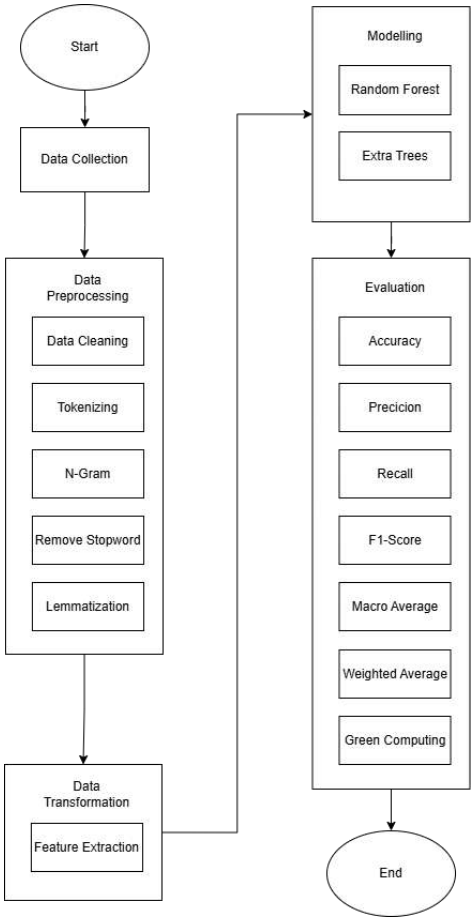
di media sosial. Penelitian ini melakukan perbandingan antara model *machine learning Extra Trees* dan *Random Forest*, dengan tujuan menemukan model yang paling akurat dalam mengklasifikasikan sentimen, sekaligus mempertimbangkan efisiensi dan akurasi sesuai dengan pendekatan komputasi hijau. Akhirnya, penelitian ini bertujuan untuk membangun *intelligent knowledge system* sebagai platform yang dapat memanfaatkan hasil analisis sentimen untuk memberikan wawasan berharga bagi pengambilan keputusan. Sistem pengetahuan cerdas ini tidak hanya menghasilkan laporan analisis sentimen tetapi juga berfungsi sebagai sarana yang efektif bagi pemangku kepentingan untuk memahami persepsi publik terhadap DPR RI dengan lebih efisien.



Gambar 3. 1 Peta Jalan Penelitian (AIS Universitas Siliwangi, 2024)

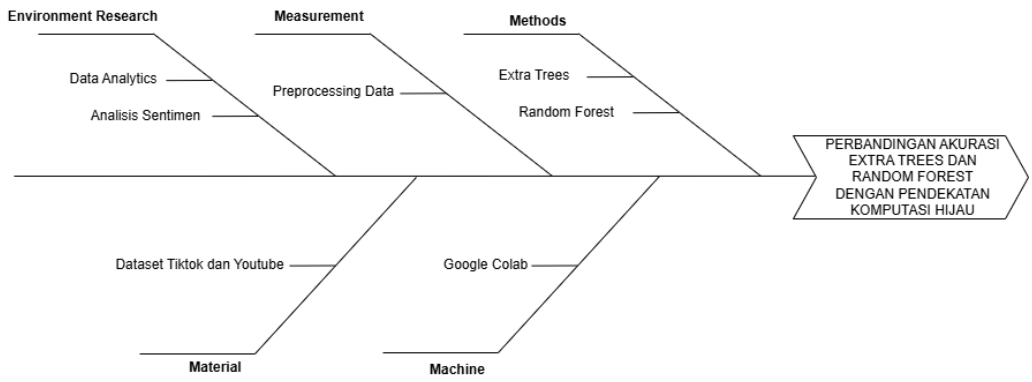
3.2 Tahapan Penelitian

Tahapan penelitian yang diterapkan dimulai dengan tahap *data collecting*, *preprocessing data, transformation*, *Modelling* dan Evaluasi. Rangkaian tahapan penelitian ini dapat diilustrasikan dalam Gambar 3.2.



Gambar 3. 2 Tahapan Penelitian

Penelitian ini juga memanfaatkan *Fishbone Diagram* sebagai alat yang sangat berguna untuk mengidentifikasi dan mengelompokkan berbagai akar penyebab potensial dari suatu masalah atau kejadian (Holifahtus Sakdiyah dkk., 2022). Dengan visualisasi faktor-faktor yang mungkin mempengaruhi masalah yang dihadapi, diagram fishbone membantu peneliti mendalami penyebab utama. Pendekatan sistematis ini tidak hanya mendukung identifikasi sumber masalah potensial tetapi juga mendukung dalam penyusunan strategi penyelesaian yang efektif (Fatmaria Tantri dkk., 2024).



Gambar 3. 3 Fishbone Diagram

Gambar 3.3 menggambarkan tujuan utama dari penelitian ini menggunakan *fishbone diagram* yang dapat dideksripsikan sebagai berikut:

a. *Environment Research,* merupakan area penelitian yang akan menjadi fokus penelitian, yaitu dalam lingkup *data analytics* khususnya pada bidang analisis sentiment.

b. *Material,* sumber data yang digunakan dalam penelitian ini adalah dataset dari platform youtube dan tiktok yang berasal dari website BRAND24 (Brand24 Global, 2024).

c. *Measurement,* merupakan tahapan persiapan data sebelum dilakukan pengujian, yang melibatkan proses *preprocessing data*.

d. *Machine,* merupakan perangkat atau alat yang digunakan dalam penelitian, yaitu menggunakan *Google Colab* sebagai *platform* komputasi.

e. *Methods,* merupakan teknik atau algoritma yang akan digunakan dalam penelitian ini, mencakup algoritma *Extra trees* dan *Random Forest*.

3.2.1 Data Collection

Data collection merupakan tahap awal dalam penelitian ini untuk mengumpulkan data yang relevan terhadap analisis sentimen yang akan dilakukan. Pengambilan data dilakukan melalui *website* Brand24, sebuah platform sosial analitik yang menyajikan informasi dari berbagai sumber digital seperti media sosial, situs berita, dan blog (Brand24 Global, 2024). Dataset yang diperoleh dari Brand24 berformat CSV, telah difilter menggunakan kata kunci "DPR RI", dan dikumpulkan dari dua platform media sosial utama, yaitu TikTok dan YouTube.

Rentang waktu pengumpulan data berlangsung mulai dari 27 April 2024 hingga 21 Mei 2024. Total jumlah data yang berhasil dikumpulkan sebanyak 2841 komentar. Dari keseluruhan data, sebanyak 1031 komentar berasal dari *platform* TikTok, sedangkan 1810 komentar berasal dari YouTube. Distribusi label sentimen dalam dataset terdiri dari 147 komentar berlabel positif, 2538 komentar netral, dan 154 komentar negatif.

3.2.2 Data Preprocessing

Data preprocessing dilakukan untuk mempersiapkan *dataset* sebelum analisis lebih lanjut.

Tahapan *preprocessing* dari penelitian ini terdiri dari:

a. *Data Cleaning*

Pada tahap awal ini, teks mentah dibersihkan dari informasi yang tidak relevan atau mengganggu, termasuk penghapusan *URL*, *HTML*, *emoji*, dan simbol yang

tidak diperlukan. Tujuannya adalah mempersiapkan teks agar lebih mudah diproses tanpa gangguan dari informasi yang tidak relevan (Davy Wiranata, 2024).

b. *Tokenizing*

Pada tahap ini, teks dipecah menjadi bagian-bagian kecil agar lebih mudah dianalisis. Setiap unit yang dihasilkan bisa berupa kata atau elemen lain yang relevan untuk pemrosesan lebih lanjut, seperti ketika sebuah kalimat diuraikan menjadi susunan kata-kata yang berdiri sendiri (Giovani dkk., 2020).

c. *N-Gram*

Dalam proses ini, kata-kata yang saling berdekatan digabungkan untuk membentuk pola tertentu yang sering muncul bersamaan. Pola tersebut bisa berupa satu kata, dua kata, atau bahkan tiga kata berturut-turut, dan sering kali membawa makna khusus dalam konteks bahasa Indonesia, terutama untuk mendeteksi kecenderungan makna atau sentimen dalam suatu teks (Nathifa Agustiana dkk., 2023)(Satya Marga dkk., 2021).

d. *Remove Stopwords*

Tahap ini bertujuan untuk menyaring elemen-elemen dalam teks yang dinilai kurang berpengaruh terhadap makna utama. Kata-kata seperti "dan", "atau", dan "sebuah" sering kali dihilangkan agar analisis dapat lebih fokus pada informasi yang benar-benar penting (Alita dkk., 2020).

e. *Lemmatization*

Dalam proses ini, bentuk kata disederhanakan menjadi bentuk dasarnya agar maknanya menjadi lebih jelas dan konsisten. Misalnya, berbagai variasi dari satu

kata seperti "mengkritik", "dikritik", hingga "kritikan" akan diperlakukan sebagai satu kesatuan makna yang seragam, yakni "kritik"(D. Tiwari & Singh, 2019).

3.2.3 Data Transformation

Transformasi data merupakan proses mengubah data yang telah dipilih ke dalam format yang dapat digunakan oleh model analisis sentimen. Proses ini melibatkan konversi teks menjadi representasi numerik atau fitur yang dapat diproses oleh algoritma. Transformasi ini sangat penting karena algoritma pembelajaran mesin dan metode statistik umumnya tidak dapat langsung mengolah data teks mentah tanpa adanya transformasi yang sesuai (T. Meisya dkk., 2021).

Tahap transformasi mencakup teknik ekstraksi fitur yang bertujuan untuk mengonversi data teks ke dalam bentuk yang lebih sesuai untuk analisis sentimen. Salah satu metode yang digunakan adalah TF-IDF (*Term Frequency - Inverse Document Frequency*), yaitu teknik yang menghitung frekuensi kemunculan suatu kata dalam sebuah dokumen (*Term Frequency*) dan mengoreksinya dengan *Inverse Document Frequency*, yang mengukur seberapa umum atau langka kata tersebut dalam keseluruhan korpus. Dengan cara ini, kata-kata umum seperti "dan" atau "dari" yang sering muncul dalam banyak dokumen tidak akan dianggap terlalu penting dalam analisis.

3.2.4 Modelling

Data hasil transformasi menggunakan TF-IDF kemudian digunakan dalam proses pembuatan model klasifikasi menggunakan dua algoritma *ensemble*, yaitu *Random Forest Classifier* dan *Extra Trees Classifier*. Model *Random Forest* dan

Extra Trees dipilih karena termasuk dalam algoritma *ensemble* berbasis pohon keputusan yang mampu menghasilkan prediksi yang stabil, akurat, dan tahan terhadap *overfitting*. Selain itu, kedua algoritma ini mendukung pemrosesan secara *paralel*, yang sangat relevan dalam konteks evaluasi efisiensi komputasi. Karakteristik ini memungkinkan penelitian untuk membandingkan performa algoritma tidak hanya dari sisi akurasi, tetapi juga dari aspek *green computing* secara komprehensif, termasuk waktu proses dan konsumsi energi.

a. Random Forest Classifier

Pada tahap ini, digunakan pendekatan berbasis sekumpulan pohon keputusan yang dibangun secara acak dari data pelatihan. Proses dimulai dengan membuat beberapa *subset* data dari keseluruhan *dataset* melalui teknik *bootstrap*, yaitu pengambilan sampel secara acak dengan pengembalian. Setiap subset digunakan untuk membentuk satu pohon keputusan yang independen.

Selain itu, pemilihan fitur yang digunakan untuk membagi *node* di setiap pohon juga dilakukan secara acak, sehingga menghasilkan keragaman antar pohon yang dibentuk. Setelah seluruh pohon selesai dibangun, proses prediksi dilakukan dengan menggabungkan hasil dari semua pohon menggunakan sistem voting mayoritas. Strategi ini membantu model menangkap pola yang kompleks dalam data tanpa terlalu bergantung pada satu bagian tertentu dari dataset. Dengan demikian, metode ini tidak hanya meningkatkan akurasi prediksi, tetapi juga lebih tahan terhadap overfitting yang sering terjadi pada pohon keputusan tunggal (H. Chyntia Morama dkk., 2022).

b. *Extra Trees Classifier*

Model ini merupakan varian dari metode *ensemble* pohon keputusan yang memiliki karakteristik unik dalam pembentukan tiap pohonnya. Berbeda dengan metode lain yang mencari pemisahan optimal berdasarkan suatu kriteria, pendekatan ini justru memperkenalkan tingkat acakan yang lebih tinggi dengan memilih fitur dan ambang batas pemisahan secara acak. Hal ini dilakukan untuk setiap *node* dalam proses pembentukan pohon. Akibatnya, model yang dihasilkan memiliki variasi struktural yang tinggi antar pohon, yang berdampak pada peningkatan kemampuan generalisasi terhadap data baru. Karena proses pemisahan tidak memerlukan pencarian nilai terbaik seperti pada *Random Forest*, waktu pelatihan menjadi lebih cepat. *Extra Trees Classifier,* yang juga dikenal sebagai *Extremely Randomized Trees*, sering kali dipilih ketika kecepatan dan efisiensi dalam pelatihan menjadi pertimbangan utama. Selain itu, karena sifat acaknya yang tinggi, model ini juga cenderung lebih *robust* terhadap *noise* dan fluktuasi kecil dalam data(N. M. Salih & M Zeebaree, 2024).

3.2.5 Evaluasi Model

Proses evaluasi dilakukan untuk mengukur kinerja model yang telah dibangun. Evaluasi ini penting untuk menentukan sejauh mana model dapat mengklasifikasikan data dengan benar dan seberapa baik model bekerja dalam kondisi berbeda. Berikut beberapa metrik dari evaluasi yang digunakan:

a. Accuracy

Pada tahap evaluasi model, salah satu metrik yang paling umum digunakan adalah akurasi. Metrik ini menghitung proporsi prediksi yang sesuai dengan

label sebenarnya dari keseluruhan data uji. Dengan membandingkan jumlah prediksi yang tepat terhadap total prediksi yang dilakukan, akurasi memberikan gambaran umum mengenai seberapa sering model menghasilkan keputusan yang benar. Nilai yang mendekati satu biasanya menunjukkan bahwa model cukup handal dalam mengenali pola-pola dalam data, khususnya dalam konteks klasifikasi sentimen. Dalam praktiknya, akurasi sering menjadi indikator awal untuk menilai kinerja model secara menyeluruh, meskipun belum tentu mencerminkan performa yang seimbang jika terdapat ketimpangan distribusi kelas dalam dataset (A. Chyntia Morama, 2022). Rumus mengukur akurasi ditunjukkan pada persamaan (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

b. *Precicion*

Presisi menjadi penting ketika fokus evaluasi diarahkan pada ketepatan model dalam mengidentifikasi suatu kelas tertentu, terutama kelas positif. Metrik ini membandingkan jumlah prediksi yang benar-benar relevan terhadap seluruh prediksi yang diklaim sebagai relevan oleh model. Dalam konteks analisis sentimen, misalnya, presisi menunjukkan seberapa akurat model dalam menyatakan bahwa sebuah teks bernuansa positif, tanpa terlalu sering salah mengenali sentimen lain sebagai positif. Nilai presisi yang tinggi mengindikasikan bahwa prediksi model terhadap kelas tersebut memiliki tingkat kepercayaan yang baik, dan tidak banyak menghasilkan kesalahan prediksi yang menyesatkan (Fikri, 2020). Rumus Precision dapat dilihat pada persamaan (2)

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

c. *Recall*

Recall lebih menyoroti kemampuan model dalam menangkap seluruh data yang seharusnya termasuk dalam kelas positif. Nilai yang tinggi menunjukkan bahwa model berhasil mengenali sebagian besar instance penting dari kelas tersebut. Metrik ini berguna saat kegagalan dalam mendeteksi suatu kelas harus ditekan sekecil mungkin (A. Salih & Zeebaree, 2024). Semakin tinggi nilai *Recall*, semakin baik model dalam mengenali kelas positif tanpa melewatkan banyak instance yang benar-benar positif. Persamaan (3) menunjukkan rumus untuk menghitung *Recall*.

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

d. *F1-Score*

F1-Score mempertimbangkan presisi dan recall secara bersamaan untuk memberikan penilaian yang lebih seimbang. Skor ini berguna ketika data memiliki distribusi kelas yang tidak merata, sehingga hanya mengandalkan salah satu metrik saja bisa memberikan gambaran yang menyesatkan. Nilai F1 yang tinggi menandakan bahwa model tidak hanya tepat, tetapi juga cakupannya baik.. Rumus *F1-Score* ditunjukkan oleh persamaan (4).

$$F1 - Score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall} \qquad (4)$$

F1-Score digunakan untuk memberikan gambaran yang lebih lengkap tentang performa model, terutama jika kita menginginkan keseimbangan antara presisi dan recall (A. Meisya, 2021).

e. *Macro average*

Metrik ini menghitung rata-rata dari nilai presisi, *recall*, dan *F1-Score* dengan memberikan perlakuan yang sama terhadap setiap kelas. Semua kelas dianggap setara tanpa memperhatikan jumlah datanya. Pendekatan ini bermanfaat ketika ingin menilai kinerja model secara menyeluruh, terutama pada data yang tidak seimbang, agar kelas yang jarang muncul tetap diperhitungkan dalam evaluasi (Wiranata, 2024).

Rumus untuk menghitung *macro average precision*, *recall*, dan *F1-score* masing-masing ditunjukkan oleh Persamaan (5), Persamaan (6), dan Persamaan (7) berikut:

$$Macro\ Precision = \frac{Precision1 + Precision2 + \cdots + Precisionn}{n} \qquad (5)$$

$$Macro\ Recall = \frac{Recall1 + Recall2 + \cdots + Recalln}{n} \qquad (6)$$

$$Macro\ F1 - Score = \frac{F1{-}Score1 + F1{-}Score2 + \cdots + F1{-}Scoren}{n} \qquad (7)$$

f. *Weighted Average*

Berbeda *dari macro, weighted average* mempertimbangkan proporsi jumlah data dari tiap kelas saat menghitung rata-rata presisi, *recall*, dan *F1-Score*. Artinya, kelas yang lebih sering muncul akan memiliki pengaruh lebih besar terhadap hasil akhir. Metrik ini cocok digunakan ketika distribusi data antar kelas tidak merata, agar penilaian kinerja model mencerminkan dominasi jumlah data yang sebenarnya (Fikri, 2020).

Perhitungan *Weighted Precision*, *Weighted Recall*, dan *Weighted F1-Score* masing-masing ditunjukkan pada Persamaan (8), Persamaan (9), dan Persamaan (10) berikut :

$$Weighted\ Precision = \frac{\sum_{i=1}^{n}(Precision_i\ x\ w_i)}{\sum_{i=1}^{n} w_i} \qquad (8)$$

$$Weighted\ Recall = \frac{\sum_{i=1}^{n}(Recall_i\ x\ w_i)}{\sum_{i=1}^{n} w_i} \qquad (9)$$

$$Weighted\ F1-Score = \frac{\sum_{i=1}^{n}(F1-Score_i\ x\ w_i)}{\sum_{i=1}^{n} w_i} \qquad (10)$$

Dimana :

n = jumlah kelas

w_i = jumlah data dalam kelas ke-i

3.2.6 Evaluasi *Green Computing*

Konsep *green computing* menekankan pentingnya efisiensi dalam penggunaan sumber daya komputasi guna mengurangi dampak negatif terhadap lingkungan. Dalam penerapan *machine learning*, evaluasi ini mencakup pengukuran aspek seperti waktu komputasi, konsumsi daya, serta estimasi emisi karbon selama proses pelatihan dan prediksi. Pendekatan ini tidak hanya mempertimbangkan akurasi model, tetapi juga dampak ekologis dari pemrosesan data yang dilakukan. Model yang mampu memberikan hasil optimal dengan kebutuhan daya yang lebih rendah dipandang lebih berkelanjutan dan ramah lingkungan (Patel, 2023). Untuk membantu proses pengukuran ini, dapat digunakan tools seperti *CodeCarbon*, yang dirancang untuk mengestimasi konsumsi energi dan jejak karbon selama eksekusi program. Dengan demikian, praktik evaluasi ini

tidak hanya mendukung efisiensi teknis, tetapi juga mendukung komputasi yang lebih bertanggung jawab terhadap lingkungan (Thompson & Gupta, 2022).

Pengukuran konsumsi energi dapat dilakukan menggunakan *CodeCarbon*, sebuah pustaka yang mengestimasi dampak lingkungan dari pemrosesan komputasi. Parameter yang diukur mencakup:

a. Total konsumsi energi

Perhitungan konsumsi energi total ditunjukkan di persamaan (11) berikut:

$$E = P \; x \; T \qquad (11)$$

dimana:

E = total energi yang digunakan (dalam kWh)

P = total daya komputasi (CPU + GPU) dalam kW

T= waktu pemrosesan dalam jam.

b. *Carbon emission estimate* (perkiraan emisi karbon yang dihasilkan)

Perhitungan estimasi emisi karbon ditunjukkan di persamaan (12) berikut :

$$C = E \; x \; EF \qquad (12)$$

dimana:

C = estimasi emisi karbon (dalam gram CO_2)

EF = faktor emisi listrik wilayah.

Data konsumsi energi ini dapat dibandingkan antara model yang berbeda, seperti Random Forest dan Extra Trees, untuk menentukan mana yang lebih hemat

energi selama proses pelatihan. Dengan pendekatan ini, pemilihan model tidak hanya berdasarkan performa prediktif tetapi juga berdasarkan efisiensi penggunaan sumber daya (Wulandari, 2024).

Sebagai acuan dalam menilai tinggi rendahnya konsumsi energi dan emisi karbon dari model *machine learning*. Tabel 3.1 menampilkan standar yang dirujuk dari berbagai studi (Lacoste dkk., 2019; Lannelongue dkk., 2020; Strubell dkk., 2019):

Tabel 3. 1 Standar Konsumsi Energi dan Emisi Karbon

Kategori	**Rentang Konsumsi Energi (kWh)**	**Rentang Emisi (CO_2eq)**
Rendah	< 0.1	< 100 gram
Sedang	0.1 – 10	100g – 10 kg
Tinggi	10 – 1000	10 – 500 kg
Sangat Tinggi	> 1000	> 500 kg

3.3 Justifikasi Pemilihan Metode

Pemilihan *Random Forest* dan *Extra Trees* didasarkan pada :

a. Kemampuan mereka dalam menangani data tidak terstruktur dan berdimensi tinggi.

   Dalam analisis data berbasis teks, jumlah fitur yang dihasilkan, seperti melalui transformasi TF-IDF, dapat mencapai ribuan. *Random Forest* dan *Extra Trees* memiliki kemampuan alami dalam menangani data berdimensi tinggi tanpa memerlukan tahapan seleksi fitur yang kompleks, sehingga efektif untuk data media sosial yang tidak terstruktur (Chen & Guestrin, 2021).

b. Kemampuan komputasi *paralel* yang efisien.

   Struktur ensemble dari kedua algoritma ini mendukung pembangunan pohon keputusan secara paralel, yang mempercepat waktu pelatihan dan meningkatkan efisiensi sumber daya, suatu keunggulan penting dalam pengolahan data berskala besar (Arrieta dkk., 2020).

c. Fleksibilitas dalam konfigurasi dan interpretabilitas hasil

   Random Forest dan *Extra Trees* memungkinkan pengaturan fleksibel terhadap jumlah pohon, kedalaman pohon, dan *subset* fitur, serta memberikan interpretasi hasil melalui analisis pentingnya fitur (*feature importance*), yang meningkatkan transparansi model (L'Heureux dkk., 2020).

d. Potensi penerapan dalam sistem analisis yang memperhatikan aspek keberlanjutan energi (*low-carbon AI)*

   Dibandingkan dengan model *deep learning* yang intensif komputasi, algoritma ini menawarkan keseimbangan antara akurasi dan konsumsi energi, mendukung konsep pengembangan kecerdasan buatan berkelanjutan berbasis *low-carbon AI* (Saran, 2021).