

BAB II

KAJIAN PUSTAKA

A. Kajian Teoritis

1. *Natural Language Processing (NLP)*

Menurut Rivaldi & Wismarini (2024), *Natural Language Processing* digunakan untuk mengolah data teks melalui beberapa tahapan, seperti *tokenization*, *stemming*, dan analisis leksikal sehingga data yang awalnya berupa teks tidak terstruktur dapat diproses menjadi informasi yang berguna dalam proses analisis sentimen.

Penelitian yang telah dilaksanakan oleh Ertansyah et al. (2025) menerangkan bahwa NLP berperan penting dalam proses analisis sentimen karena mampu membantu komputer mengidentifikasi dan mengelompokkan opini masyarakat berdasarkan data teks dari media sosial. Melalui tahapan *preprocessing*, data teks dapat dipersiapkan sebelum dilakukan proses klasifikasi sehingga menghasilkan informasi yang lebih akurat.

Penelitian lain yang telah dilaksanakan oleh Tribuana et al. (2025) membuktikan bahwa metode Natural Language Processing (NLP) sangat efektif digunakan untuk memproses bahasa informal dan tidak baku di media sosial. Penggunaan NLP yang dikombinasikan dengan algoritma pembelajaran mesin terbukti mampu mengklasifikasikan opini publik secara otomatis dan dapat digunakan sebagai sistem pemantauan sentimen masyarakat secara *real-time*.

Dari pemaparan sebelumnya, dapat ditarik kesimpulan bahwa NLP berfungsi mentransformasi bahasa alami manusia agar dapat dikenali dan diproses oleh sistem komputer. Peran NLP dalam studi ini difokuskan pada tahap awal persiapan data komentar TikTok, guna mendukung kelancaran proses pengelompokan sentimen yang mengandalkan model *Bidirectional Long Short-Term Memory* (BiLSTM).

2. *Deep Learning*

Sebagai sub-bidang dari *Machine Learning*, *Deep Learning* beroperasi dengan mengandalkan jaringan saraf tiruan yang tersusun atas berbagai *hidden layer* guna mengekstraksi relasi data tanpa intervensi manusia. Keunggulan esensialnya dibandingkan algoritma *Machine Learning* klasik terletak pada kapabilitasnya melakukan ekstraksi fitur secara otonom, tanpa perlu didefinisikan secara manual. Karakteristik ini menjadikan pendekatan *Deep Learning* sangat tangguh ketika dihadapkan pada pemrosesan data yang rumit dan tidak terstruktur, seperti halnya klasifikasi teks, pengolahan gambar, hingga analisis suara (Pardede, 2023).

Menurut Pardede (2023), perkembangan *Deep Learning* memungkinkan sistem komputer mempelajari pola data dalam jumlah besar melalui proses pelatihan yang dilakukan secara bertahap pada setiap lapisan jaringan saraf. Kemampuan tersebut menjadikan *Deep Learning* banyak diterapkan pada berbagai bidang, seperti pengenalan citra, pengenalan suara, Natural Language Processing (NLP), dan analisis data berskala besar.

Penelitian yang sudah dijalankan oleh Naufal & Kusuma (2023) menjelaskan bahwa metode berbasis *Deep Learning* punya kemampuan yang baik untuk mempelajari karakteristik data yang kompleks sehingga mampu mendapatkan performa klasifikasi yang tinggi pada berbagai kasus. Selain itu, penelitian Qososyi et al. (2025) juga menunjukkan bahwa penerapan *Deep Learning* menggunakan arsitektur *Bidirectional Long Short-Term Memory* (BiLSTM) mampu memberikan hasil yang baik dalam proses klasifikasi sentimen pada data teks berbahasa Indonesia.

Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa *Deep Learning* merupakan pengembangan dari *Machine Learning* yang memanfaatkan jaringan saraf tiruan berlapis untuk mempelajari pola data secara otomatis tanpa membutuhkan ekstraksi fitur secara manual. Kemampuan tersebut menjadikan *Deep Learning* banyak digunakan dalam berbagai bidang, termasuk analisis sentimen. Pada penelitian ini, metode *Deep Learning* yang digunakan adalah arsitektur *Bidirectional Long Short-Term Memory* (BiLSTM) untuk mengelompokkan komentar TikTok mengenai Program Makan Bergizi Gratis ke dalam kelompok positif, netral, dan negatif.

3. *Bidirectional Long Short-Term Memory* (BiLSTM)

Sebagai perluasan dari metode *Long Short-Term Memory* (LSTM), *Bidirectional Long Short-Term Memory* (BiLSTM) dirancang untuk memproses aliran data sekuensial secara bolak-balik (*forward* dan *backward*). Cara kerja ini memberikan keuntungan bagi model untuk

menangkap makna dari kata sebelum dan sesudahnya secara bersamaan. Alhasil, representasi konteks kalimat yang dibentuk menjadi jauh lebih utuh jika dibandingkan dengan pendekatan LSTM satu arah (Graves et al., 2005).

Menurut Zahwa et al. (2023), BiLSTM banyak diterapkan pada analisis sentimen karena memiliki kemampuan untuk mengerti secara kontekstual antar kata dalam suatu kalimat. Kemampuan tersebut menjadikan BiLSTM lebih efektif dalam mengolah data teks yang memiliki struktur kalimat beragam, seperti komentar pada media sosial, sehingga mampu meningkatkan kualitas proses klasifikasi sentimen.

Penelitian lain yang telah dijalankan oleh Qososyi et al. (2025) membuktikan bahwa penerapan BiLSTM pada data teks berbahasa Indonesia menghasilkan performa klasifikasi sentimen yang sangat baik. Arsitektur BiLSTM dipilih karena kemampuannya yang unggul dalam menangkap hubungan temporal dan konteks yang sangat kompleks di dalam sebuah teks. Hal ini menjadikan BiLSTM sangat ideal untuk membedah opini dan diskusi publik di media sosial yang dinamis.

Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa *Bidirectional Long Short-Term Memory* (BiLSTM) ialah pengembangan dari LSTM yang memanfaatkan proses pembelajaran dari dua arah untuk memahami makna kalimat secara utuh. Kemampuan tersebut menjadikan BiLSTM banyak digunakan dalam berbagai penelitian *Natural Language Processing*, khususnya analisis sentimen. Oleh karena itu, penelitian ini menggunakan BiLSTM sebagai metode utama dalam mengklasifikasikan

komentar TikTok mengenai Program Makan Bergizi Gratis ke dalam kategori positif, netral, dan negatif.

4. **Web Scraping**

Menurut Kusumo & Somya (2022), *web scraping* merupakan proses pengambilan dokumen semi-terstruktur dari internet yang umumnya berbentuk HTML atau XHTML untuk memperoleh data atau informasi tertentu sesuai dengan kebutuhan pengguna. Data yang berhasil diekstraksi kemudian dapat disimpan dan dimanfaatkan sebagai sumber informasi maupun dataset dalam berbagai penelitian.

Chrisinta & Simarmata (2024), menjelaskan bahwa *web scraping* merupakan teknik ekstraksi data dari halaman web yang dilakukan secara otomatis menggunakan program komputer. Teknik ini mampu mengambil data dalam jumlah besar secara lebih cepat dan efisien dibandingkan proses pengumpulan data secara manual, sehingga banyak dimanfaatkan dalam proses *data mining*, analisis data, maupun pengembangan aplikasi.

Penelitian yang dilakukan oleh Dwicahyo & Ratnasari (2023) menyimpulkan bahwa web scraping sangat populer digunakan sebagai metode pengumpulan data pada penelitian saat ini, khususnya untuk mengotomatisasi proses ekstraksi data dari situs web tertentu. Teknik ini membantu peneliti mengumpulkan data secara lebih efektif dan efisien dari segi waktu sebelum digunakan pada tahap analisis.

Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa *web scraping* merupakan teknik pengambilan dan ekstraksi data secara

otomatis dari halaman web menggunakan program komputer untuk memperoleh informasi sesuai kebutuhan. Teknik ini mampu mengumpulkan data dalam jumlah besar secara cepat dan efisien sehingga banyak dimanfaatkan dalam berbagai penelitian yang membutuhkan data dari internet. Pada penelitian ini, *web scraping* digunakan untuk mengambil komentar pada video TikTok yang membahas tentang program makan bergizi gratis sebagai dataset yang akan dianalisis menggunakan metode *Bidirectional Long Short-Term Memory* (BiLSTM).

5. Analisis Sentimen

Dalam ranah *Natural Language Processing* (NLP), analisis sentimen bertindak sebagai instrumen komputasi untuk mendeteksi dan mengelompokkan kecenderungan opini dalam wujud teks ke dalam kelas sentimen positif, netral, atau negatif. Penerapan metode ini sangat esensial untuk memetakan respons maupun persepsi masyarakat luas terhadap suatu kebijakan publik, seperti Program Makan Bergizi Gratis (Anggriyani & Fakhriza, 2024). Berlandaskan pendekatan tersebut, penelitian ini akan memanfaatkan ketersediaan data teks tidak terstruktur dari platform jejaring sosial TikTok.

Menurut Rivaldi & Wismarini (2024), analisis sentimen adalah proses pengolahan data teks menggunakan teknik *Natural Language Processing* (NLP) untuk mengelompokkan opini pengguna secara otomatis sehingga informasi yang terkandung di dalam ulasan atau komentar dapat dimanfaatkan sebagai dasar pengambilan keputusan. Penerapan analisis

sentimen mampu membantu proses analisis terhadap data dalam jumlah besar yang sulit dilakukan secara manual.

Penelitian lain yang dilakukan oleh Ertansyah et al. (2025) menjelaskan bahwa analisis sentimen bertujuan mengelompokkan pendapat masyarakat ke dalam kelompok positif, negatif, dan netral berdasarkan data media sosial. Dengan menggunakan teknik *Natural Language Processing* (NLP), proses klasifikasi dapat dilakukan secara otomatis sehingga mampu menyediakan gambaran mengenai sudut pandang masyarakat terhadap suatu isu secara lebih cepat dan efisien.

Mengacu pada penjabaran di atas, dapat disimpulkan bahwa analisis sentimen bertindak sebagai instrumen NLP untuk mengotomatisasi pengelompokan opini menjadi tiga polaritas: positif, negatif, dan netral. Kemampuannya dalam menerjemahkan teks tidak terstruktur menjadi landasan pengambilan keputusan sangatlah esensial. Oleh sebab itu, penelitian ini menerapkan teknik tersebut untuk mendeteksi dan memilah arah pandangan masyarakat di kolom komentar TikTok mengenai pelaksanaan Program Makan Bergizi Gratis.

6. *Text Preprocessing*

Menurut Hakim (2021), *text preprocessing* adalah langkah awal dalam pengolahan data teks yang digunakan untuk mempersiapkan data sebelum diproses menggunakan algoritma *machine learning*. Pada tahap ini dilakukan berbagai proses, seperti pembersihan data (*cleaning*), normalisasi kata, penghapusan *stopword*, maupun teknik lainnya agar data yang semula

tidak terstruktur menjadi lebih teratur. Tahapan ini memiliki pengaruh terhadap kualitas data yang digunakan sehingga dapat memengaruhi hasil klasifikasi sentimen.

Penelitian yang dilakukan oleh Rifaldi et al. (2023) menjelaskan bahwa *text preprocessing* merupakan proses pengolahan awal terhadap data hasil *crawling* yang masih mengandung banyak *noise*. Proses tersebut bertujuan menghasilkan data yang lebih bersih dan siap digunakan pada tahap pembentukan model. Beberapa tahapan preprocessing yang umum dilakukan meliputi *case folding*, *tokenizing*, *stopword removal*, normalisasi kata, dan *stemming*, yang penerapannya dapat disesuaikan dengan kebutuhan penelitian dan karakteristik data.

Khairunnisa et al. (2021) menjelaskan bahwa penerapan teknik *text preprocessing* berpengaruh terhadap performa analisis sentimen. Pemilihan tahapan preprocessing yang sesuai mampu meningkatkan kualitas representasi data teks sehingga model klasifikasi dapat bekerja lebih optimal. Oleh karena itu, penentuan tahapan preprocessing perlu disesuaikan dengan karakteristik data yang digunakan agar informasi penting pada teks tetap dipertahankan selama proses pengolahan.

Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa *text preprocessing* merupakan tahapan awal yang bertujuan membersihkan, menata, dan menstandarkan data teks agar siap digunakan dalam proses pembelajaran mesin maupun *deep learning*. Pada penelitian ini, *text preprocessing* dilakukan melalui tahapan *cleaning*, *case folding*,

normalisasi kata, dan *stopword removal* sehingga komentar TikTok memiliki format yang lebih konsisten sebelum diproses menggunakan *Tokenizer* dan model Bidirectional Long Short-Term Memory (BiLSTM).

7. Tokenisasi (*Tokenization*)

Menurut Susilawati et al. (2025), tokenisasi merupakan proses mengubah data teks menjadi unit-unit kata (*token*) yang selanjutnya direpresentasikan dalam bentuk indeks numerik agar dapat diproses oleh model *deep learning*. Pada implementasinya, proses tokenisasi dilakukan menggunakan *Tokenizer* dari Keras yang membangun kamus kata berdasarkan frekuensi kemunculan setiap kata pada korpus, kemudian mengonversi setiap kalimat menjadi urutan bilangan bulat (*integer Sequence*).

Penelitian yang dilakukan oleh Ramadhan & Kusnawi (2026), menjelaskan bahwa tokenisasi merupakan salah satu tahapan penting dalam *Natural Language Processing* karena bertujuan memecah teks menjadi unit-unit yang lebih kecil sehingga lebih mudah diproses oleh komputer. Hasil tokenisasi selanjutnya diubah menjadi representasi numerik melalui proses *encoding* dan digunakan sebagai masukan bagi model *Bidirectional Long Short-Term Memory* (BiLSTM). Selain itu, tokenisasi membantu sistem mengenali pola hubungan antar kata sehingga proses klasifikasi dapat dilakukan dengan lebih efektif.

Perdana & Pramudwiatmoko (2024) menyatakan bahwa tokenisasi merupakan salah satu tahapan *preprocessing* yang umum diterapkan pada

analisis sentimen. Tahapan ini dilakukan setelah proses pembersihan data untuk memecah kalimat menjadi token sehingga setiap kata dapat diproses pada tahap berikutnya. Tokenisasi menjadi dasar dalam pembentukan representasi data teks sebelum dilakukan proses *padding* maupun pelatihan model klasifikasi.

Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa tokenisasi merupakan proses mengubah data teks menjadi unit-unit kata yang kemudian direpresentasikan dalam bentuk numerik agar dapat diproses oleh model *deep learning*. Pada penelitian ini, proses tokenisasi dilakukan menggunakan *Tokenizer* dari TensorFlow/Keras untuk membangun kosakata (*vocabulary*) serta mengubah komentar TikTok hasil *preprocessing* menjadi urutan indeks kata sebelum dilakukan proses *padding Sequence* dan klasifikasi menggunakan metode *Bidirectional Long Short-Term Memory* (BiLSTM).

8. *Flowchart*

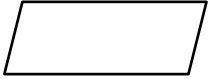
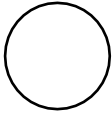
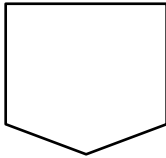
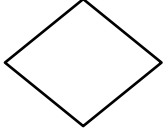
Penelitian yang dilakukan oleh Listyoningrum et al. (2023) menjelaskan bahwa *flowchart* merupakan alat bantu visual yang efektif untuk memahami proses yang kompleks serta memodelkan alur kerja secara sistematis. Melalui *flowchart*, setiap tahapan proses dapat digambarkan secara runtut sehingga mempermudah komunikasi antaranggota tim, membantu mengidentifikasi peluang perbaikan proses, serta meminimalkan kesalahan dalam pelaksanaan suatu kegiatan. Selain itu, *flowchart* juga

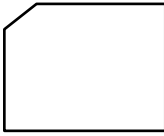
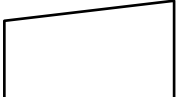
berperan sebagai media dokumentasi yang memudahkan proses analisis dan evaluasi suatu sistem.

Penelitian lain yang dilakukan oleh Budiarto et al. (2017) menyatakan bahwa *flowchart* atau bagan alir merupakan skema yang menunjukkan aliran logika suatu program dan menjadi salah satu alat yang banyak digunakan untuk menggambarkan algoritma dalam bentuk notasi tertentu. Penggunaan *flowchart* dalam pengembangan perangkat lunak membantu pengembang memahami hubungan antarproses sebelum dilakukan implementasi sehingga proses pembangunan sistem menjadi lebih terstruktur dan mudah dipelihara.

Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa *flowchart* merupakan diagram yang digunakan untuk menggambarkan urutan proses atau alur kerja suatu sistem secara logis dan sistematis menggunakan simbol-simbol standar. Dalam penelitian ini, *flowchart* digunakan untuk memvisualisasikan alur proses sistem analisis sentimen berbasis web, mulai dari proses pengambilan data komentar TikTok, *text preprocessing*, klasifikasi sentimen menggunakan metode *Bidirectional Long Short-Term Memory* (BiLSTM), hingga penyajian hasil analisis kepada pengguna.

Tabel 2. 1 Simbol – Simbol *Flowchart*

Simbol	Nama	Keterangan
	<i>Input/Output</i>	Berfungsi sebagai simbol untuk data yang masuk maupun hasil keluaran yang telah diproses
	Proses	Menandakan terjadinya suatu aktivitas operasional, instruksi, atau tahapan pemrosesan data di dalam sistem
	<i>Connector</i>	Berfungsi sebagai titik sambung alur proses yang terpisah, namun masih berada di dalam satu halaman yang sama
	<i>Off-page Connector</i>	Berfungsi untuk menghubungkan alur diagram yang terputus dan berlanjut pada halaman yang berbeda
	Anak Panah	Merepresentasikan alur kerja
	Keputusan	Penentuan kondisi atau percabangan alur dalam program
	Preparation	Deklarasi variabel atau pemberian nilai awal program
	Terminal Point	Titik mulai (<i>start</i>) dan selesainya (<i>end</i>) sebuah sistem atau program

	Punched Card	Operasi input maupun output data yang memanfaatkan media kartu berlubang
	Dokumen	Proses input maupun output informasi dalam bentuk cetakan (<i>hardcopy</i>)
	Manual Input	Proses entri data yang dilakukan secara langsung oleh pengguna menggunakan papan ketik (<i>keyboard</i>)
	Database	Pencatatan dan penyimpanan data ke dalam <i>database</i>

Sumber : (Smrti et al., 2023)

9. *Unified Modelling Language (UML)*

Menurut Havaluddin (2011), *Unified Modeling Language (UML)* adalah bahasa pemodelan visual yang bertujuan untuk menspesifikasikan, menggambarkan, membangun, serta mendokumentasikan sistem *software* berbasis objek. UML dikembangkan sebagai standar pemodelan yang mampu menggambarkan struktur maupun perilaku suatu sistem melalui notasi-notasi yang telah ditetapkan. Selain itu, UML membantu pengembang dalam memahami kebutuhan sistem, mempermudah komunikasi antaranggota tim, serta menghasilkan dokumentasi sistem yang lebih terstruktur.

Penelitian yang dilakukan oleh Nurshadrina & Voutama (2022) menjelaskan bahwa UML merupakan bahasa pemodelan yang digunakan

untuk menentukan, memvisualisasikan, membangun, dan mendokumentasikan sistem perangkat lunak. UML telah menjadi standar dalam pengembangan sistem berorientasi objek karena mampu menggambarkan kebutuhan sistem secara jelas sebelum proses implementasi dilakukan. Penggunaan UML juga membantu pengembang dalam merancang sistem yang lebih terstruktur sehingga proses pengembangan perangkat lunak menjadi lebih efektif.

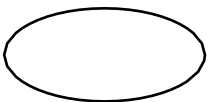
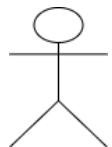
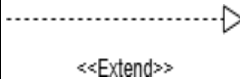
Penelitian lain yang dilakukan oleh Sastra & Setiaji (2021) menjelaskan bahwa proses pemodelan menggunakan UML dimulai dari analisis kebutuhan sistem, kemudian dilanjutkan dengan penyusunan *Use Case Diagram*, *Class Diagram*, *Sequence Diagram*, dan *Activity Diagram* sebagai dokumentasi rancangan sistem. Penggunaan UML mampu memberikan gambaran menyeluruh mengenai hubungan antaraktor, proses bisnis, struktur kelas, serta interaksi antarobjek sehingga memudahkan proses implementasi perangkat lunak.

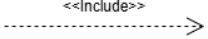
Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa *Unified Modeling Language* (UML) merupakan bahasa pemodelan visual yang digunakan untuk menganalisis, merancang, dan mendokumentasikan sistem perangkat lunak menggunakan notasi standar. Dalam penelitian ini, UML digunakan sebagai alat bantu untuk merancang sistem analisis sentimen berbasis web melalui *Use Case Diagram*, *Activity Diagram*, *Sequence Diagram*, dan *Class Diagram*.

a) *Use Case Diagram*

Menurut Haviluddin (2011), *Use Case Diagram* merupakan salah satu diagram dalam *Unified Modeling Language* (UML) yang digunakan untuk menggambarkan interaksi antara aktor dengan sistem berdasarkan fungsi atau layanan yang disediakan. Diagram ini berfungsi untuk mengidentifikasi kebutuhan fungsional sistem, menunjukkan hak akses setiap aktor, serta memberikan gambaran mengenai ruang lingkup sistem yang akan dikembangkan sehingga memudahkan proses analisis dan perancangan perangkat lunak.

Tabel 2. 1 Simbol – Simbol *Use Case Diagram*

Simbol	Nama	Keterangan
	<i>Use Case</i>	Menggambarkan fungsionalitas yang disediakan oleh sistem serta bagaimana pengguna berinteraksi dengan layanan tersebut
	<i>Actor</i>	Pihak luar atau pengguna yang memberikan masukan (instruksi) maupun menerima informasi langsung dari sistem
	<i>Association</i>	Garis relasi yang merepresentasikan jalur komunikasi langsung antara entitas aktor dengan fungsionalitas sistem (use case)
	<i>Extend</i>	Menggambarkan use case perluasan yang sifatnya independen dan tidak wajib dieksekusi oleh use case utamanya

	<i>Generalization</i>	Relasi pewarisan (<i>inheritance</i>) antar <i>use case</i> , di mana satu <i>use case</i> bertindak sebagai bentuk umum (induk) dari <i>use case</i> lain yang lebih spesifik
	<i>Include</i>	Menunjukkan pemanggilan fungsionalitas tambahan yang sifatnya wajib dan otomatis tereksekusi saat <i>use case</i> utama dijalankan

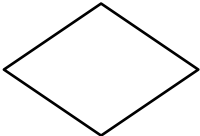
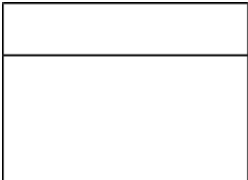
Sumber: (Haviluddin, 2011)

b) *Activity Diagram*

Menurut Haviluddin (2011), *Activity Diagram* yaitu diagram pada *Unified Modeling Language* (UML) yang digunakan untuk menggambarkan alur aktivitas atau proses bisnis dalam suatu sistem secara berurutan. *Activity Diagram* menunjukkan urutan pekerjaan mulai dari kondisi awal, proses yang dilakukan, percabangan keputusan, hingga kondisi akhir sehingga memudahkan pengembang dalam memahami logika proses yang terjadi pada sistem yang akan dibangun.

Tabel 2. 2 Simbol – Simbol *Activity Diagram*

Simbol	Nama	Keterangan
	Status awal	Penanda dimulainya seluruh rangkaian aktivitas atau alur kerja di dalam sistem
	<i>Activity</i>	Merepresentasikan sebuah tindakan, tahapan pekerjaan, atau proses eksekusi yang terjadi di dalam alur sistem

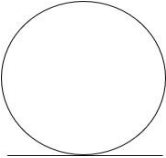
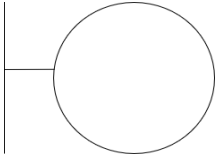
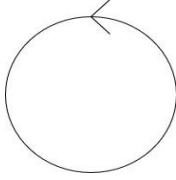
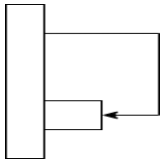
	<i>Decision</i>	Titik evaluasi logika yang memecah alur proses menjadi beberapa alternatif jalur berdasarkan syarat atau kondisi tertentu
	<i>Join</i>	Menggabungkan kembali dua atau lebih alur percabangan sistem menjadi satu aliran proses tunggal
	Status akhir	Menandakan titik penyelesaian atau berhentinya seluruh rangkaian proses dalam aliran aktivitas sistem
	<i>Swimlane</i>	Mengelompokkan rangkaian aktivitas berdasarkan peran, entitas, atau unit organisasi yang bertanggung jawab atas pelaksanaannya

Sumber: (Haviluddin, 2011)

c) *Sequence Diagram*

Menurut Haviluddin (2011), *Sequence Diagram* merupakan salah satu diagram *Unified Modeling Language* (UML) yang digunakan untuk menggambarkan urutan interaksi antar objek dalam suatu sistem berdasarkan dimensi waktu. Diagram ini memperlihatkan pesan yang dikirim dan diterima oleh setiap objek secara berurutan sehingga memudahkan pengembang dalam memahami alur komunikasi yang terjadi selama suatu proses dijalankan.

Tabel 2. 3 Simbol – Simbol *Sequence* diagram

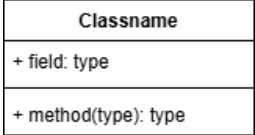
Gambar	Nama	Keterangan
	<i>Entity Class</i>	Representasi data sistem yang wujud fisiknya akan diimplementasikan sebagai ruang penyimpanan di dalam <i>database</i>
	<i>Boundary Class</i>	Merepresentasikan antarmuka (<i>interface</i>) yang menjembatani interaksi dan komunikasi antara entitas luar (aktor) dengan logika internal sistem
	<i>Controll Class</i>	Komponen yang memuat logika bisnis sistem dan bertugas mengoordinasikan interaksi antara <i>boundary class</i> dengan <i>entity class</i>
	<i>Recursive</i>	Eksekusi perintah atau instruksi internal yang ditujukan secara langsung pada objek itu sendiri
	<i>Activation</i>	Menggambarkan periode waktu atau durasi ketika sebuah objek sedang aktif mengeksekusi suatu operasi
	<i>Life Line</i>	Garis waktu vertikal yang merepresentasikan masa hidup (<i>lifeline</i>) atau eksistensi dari sebuah objek selama proses interaksi berlangsung

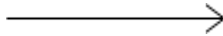
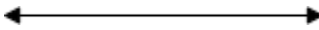
Sumber: (Haviluddin, 2011)

d) *Class Diagram*

Menurut Haviluddin (2011) *Class Diagram* merupakan diagram dalam *Unified Modeling Language* (UML) yang digunakan untuk menggambarkan struktur statis suatu sistem melalui kelas, atribut, metode, serta hubungan antar kelas. *Class Diagram* berfungsi sebagai dasar dalam perancangan basis data maupun struktur program karena mampu menunjukkan keterkaitan antar kelas secara jelas sehingga memudahkan proses implementasi sistem berbasis objek.

Tabel 2. 4 Simbol – Simbol *Class Diagram*

Simbol	Nama	Keterangan
	<i>Class</i>	Merepresentasikan cetak biru (<i>blueprint</i>) yang mendefinisikan struktur informasi, atribut, serta operasi (metode) dari suatu entitas dalam perangkat lunak
	<i>Interface</i>	Mendefinisikan sebuah kontrak fungsionalitas berupa sekumpulan operasi abstrak yang wujud detailnya wajib diimplementasikan oleh kelas lain di dalam sistem
	Asosiasi	Hubungan struktural yang mengikat dan mengkoneksikan berbagai kelas di dalam model perangkat lunak

	Asosiasi berarah	Relasi struktural yang menunjukkan arah navigasi tunggal, di mana sebuah kelas mengenali dan dapat mengakses kelas tujuan, namun tidak berlaku sebaliknya
	Kebergantungan	Relasi keterikatan struktural di mana perubahan pada satu kelas mandiri akan memengaruhi kelas lain sebagai penggunaanya
	<i>Aggregation</i>	Bentuk khusus dari asosiasi yang merepresentasikan hubungan keseluruhan-bagian antara kelas agregat (induk) dengan kelas komponennya
	<i>Generalization</i>	Relasi pewarisan yang menunjukkan tingkatan spesialisasi dari kelas yang bersifat umum menuju kelas yang lebih khusus

Sumber: (Haviluddin, 2011)

B. Kajian Empiris

Kajian empiris mengenai analisis sentimen berbasis teks menunjukkan bahwa pemanfaatan algoritma *deep learning* seperti *Bidirectional Long Short-Term Memory* (BiLSTM) memberikan hasil yang menjanjikan, terutama dalam konteks klasifikasi sentimen terhadap data teks berbahasa Indonesia yang bersumber dari media sosial. Sejumlah penelitian telah dilakukan untuk mengevaluasi performa BiLSTM dalam menangani klasifikasi sentimen dari

berbagai platform seperti Twitter/X, YouTube, dan kanal media sosial lainnya. Kajian-kajian ini menjadi pijakan penting dalam menyusun strategi dan pendekatan yang tepat dalam pengembangan sistem klasifikasi sentimen yang adaptif dan akurat.

Penelitian oleh Qososi et al. (2025) dengan judul "Implementasi BiLSTM untuk Klasifikasi Sentimen pada Kasus Pemilihan Umum 2024" menjadi salah satu studi yang relevan. Penelitian ini menerapkan BiLSTM untuk mengklasifikasikan opini masyarakat berbahasa Indonesia terhadap Pemilu 2024 yang bersumber dari Twitter sebanyak 3.085 data. Dengan menggunakan tiga jenis word embedding yang berbeda, hasil kategorisasi sentimen menggunakan BiLSTM memperlihatkan akurasi terbaik 83% menggunakan *embedding* FastText, diikuti oleh Word2Vec dan Glove dengan akurasi 82%. Temuan ini menegaskan bahwa pemilihan teknik *word embedding* turut memengaruhi performa BiLSTM dalam mengenali pola sentimen pada teks berbahasa Indonesia.

Menguatkan temuan tersebut, penelitian oleh Marcelina (2024) dalam studi berjudul "Implementasi Algoritma Bidirectional Long Short Term Memory (BiLSTM) dengan Analisis Sentimen Emoji Embedded terhadap Tingkat Kecemasan Covid-19 dan Sosial Ekonomi Masyarakat" menunjukkan efektivitas modifikasi BiLSTM dalam memproses opini publik. Dengan memanfaatkan 3.707 data *tweet* dan mengintegrasikan analisis sentimen melalui penyematan emoji (*emoji embedding*), algoritma BiLSTM berhasil mencapai tingkat akurasi yang sangat tinggi, yakni sebesar 98,08%. Hasil ini memberikan

bukti bahwa pemahaman konteks emosi (melalui emoji) dikombinasikan dengan BiLSTM sangat relevan untuk menangani teks di ruang digital publik.

Penelitian selanjutnya oleh Aini et al. (2024) pada studi "Analisis Sentimen Hasil Pemilu (*Quick Count*) Calon Presiden dan Wakil Presiden 2024 di Media Sosial Media X Menggunakan Metode *Bidirectional Long Short-Term Memory* (BiLSTM)" memperluas eksplorasi BiLSTM pada data dengan distribusi kelas yang tidak seimbang. Hasil penelitian menunjukkan bahwa model BiLSTM mencapai akurasi sebesar 86,89% dengan akurasi keseluruhan sistem sebesar 97%, di mana nilai *precision*, *recall*, dan *F1-score* tertinggi ditemukan pada kelas netral. Studi ini menunjukkan bahwa BiLSTM tetap mampu memberikan performa klasifikasi yang baik meskipun menangani teks kompleks seperti hasil pemilu yang didominasi opini netral.

Sementara itu, penelitian oleh Sitorus et al. (2025), dalam studi "Penerapan *Deep Learning* dan Analisis Sentimen" yang membandingkan tiga varian model berbasis LSTM menunjukkan keunggulan arsitektur dua arah dibandingkan model satu arah. Hasil penelitian menunjukkan bahwa model BiLSTM memberikan performa terbaik dengan akurasi sebesar 90% dan nilai weighted F1-score sebesar 0,91, unggul dalam menangkap konteks dua arah teks dibandingkan model LSTM murni maupun LSTM dengan *class weight*. Studi ini memperlihatkan keunggulan BiLSTM dalam menangkap hubungan antar kata dari dua arah (maju dan mundur) sehingga konteks kalimat dapat dipahami lebih utuh.

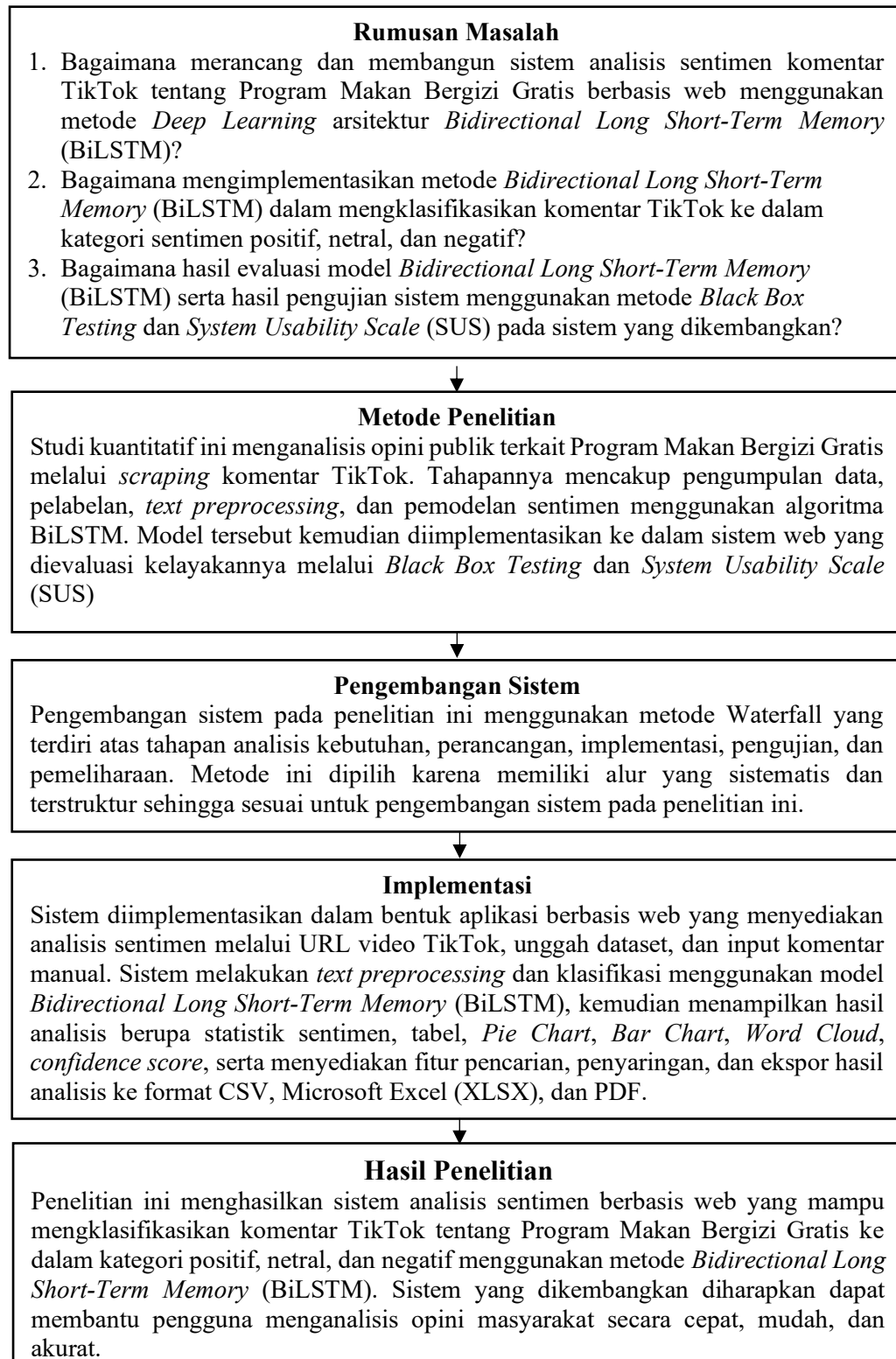
Tidak hanya terbatas pada parameter standar, efektivitas BiLSTM juga diuji melalui optimasi model. Penelitian Nasir (2023), dengan judul "Optimasi Model BiLSTM untuk Analisis Sentimen Ulasan Film Menggunakan Hyperparameter Tuning Random Search" membuktikan bahwa performa BiLSTM dapat ditingkatkan secara signifikan melalui pencarian konfigurasi parameter yang optimal. Dengan menggunakan metode *Random Search*, penelitian ini berhasil meningkatkan akurasi rata-rata keseluruhan model dari rentang 88%-89% menjadi 96,99%. Hal ini menegaskan bahwa performa BiLSTM sangat bergantung pada konfigurasi *hyperparameter* yang digunakan.

Dari keseluruhan kajian tersebut, dapat disimpulkan bahwa BiLSTM merupakan salah satu algoritma *deep learning* yang unggul dan konsisten dalam tugas klasifikasi sentimen berbasis teks, dengan rentang akurasi yang bervariasi antara 83% hingga 97% bergantung pada karakteristik dataset, teknik *word embedding*, dan konfigurasi hyperparameter yang digunakan. Keunggulan performanya tidak hanya berlaku pada teks formal, tetapi juga pada teks informal dan tidak baku yang umum ditemukan pada komentar media sosial. Efektivitas BiLSTM dalam klasifikasi sentimen sangat ditentukan oleh kualitas *preprocessing*, pemilihan *word embedding*, serta jenis dan jumlah data yang digunakan.

Adapun dalam penelitian ini, peneliti akan mengembangkan sistem analisis sentimen berbasis teks menggunakan algoritma BiLSTM, yang diimplementasikan dalam bentuk aplikasi web menggunakan *framework* FastAPI sebagai *Backend* dan React.js sebagai *Frontend*. Data yang digunakan

bersumber dari komentar konten TikTok yang membahas terkait Program Makan Bergizi Gratis (MBG), yang diperoleh melalui teknik *web scraping*. Berbeda dengan penelitian-penelitian sebelumnya yang umumnya menggunakan data dari Twitter/X atau *platform* ulasan seperti IMDB, penelitian ini berfokus pada karakteristik komentar TikTok yang cenderung lebih singkat, informal, dan kaya akan singkatan serta bahasa gaul. Fokus utama penelitian ini adalah pada penerapan BiLSTM sebagai model klasifikasi sentimen tiga kelas (positif, netral, negatif) serta penyajian hasil analisis melalui *Dashboard* interaktif yang dilengkapi visualisasi data dan fitur ekspor laporan. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengembangan sistem analisis sentimen berbasis web yang akurat, informatif, dan mudah digunakan oleh pemangku kebijakan dalam memantau opini masyarakat terhadap Program Makan Bergizi Gratis.

C. Kerangka Berpikir



Gambar 2. 1 Kerangka Berpikir