

TESIS

**ANALISIS SENTIMEN DENGAN METODE HIBRID PADA ULASAN
PELANGGAN TERHADAP TRANSAKSI PEMBELIAN PRODUK E-
COMMERCE**

*Sentiment Analysis with Soft-Voting Method on Customer Reviews for
Purchasing Transactions of E-Commerce*

MUH ZULFADLI A SUYUTI

D082192011



**PROGRAM S2 TEKNIK INFORMATIKA
DEPARTEMEN TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

**ANALISIS SENTIMEN DENGAN METODE HIBRID
PADA ULASAN PELANGGAN TERHADAP TRANSAKSI
PEMBELIAN PRODUK E-COMMERCE**

Tesis
Sebagai Salah Satu Syarat untuk Mencapai Gelar Magister
Program Studi Teknik Informatika

Disusun dan diajukan oleh

**MUH ZULFADLI A SUYUTI
D082192011**

Kepada

**FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2023**

TESIS

ANALISIS SENTIMEN DENGAN METODE HIBRID PADA ULASAN PELANGGAN TERHADAP TRANSAKSI PEMBELIAN PRODUK E-COMMERCE

MUH ZULFADLI A SUYUTI
D082192011

Telah dipertahankan di hadapan Panitia Ujian Tesis yang dibentuk dalam rangka Penyelesaian studi pada Program Magister Teknik Informatika Fakultas Teknik Universitas Hasanuddin pada tanggal 10 Agustus 2023 dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,



Dr. Amil Ahmad Ilham, S.T., M.IT.

NIP. 19731010 199802 1 001



Prof. Dr. Ir. Indrabayu, ST., MT.,

M.Bus.Sys., IPM, ASEAN, Eng.

NIP. 19750716 200212 1 004

Dekan Fakultas Teknik
Universitas Hasanuddin



Prof. Dr. Eng. Ir. Muh. Isran. Ramli, S.T., M.T.

NIP. 19730926 200012 1 002

Ketua Program Studi
S2 Teknik Informatika



Dr. Ir. Zahir Zainuddin, M.Sc.

NIP. 19640427 198910 1 002

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Yang bertanda tangan di bawah ini

Nama : Muh Zulfadli A Suyuti
Nomor mahasiswa : D082192011
Program studi : S2 Teknik Informatika

Dengan ini menyatakan bahwa, tesis berjudul “Analisis Sentimen Dengan Metode hybrid pada Ulasan Pelanggan terhadap Transaksi Pembelian Produk *E-Commerce*” adalah benar karya saya dengan arahan dari komisi pembimbing (Dr. Amil Ahmad Ilham, S.T., M.IT. dan Prof. Dr. Ir. Indrabayu, S.T., M.T., M.Bus.Sys., IPM. ASEAN. Eng.). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Prosiding (*2023 IEEE International Conference On Industry 4.0, Artificial Intelligence, And Communication Technology (IAICT)*), Bali, Indonesia, 2023, pp 295-299, DOI : 10.1109/IAICT59002.2023.10205954) sebagai artikel dengan judul “*Sentiment Analysis with Soft-Voting Method on Cumstomer Reviews for Purchasing Transaction of E-Commerce*”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Gowa, 14 Agustus 2023

Yang menyatakan



Muh Zulfadli A Suyuti

KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh

Puji Syukur Senantiasa kita panjatkan kehadirat Allah SWT, yang telah mencurahkan Kasih Sayang serta Rahmat dan Hidayah-Nya sehingga penelitian dengan judul “**Analisis Sentimen Dengan Metode Hibrid Pada Ulasan Pelanggan Terhadap Transaksi Pembelian Produk E-Commerce**” dapat terselesaikan dengan baik. Sholawat serta Salam tidak lupa kita senantiasa kirimkan kepada junjungan Nabi Besar Muhammad SAW, yang menjadi suri tauladan bagi seluruh umat manusia.

Tesis ini disusun untuk memenuhi persyaratan untuk memperoleh gelar **Magister Komputer (M.Kom)** pada program S2 Teknik Informatika, Departemen Teknik Informatika, Universitas Hasanuddin, Makasar. Dengan mengucap syukur yang sedalam-dalamnya, gelar ini penulis persembahkan kepada kedua orang tua tercinta, **Ansar Suyuti** dan ibunda Satriani Said Akhmad yang senantiasa memberikan dukungan baik dalam segi moral dan materil, motivasi dan doa yang tidak henti-hentinya dipanjatkan sehingga penulis dapat menyelesaikan penelitian ini dengan baik serta istri tercinta, Musdalifa Zaenal dan anakku tersayang Fathan Arshaka ZA Suyuti yang senantiasa memberikan semangat dan dukungan kepada penulis.

Dalam penyusunan tesis ini, tentunya tidak lepas dari dukungan dari seluruh pihak. Dengan segala kerendahan hati, penulis menyampaikan terima kasih yang sebesar-besarnya kepada :

1. Dewan pembimbing, Bapak Dr. Amil Ahmad Ilham, S.T., M.IT. dan Bapak Prof. Dr. Ir. Indrabayu, ST., MT., M.Bus.Sys., IPM, ASEAN. Eng., yang dengan sabar dan penuh tanggung jawab memberikan bimbingan, masukan serta motivasi kepada penulis dalam penyelesaian tesis ini.
2. Dewan Penguji, Bapak Dr. Eng. Muhammad Nizwar, ST., M, IT., Bapak Dr. Eng. Zulkifli Tahir, S.T., M.Sc., dan Ibu Prof. Dr. Eng. Intan Sari Areni, ST.

MT. yang senantiasa memberikan saran yang membangun selama penelitian ini dilakukan.

3. Bapak Dr. Ir. Zahir Zainuddin, M.Sc selaku Ketua Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin yang telah memberikan motivasi, bimbingan, dan semangatnya selama masa perkuliahan penulis
4. Teman-teman Pascasarjana UNHAS Angkatan 1 atas dukungan dan semangat yang diberikan selama ini;
5. Ibu Yuanita serta segenap Staf Departemen Magister Teknik Informatika yang telah banyak membantu penulis selama pengurusan administrasi;
6. Rekan-rekan mahasiswa Pascasarjana Teknik Informatika Universitas Hasanuddin yang selalu aktif dalam berdiskusi dan memberikan masukan dalam penyelesaian penelitian ini.

Penulis menyadari bahwa penyusunan tesis ini masih sangat jauh dari kata sempurna dan mempunyai banyak kekurangan. Sehingga penulis sangat mengharapkan kritik dan saran untuk kedepannya sehingga dapat memberikan manfaat bagi seluruh pembaca. Akhir kata, penulis menyampaikan permohonan yang sebesar-besarnya kepada pembaca sekiranya terdapat kesalahan-kesalahan dalam penyusunan tesis ini, wassalamualaikum warahmatullahi wabarakatuh.

Makassar, 1 Agustus 2023

Muh Zulfadli A Suyuti

ABSTRAK

MUH ZULFADLI A SUYUTI. *Analisis Sentimen Dengan Metode Hibrid Pada Ulasan Pelanggan Terhadap Transaksi Pembelian Produk E-Commerce.* (dibimbing oleh **Amil Ahmad Ilham** dan **Indrabayu**).

Akurasi ulasan pelanggan sangat penting bagi platform e-commerce untuk membantu pembeli dalam memilih produk berkualitas tinggi dari berbagai pilihan. Penelitian ini bertujuan untuk mengembangkan model analisis sentimen untuk mengevaluasi pendapat pelanggan yang diungkapkan dalam ulasan produk e-commerce. Pendekatan yang diusulkan menggunakan teknik Soft Voting (SV), yang menunjukkan akurasi yang lebih baik dibandingkan metode Sentiment Selector (SS) konvensional. Akurasi model analisis sentimen ditentukan dengan mengumpulkan nilai probabilitas dari tiga klasifikasi (Support Vector Machine (SVM), Random Forest (RF), dan Naive Bayes (NB)) untuk setiap kategori sentimen (positif, netral, negatif). Evaluasi dilakukan menggunakan dataset ulasan produk Tokopedia. Temuan menunjukkan bahwa model Soft Voting (SV) lebih baik daripada pendekatan Sentiment Selector (SS). Model SV yang diusulkan mencapai akurasi, presisi, recall, dan f1-score sebesar 69%, 70%, 69%, dan 69%, secara berturut-turut.

Kata Kunci : *Soft Voting, Sentiment Selector, Support Vector Machine, Random Forest, Naive Bayes*

ABSTRACT

MUH ZULFADLI A SUYUTI. *Sentiment Analysis with Soft-Voting Method On Customer Reviews For Purchasing Transactions Of E-Commerce.* (Supervised by **Amil Ahmad Ilham** and **Indrabayu**)

The accuracy of customer reviews is crucial for an e-commerce platform to assist buyers in selecting high-quality products from a vast array of options. This research aims to develop a sentiment analysis model for evaluating customer opinions expressed in e-commerce product reviews. The proposed approach utilizes the Soft Voting (SV) technique, which demonstrates superior accuracy compared to the conventional Sentiment Selector (SS) method. The sentiment analysis model's accuracy is determined by gathering probability values from three classifiers (Support Vector Machine (SVM), Random Forest (RF), and Naive Bayes (NB)) for each sentiment category (positive, neutral, negative). The evaluation is conducted using the Tokopedia product review dataset. The findings indicate that the Soft Voting (SV) model outperforms the Sentiment Selector (SS) approach. The proposed SV model achieves an accuracy, precision, recall, and f1-score of 69%, 70%, 69%, and 69%, respectively.

Keyword: *Soft Voting, Sentiment Selector, Support Vector Machine, Random Forest, Naive Bayes*

DAFTAR ISI

	<u>Halaman</u>
HALAMAN JUDUL	i
LEMBAR PENGESAHAN TESIS	ii
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
BAB I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
1.1 Kajian Pustaka	5
BAB II. KAJIAN LITERATUR DAN METODE PENYELESAIAN MASALAH	15
2.1 Metode Penyelesaian Masalah	15
2.2 Metode yang Diusulkan	25
2.3 Target Hasil Penelitian	25
2.4 Kerangka Pikir	26
BAB III. METODOLOGI PENELITIAN	27
3.1 Tahapan Penelitian	27
3.2 Skenario Pengujian	39
BAB IV. HASIL DAN PEMBAHASAN	41
4.1 Hasil	41
BAB V. PENUTUP	53
5.1 Kesimpulan	53
5.2 Saran	54
DAFTAR PUSTAKA	56

DAFTAR GAMBAR

Nomor	Halaman
Gambar 1. Kerangka Pikir	26
Gambar 2. Alur Penelitian	27
Gambar 3. Diagram Proses	28
Gambar 4. Model Sentiment Selector	33
Gambar 5. Model Soft Voting	36

DAFTAR TABEL

Nomor	Halaman
Tabel 1. <i>State of the Art</i>	15
Tabel 2. Analisis Perbandingan dari Setiap Klasifikasi	46
Tabel 3. Contoh Data Pada Kasus Prediksi Untuk Review pada Tokopedia	49
Tabel 4. Analisis Komparatif antara Sentiment Selector dan Soft Voting	50

BAB I. PENDAHULUAN

1.1 Latar Belakang

Gaya hidup masyarakat semakin menjadi-jadi karena mayoritas dari mereka memiliki mobilitas tinggi dalam aktivitas sehari-hari. Kebutuhan masyarakat terhadap produk telah terpenuhi dengan munculnya platform e-commerce saat ini. Hal ini diikuti oleh penyediaan berbagai layanan yang membantu masyarakat (Kumara *et.al.*, 2022). Setiap perusahaan menawarkan berbagai keunggulan kepada konsumennya agar layanan mereka dapat terus bertahan. Platform e-commerce yang menarik, mudah, dan menguntungkan bagi konsumen semakin memberikan keuntungan yang lebih besar bagi perusahaan.

E-commerce merujuk pada pembelian dan penjualan elektronik berbagai produk antara bisnis, yang difasilitasi oleh teknologi komputer sebagai perantara untuk aktivitas transaksional. E-commerce telah mengalami pertumbuhan yang luar biasa dan saat ini sedang mengalami tahap transformasi, yang pada akhirnya akan membuatnya menjadi pasar yang paling kompetitif dan rumit (Chenyue *et.al.*, 2022). Angka yang diantisipasi menunjukkan bahwa jumlah pembeli digital di seluruh dunia akan melebihi 2 miliar pada tahun 2020, dan proporsi perdagangan melalui ponsel dalam semua transaksi e-commerce akan meningkat menjadi sekitar 73% pada tahun 2021, menghasilkan penjualan hingga \$4,88 triliun. Pada tahun 2040, diprediksi bahwa e-commerce akan mencakup 95% dari semua pembelian

oleh konsumen, dan 85% dari mereka akan melakukan riset produk secara online sebelum melakukan pembelian (Clement *et.al.*, 2020).

Saat ini, konsumen dapat membeli sejumlah produk yang tidak terbatas melalui online. Saat ini di Indonesia terdapat banyak situs e-commerce lokal yang dibuat oleh perusahaan-perusahaan dalam negeri, seperti Tokopedia, Bukalapak, dan lainnya (Kowanda *et.al.*, 2018). Platform ini berisi ribuan jenis produk yang dibutuhkan oleh masyarakat dalam kehidupan sehari-hari. Sayangnya, produk-produk yang muncul di setiap platform tersebut sangat mirip, sehingga konsumen bingung dan sulit untuk memutuskan di mana mereka harus membeli produk yang paling cocok untuk mereka. Konsumen perlu menghabiskan lebih banyak waktu memilih produk di setiap platform dengan pertimbangan-pertimbangan seperti melihat rating, harga, dan ulasan produk tersebut. Hal ini dapat membuat konsumen bosan dan akhirnya membatalkan pembelian produk. Peluang bagi konsumen untuk mendapatkan produk dengan kualitas rendah juga sangat tinggi karena mereka tertipu oleh rating platform tersebut (Casas-Valadez *et.al.*, 2020).

Sentimen konsumen sangat penting karena dapat mempengaruhi keputusan pembelian mereka. Oleh karena itu, perusahaan e-commerce perlu secara rutin memantau sentimen konsumen untuk meningkatkan pengalaman pelanggan dan peningkatan penjualan. Untuk memantau sentimen konsumen, dapat digunakan analisis sentimen. Analisis sentimen (SA) adalah teknik pemrosesan bahasa alami yang memungkinkan komputer untuk menganalisis dan memahami opini atau sentimen dalam teks (Appel *et.al.*, 2016). Teknik SA dapat digunakan untuk menganalisis ulasan produk, komentar pelanggan,

dan pesan media sosial untuk memahami sentimen konsumen terhadap merek atau produk tertentu (Weidong *et.al.*, 2020).

Kinerja model pembelajaran mesin tradisional dibandingkan dengan model deep learning pada tugas analisis sentimen (Narayama *et.al.*, 2018). Data yang digunakan adalah dataset Ulasan IMDB, Yelp, dan Amazon. Dataset ini terdiri dari ulasan film, restoran, dan produk dari platform online. Hasil pengujian menunjukkan bahwa model deep learning, khususnya Convolutional Neural Network (CNN) dan Recurrent Neural Network (RNN), menunjukkan kinerja yang lebih baik dalam tugas analisis sentimen dibandingkan dengan model pembelajaran mesin konvensional seperti Naive Bayes, Decision Tree, dan SVM. Model deep learning mampu menangani pola-pola kompleks dan menghasilkan akurasi yang lebih baik dibandingkan dengan model pembelajaran mesin tradisional. Kesimpulan hasil penelitian ini ke dataset lain harus dilakukan dengan hati-hati. Penelitian ini tidak mempertimbangkan faktor-faktor eksternal seperti konteks sosial dan budaya dalam analisis sentimen.

Penelitian ini memperkenalkan model analisis sentimen untuk e-commerce yang menggunakan teknik soft voting. Ini melibatkan penggunaan kumpulan klasifikasi, termasuk Support Vector Machine (SVM), Random Forest (RF), dan Naive Bayes (NB), untuk memprediksi sentimen kalimat yang dipilih. Model soft voting kemudian menggabungkan prediksi dari klasifikasi-klasifikasi tersebut untuk mendapatkan prediksi sentimen akhir.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang dibahas, maka ditarik suatu rumusan masalah pada penelitian ini yaitu bagaimana memberikan gambaran yang tepat terhadap transaksi pembelian suatu produk e-commerce berdasarkan hasil analisis sentimen menggunakan metode hibrid pada ulasan pengguna?

1.3 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini yaitu mengembangkan model analisis sentimen pada ulasan pelanggan terhadap transaksi pembelian suatu produk e-commerce menggunakan metode hibrid dengan nilai akurasi hasil analisis sentimen yang lebih tinggi dibandingkan dengan metode non hybrid.

1.4 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah diharapkan memberikan kontribusi keilmuan dalam pengembangan metode hibrid analisis sentimen pada bidang e-commerce dan memberikan manfaat bagi peningkatan kepuasan penngguna/pelanggan dalam melakukan transaksi pembelian suatu produk e-commerce.

1.5 Batasan Masalah

Adapun batasan masalah pada penelitian ini yaitu:

1. Analisis dilakukan terhadap komentar pelanggan pada proses transaksi suatu produk e-commerce (bukan review sepesifikasi teknis suatu produk e-commerce).
2. Data ulasan transaksi produk diambil dari Tokopedia.

1.1 Kajian Pustaka

Sebuah survei yang diselenggarakan Asosiasi Penyelenggaraan Jasa Internet Indonesia (APJII) menunjukkan bahwa jumlah pengguna internet di Indonesia meningkat tiap tahunnya hingga akhir tahun 2018 mencapai 171,17 juta orang, dari total penduduk sebanyak 264,16 juta orang. perkembangan tersebut berdampak pada berbagai bidang. Salah satunya ialah maraknya kegiatan berbelanja melalui media internet. Tokopedia dan Bukalapak mempunyai sebuah akun Twitter customer care yaitu @tokopediacare untuk Tokopedia dan @bukabantuan untuk Bukalapak (Srivats *et.al.*, 2018). Akun tersebut adalah salah satu bentuk layanan pelanggan khusus melalui online yang disediakan untuk menanggapi tanggapan, pendapat, kritik, saran dan masalah komplain. Komentar dalam twitter berbentuk teks, sehingga perlu dilakukan analisis text mining. Tujuan dari text mining yaitu untuk memperoleh informasi dari sekumpulan dokumen serta mendukung proses knowledge discovery pada koleksi dokumen yang besar, sumber data yang digunakan pada text mining merupakan kumpulan teks dengan format yang tidak teratur atau minimal semi teratur (Juanita *et.al.*, 2022).

Konsumen dari belanja online dapat memberikan feedback secara terbuka, saling berkomentar dalam waktu cepat dan tidak terbatas terhadap pelayanan dari Tokopedia dan Bukalapak melalui twitter customer care. Mengetahui sentiment dari konsumen belanja online secara manual dapat merugikan waktu, dan tenaga. Oleh karena itu, pada penelitian ini dilakukan analisis text mining dengan membandingkan metode klasifikasi Support Vector Machine dan Decision Tree. Informasi yang didapatkan dari e-

commerce tidak dapat menggambarkan struktur komunikasi dan tingkat partisipasi dari setiap pelanggan. Oleh karena itu diperlukan suatu metode yang dapat memvisualisasikan atau memeriksa pola interaksi pengguna Tokopedia dan Bukalapak menggunakan Word. Tujuan yang akan dicapai dalam penelitian ini adalah mengetahui metode klasifikasi terbaik antara Support Vector Machine dan Naive Bayes. Selain itu, juga ingin mengetahui kata-kata yang sering muncul pada masing-masing sentimen menggunakan visualisasi word cloud. Manfaat yang dapat diperoleh dari penelitian ini adalah mampu mengetahui proses mengekstrak dan mengolah data teks secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini dengan hasil metode klasifikasi terbaik (Rohman *et.al.*, 2022). Selain itu juga dapat mengidentifikasi pola pengguna twitter konsumen Tokopedia dan Bukalapak sehingga dapat dimanfaatkan lebih maksimal untuk menyebarkan informasi secara lebih efektif. Dengan diketahuinya hasil visualisasi kata berdasarkan hasil tampilan word cloud maka dapat diketahui kata apa saja yang paling sering muncul dalam tanggapan positif negatif, maupun tanggapan netral (Sutabri *et.al.*, 2018).

Random Forest adalah sebuah metode ensemble learning yang digunakan dalam pembelajaran mesin untuk tugas klasifikasi dan regresi. Ini adalah kombinasi dari beberapa pohon keputusan yang dikenal sebagai "pohon acak" (random trees). Metode ini dikembangkan oleh Leo Breiman dan Adele Cutler (Singh *et.al.*, 2020).

Berikut ini adalah penjelasan lengkap tentang Random Forest:

1. **Pohon Keputusan (Decision Trees):** Sebelum memahami Random Forest, penting untuk memahami pohon keputusan. Pohon keputusan adalah model pembelajaran mesin yang menggambarkan keputusan dan konsekuensi mereka dalam bentuk struktur pohon yang terdiri dari simpul dan cabang. Setiap simpul dalam pohon ini mewakili atribut atau fitur, sedangkan cabang-cabangnya mewakili keputusan dan hubungan antara atribut tersebut.
2. **Ensemble Learning:** Ensemble Learning adalah pendekatan untuk membangun model yang kuat dengan menggabungkan beberapa model yang lebih lemah. Ini dilakukan dengan menggabungkan prediksi dari model-model individu untuk menghasilkan prediksi akhir yang lebih baik. Random Forest adalah salah satu teknik Ensemble Learning yang populer.
3. **Pembentukan Random Forest:** Random Forest terdiri dari beberapa pohon keputusan. Proses pembentukannya melibatkan dua tingkat acak: pertama, setiap pohon keputusan dibangun dengan menggunakan sampel acak dengan penggantian dari dataset pelatihan. Kedua, setiap simpul pembentukan pohon dipilih secara acak dari subset fitur yang dipilih secara acak. Dengan menggabungkan acak ini, setiap pohon dalam Random Forest menjadi sedikit berbeda dari yang lainnya.
4. **Proses Pelatihan:** Proses pelatihan Random Forest dimulai dengan membagi dataset pelatihan menjadi subset acak. Kemudian, setiap subset digunakan untuk membangun satu pohon keputusan. Saat membangun pohon, algoritma membagi setiap simpul berdasarkan aturan yang mengoptimalkan pemisahan kelas target. Proses ini berlanjut hingga pohon

mencapai kriteria berhenti, seperti mencapai kedalaman maksimum atau jumlah sampel minimum dalam simpul.

5. Proses Prediksi: Setelah Random Forest terbentuk, prediksi dilakukan dengan mengambil mayoritas suara dari semua pohon dalam ensemble. Dalam kasus klasifikasi, kelas yang mendapatkan suara terbanyak dipilih sebagai prediksi akhir. Sedangkan dalam kasus regresi, prediksi akhir dapat berupa nilai rata-rata dari semua prediksi pohon.
6. Keuntungan Random Forest: Random Forest memiliki beberapa keuntungan, termasuk:
 - Mampu menangani dataset yang besar dan kompleks dengan baik.
 - Mampu mengatasi masalah overfitting karena variasi yang diperkenalkan oleh pohon acak.
 - Menghasilkan estimasi pentingnya fitur, yang dapat membantu dalam pemilihan fitur.
 - Relatif mudah diimplementasikan dan memiliki waktu prediksi yang cepat.

Random Forest telah terbukti berhasil dalam berbagai aplikasi, termasuk klasifikasi teks, pengenalan citra, dan bioinformatika. Namun, seperti halnya dengan metode pembelajaran mesin lainnya, pemilihan dan tuning parameter yang tepat sangat penting dalam mencapai performa yang optimal dengan menggunakan Random Forest. Setiap pohon dalam kumpulan dibangun dari sampel yang diambil dengan penggantian (bootstrap) dari set data pelatihan.

Selain itu, saat membagi setiap node saat membangun pohon, pembagian terbaik ditemukan dari semua fitur masukan atau dari subset acak dengan ukuran *max_features*.

Tujuan dari dua sumber ketidakteraturan ini adalah untuk mengurangi varians dari estimator hutan karena pohon keputusan individual biasanya menunjukkan varians tinggi dan cenderung overfitting. Ketidakteraturan yang dimasukkan ke dalam hutan menghasilkan pohon keputusan dengan kesalahan prediksi yang agak terlepas. Dengan mengambil rata-rata dari prediksi tersebut, beberapa kesalahan dapat dibatalkan. Hutan acak mencapai pengurangan varians dengan menggabungkan pohon-pohon yang beragam, kadang-kadang dengan biaya peningkatan sedikit pada bias. Dalam praktiknya, pengurangan varians seringkali signifikan sehingga menghasilkan model yang secara keseluruhan lebih baik.

Implementasi *scikit-learn* menggabungkan *classifier* dengan mengambil rata-rata dari prediksi probabilistik mereka, bukan dengan membiarkan setiap *classifier* memilih untuk satu kelas saja. Modul *sklearn.ensemble* mencakup dua algoritma perataan berdasarkan pohon keputusan acak: algoritma RandomForest dan metode Extra-Trees. Kedua algoritma ini merupakan teknik perubahan-dan-gabung yang dirancang khusus untuk pohon keputusan. Artinya, sebuah set beragam classifier dibuat dengan memasukkan ketidakteraturan dalam konstruksi classifier. Prediksi dari kumpulan ini diberikan sebagai prediksi rata-rata dari classifier-individu.

Algoritma Naive Bayes adalah sebuah metode klasifikasi yang berdasarkan pada Teorema Bayes dengan asumsi naif (naive assumption)

bahwa semua fitur yang terlibat dalam klasifikasi adalah independen secara bersyarat (Mahyarani *et.al.*, 2021). Algoritma ini sangat efisien dan efektif dalam banyak aplikasi, terutama dalam pengolahan teks dan klasifikasi dokumen. Klasifikasi Naive Bayes telah bekerja cukup baik dalam banyak situasi dunia nyata, terutama klasifikasi dokumen dan penyaringan spam. Metode ini memerlukan sedikit data pelatihan untuk mengestimasi parameter yang diperlukan. Penguraian distribusi fitur bersyarat kelas berarti setiap distribusi dapat diestimasi secara independen sebagai distribusi satu dimensi. Hal ini membantu mengatasi masalah yang berasal dari kutukan dimensionalitas.

Di sisi lain, meskipun Naive Bayes dikenal sebagai klasifikasi yang baik, ia juga dikenal sebagai estimator yang buruk, sehingga output probabilitas dari *predict_proba* tidak boleh dianggap terlalu serius. Metode dalam pengklasifikasi dan pengelompokan yang dapat menghasilkan estimasi probabilitas untuk setiap kelas/kluster. Inputnya biasanya hanya beberapa data yang diamati, X . Jika estimator belum di-fitkan sebelumnya, pemanggilan metode ini harus menyebabkan pengecualian `NotFittedError`.

Konvensi output mirip dengan `decision_function` kecuali dalam kasus klasifikasi biner, di mana satu kolom keluaran diberikan untuk setiap kelas (sedangkan `decision_function` mengeluarkan array 1 dimensi). Untuk prediksi biner dan multikelas, setiap baris harus berjumlah 1.

predict_proba hanya akan ada jika estimator dapat membuat prediksi probabilitas (lihat duck typing). Ini berarti keberadaan metode tersebut mungkin tergantung pada parameter estimator (misalnya pada

`linear_model.SGDClassifier`) atau data pelatihan (misalnya pada `model_selection.GridSearchCV`) dan mungkin hanya muncul setelah proses fitting dilakukan.

Berikut ini adalah penjelasan lengkap tentang algoritma Naive Bayes:

1. Teorema Bayes: Naive Bayes didasarkan pada Teorema Bayes, yang memberikan cara untuk menghitung probabilitas suatu kejadian berdasarkan informasi yang ada sebelumnya. Teorema Bayes dinyatakan sebagai berikut:

$$P(A|B) = (P(B|A) * P(A)) / P(B) \quad (1)$$

Di mana:

- $P(A|B)$ adalah probabilitas suatu kejadian A terjadi, diberikan bahwa kejadian B telah terjadi.
 - $P(B|A)$ adalah probabilitas kejadian B terjadi, diberikan bahwa kejadian A telah terjadi.
 - $P(A)$ adalah probabilitas kejadian A terjadi secara independen.
 - $P(B)$ adalah probabilitas kejadian B terjadi secara independen.
2. Asumsi Naif: Naive Bayes membuat asumsi naif bahwa semua fitur yang digunakan dalam klasifikasi adalah independen secara bersyarat. Dalam konteks klasifikasi, ini berarti bahwa nilai fitur tidak saling mempengaruhi satu sama lain setelah diketahui kelasnya. Meskipun asumsi ini jarang terpenuhi secara benar dalam dunia nyata, Naive Bayes sering kali memberikan hasil yang baik dan memiliki kinerja yang baik dalam berbagai aplikasi.

3. **Proses Pelatihan:** Proses pelatihan Naive Bayes melibatkan menghitung probabilitas setiap kelas dan probabilitas kondisional dari setiap fitur, diberikan kelasnya. Probabilitas kelas diperoleh dengan menghitung jumlah kemunculan setiap kelas dalam dataset pelatihan. Probabilitas kondisional fitur diperoleh dengan menghitung jumlah kemunculan nilai fitur tertentu dalam setiap kelas. Hal ini dilakukan untuk setiap fitur yang digunakan dalam klasifikasi.
4. **Proses Prediksi:** Setelah proses pelatihan selesai, Naive Bayes dapat digunakan untuk membuat prediksi pada data uji. Untuk setiap sampel data uji, Naive Bayes menghitung probabilitas kelas yang mungkin berdasarkan probabilitas yang diestimasi selama pelatihan. Prediksi akhir dilakukan dengan memilih kelas dengan probabilitas tertinggi.
5. **Jenis Naive Bayes:** Terdapat beberapa jenis algoritma Naive Bayes yang umum digunakan, termasuk:
 - **Naive Bayes Multinomial:** Digunakan untuk data dengan distribusi multinomial, seperti pengolahan teks.
 - **Naive Bayes Gaussian:** Digunakan untuk data numerik yang mengikuti distribusi Gaussian.
 - **Naive Bayes Bernoulli:** Digunakan untuk data biner atau boolean.
6. **Keuntungan Naive Bayes:** Naive Bayes memiliki beberapa keuntungan, termasuk:
 - Sederhana dan mudah diimplementasikan.
 - Berkinerja cepat dalam pelatihan dan prediksi.
 - Stabil dalam menghadapi variabel yang tidak relevan.

- Tidak memerlukan jumlah data pelatihan yang besar untuk menghasilkan model yang baik.

Namun, perlu diingat bahwa asumsi naif pada Naive Bayes dapat menjadi pembatasan. Jika fitur-fitur yang digunakan saling terkait secara kuat, asumsi tersebut mungkin tidak terpenuhi, dan hasil klasifikasi bisa menjadi tidak akurat.

Algoritma SVM (Support Vector Machine) adalah sebuah metode pembelajaran mesin yang digunakan untuk tugas klasifikasi dan regresi (Uma *et.al.*, 2022). SVM bekerja dengan membangun sebuah hyperplane dalam ruang fitur yang memaksimalkan pemisahan antara kelas-kelas yang berbeda. Metode ini sangat populer dan efektif dalam banyak aplikasi, termasuk pengenalan citra, analisis teks, dan bioinformatika.

1. Keuntungan SVM:

- SVM memiliki kemampuan yang baik dalam menangani data dengan dimensi tinggi.
- SVM efektif dalam kasus data yang terpisah secara linier maupun tidak linier.
- Dalam kasus data yang terdapat outlier, SVM cenderung lebih tahan terhadap pengaruhnya.
- SVM memiliki sedikit parameter yang harus ditentukan, seperti C (parameter penalti) dan jenis kernel.

2. Keterbatasan SVM:

- SVM cenderung menjadi lebih lambat dalam pelatihan dan prediksi ketika ukuran dataset sangat besar.

- Memilih kernel yang tepat bisa menjadi tantangan. Penggunaan kernel yang tidak tepat dapat menghasilkan performa yang buruk.

SVM telah terbukti berhasil dalam berbagai aplikasi, terutama dalam kasus klasifikasi biner dan multi-kelas. Dalam beberapa kasus, SVM dapat mengungguli metode pembelajaran mesin lainnya dalam hal akurasi klasifikasi. SVM dalam *scikit-learn* mendukung vektor sampel yang padat (*numpy.ndarray* dan dapat diubah menjadi itu melalui *numpy.asarray*) dan vektor sampel yang jarang (dalam format *scipy.sparse*) sebagai input. Namun, untuk menggunakan SVM untuk membuat prediksi pada data yang jarang, SVM harus telah dipasangkan pada data semacam itu. Untuk performa yang optimal, gunakan *numpy.ndarray* berurutan-C (dense) atau *scipy.sparse.csr_matrix* (sparse) dengan `dtype=float64`.

Kelas ini menggunakan validasi silang untuk mengestimasi parameter klasifikasi dan selanjutnya mengkalibrasi klasifikasi. Dengan `ensemble=True` (default), untuk setiap pembagian cv, ia memasang salinan estimator dasar ke subset pelatihan, dan mengkalibrasinya menggunakan subset pengujian. Untuk prediksi, probabilitas yang diprediksi dirata-ratakan di antara klasifikasi terkalibrasi individu ini. Ketika `ensemble=False`, validasi silang digunakan untuk mendapatkan prediksi yang tidak bias, melalui `cross_val_predict`, yang kemudian digunakan untuk kalibrasi. Untuk prediksi, estimator dasar yang dilatih menggunakan semua data digunakan. Ini adalah metode yang diimplementasikan ketika `probabilities=True` untuk estimator `sklearn.svm`.

Klasifikasi yang sudah difitkan dapat dikalibrasi melalui parameter $cv="prefit"$. Dalam kasus ini, tidak ada validasi silang yang digunakan dan semua data yang disediakan digunakan untuk kalibrasi. Pengguna harus memastikan secara manual bahwa data untuk pelatihan model dan kalibrasi tidak tumpang tindih.

Kalibrasi didasarkan pada metode `decision_function` dan jika ada, jika tidak, digunakan metode `predict_proba`.

BAB II. KAJIAN LITERATUR DAN METODE PENYELESAIAN MASALAH

2.1 Metode Penyelesaian Masalah

2.2.1 State of The Art

Beberapa Penelitian rujukan yang penulis ambil dijabarkan pada Tabel 1 *State of The Art* sebagai berikut:

Tabel 1. *State of the Art*

No	Judul, Nama, Tahun, Penerbit	Objek dan Permasalahan	Metode Penyelesaian	Kinerja	Korelasi <, =, >
1.	Judul: Analisis Sentimen Dengan Metode Hibrid pada Ulasan Pelanggan Terhadap Transaksi Pembelian Produk E-	Objek : Review produk E-Commerce Permasalahan: Bagaimana meningkatkan nilai akurasi hasil sentimen analisis terhadap transaksi pembelian suatu produk e-commerce?	Metode Hybrid dengan menggabungkan Soft Voting dan Sentiment Selector	Kinerja dari topik yang diusulkan mampu meningkatkan akurasi sentiment analisis.	>

	Commerce Penulis : Muh Zufadli A Suyuti				
2.	Judul : Sentiment Analysis of Online Product Reviews Based On SenBERT- CNN [13] Penulis : F. Wu, Z. Shi, Z. Dong, C. Pang, dan B. Zhang Tahun: 2020 Penerbit : International Conference on Machine Learning and Cybernetics (ICMLC)	Objek : Review Produk E-Commerce Permasalahan: mengembangkan sebuah metode untuk analisis sentimen pada ulasan produk online menggunakan pendekatan gabungan antara model bahasa bertingkat dan jaringan saraf konvolusi?	Pre-trained model BERT yang disebut "SenBERT" yang telah dilatih pada data ulasan produk dalam bahasa Tionghoa	Hasil evaluasi ulasan produk dunia nyata menunjukkan bahwa model SenBERT-CNN yang diusulkan jauh lebih baik daripada beberapa metode dasar serta model pembelajaran mendalam yang sudah ada.	<
3.	Judul : Sentiment Analysis on Marketplace Review using Hybrid Lexicon	Objek: Review Produk E- Commerce Permasalahan: Bagaimana	Sentiwordnet 3.0 and Support Vector Machine (SVM).	Hasil eksperimen menunjukkan bahwa metode gabungan ini mengungguli pendekatan	=

	<p>and SVM Method [14]</p> <p>Penulis : M. M. Dwi Putra, W. Fawwaz Al Maki and A. Romadhony</p> <p>Tahun : 2021</p> <p>Penerbit : ICoICT</p>	<p>meningkatkan hasil sentiment analisis pada ulasan produk e-commerce?</p>		<p>leksikon dan pendekatan SVM</p>	
4.	<p>Judul : Application of Bidirectional Gated Recurrent Unit (BiGRU) in Sentiment Analysis of Tokopedia Application Users [15]</p> <p>Penulis : D. Ananda, T. A. Taqiyyuddin, I. Nugraha Faqih, R. Badrahadipura and A. A. Pravitasari</p> <p>Tahun : 2021</p> <p>Penerbit :</p>	<p>Objek : Review Produk Aplikasi Tokopedia pada Google Play.</p> <p>Permasalahan: Bagaimana menghasilkan prediksi dengan akurasi tinggi terhadap suatu aplikasi?</p>	<p>Docker Bidirectional gated recurrent unit (BiGRU)</p>	<p>BiGRU mampu memprediksi ulasan pelanggan dengan akurasi 74,79%, recall 74,80%, dan F1-Score 70,88%, yang dianggap cukup baik dan dapat digunakan untuk membantu pengembangan Tokopedia.</p>	=

	International Conference on Artificial Intelligence and Big Data Analytics				
5.	<p>Judul : Value Customer Reviews in E-Commerce Platforms using Sentiment Analysis Tactics: A Review [1]</p> <p>Penulis : W. Kumara, P. Kumara and F. Sourjah</p> <p>Tahun : 2022</p> <p>Penerbit : International Conference for Emerging Technology (INCET)</p>	<p>Objek : Teknik Sentiment Analysis</p> <p>Permasalahan: Bagaimana memilih teknik SA yang paling relevant?</p>	Literature Review	Rekomendasi pendekatan baru yang menggabungkan teknik leksikon dan pembelajaran mesin.	<
6.	<p>Judul : Sentiment Analysis for E-Commerce Reviews Based</p>	<p>Objek : Review Produk pada E-Commerce</p> <p>Permasalahan:</p>	CNN RNN Algoritma MWVH	Hasil eksperimen menunjukkan bahwa model hibrida dua saluran yang diusulkan	=

	<p>on Deep Learning Hybrid Model [2]</p> <p>Penulis : Chenyue Wang, Xiaodong Zhu, and Lirong Yan</p> <p>Tahun : 2022</p> <p>Penerbit : International Conference on Signal Processing and Machine Learning</p>	<p>Bagaimana menganalisa review yang banyak dari sebuah platform yang dapat meningkatkan kepuasan pelanggan?</p>		<p>(MWVH1.0 dan MWVH2.0) berdasarkan fitur embedding dari beberapa kata dapat meningkatkan kinerja klasifikasi. Ketika jumlah data sedikit, model MWVH1.0 lebih baik, dengan akurasi rata-rata sebesar 94,62%. Namun, model MWVH2.0 tampil lebih baik pada kumpulan data yang lebih besar, dengan akurasi rata-rata sebesar 95,14%.</p>	
7.	<p>Judul : SFNN: Semantic Features Fusion Neural Network for Multimodal Sentiment Analysis [7]</p> <p>Penulis : W. Wu, Y.</p>	<p>Objek : Sentiment analysis pada aplikasi online</p> <p>Permasalahan: Bagaimana mendeteksi sentimen dalam ulasan online secara efektif untuk mendapatkan pemodelan</p>	<p>SFNN SEMANTIC FEATURE FUSION</p>	<p>Akurasi rata-rata mencapai 62,80%. Data eksperimental dari lima kota besar di Amerika Serikat menunjukkan bahwa model SFNN kami lebih baik daripada</p>	<

	<p>Wang, S. Xu and K. Yan</p> <p>Tahun : 2020</p> <p>Penerbit : International Conference on Automation, Control and Robotics Engineering (CACRE)</p>	<p>preferensi pengguna, pemantauan perilaku konsumen, dan analisis opini publik?</p>		<p>baseline multi-modal dalam analisis sentimen</p>	
8.	<p>Judul : Comparative Study between Traditional Machine Learning and Deep Learning Approaches for Text Classification [8]</p> <p>Penulis: Narayana Kamath, Cannannore Nidhi & Bukhari, Syed & Dengel, Andreas.</p> <p>Tahun : 2018</p>	<p>Objek: Klasifikasi dokumen dengan text mining</p> <p>Permasalahan: Bagaimana membangun sistem yang mampu mengklasifikasikan dokumen berdasarkan sentiment analisis?</p>	<p>Naive Bayes, Logistic Regression, Support Vector Machine, Random Forest Classifier and Multi-Layer Perceptron</p>	<p>Sistem ini mengklasifikasikan data dengan lebih cepat dan dengan akurasi yang lebih baik daripada jika dilakukan secara manual. Hasilnya dibahas dalam bagian hasil dan evaluasi.</p>	<

	<p>Penerbit : ACM Symposium on Document Engineering</p>				
9.	<p>Judul : Sentiment analysis on E- Marketplace User Opinions Using Lexicon- Based and Naïve Bayes Model [10] Penulis : S. Juanita, K. Adiyarta and M. Syafrollah Tahun : 2022 Penerbit : International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)</p>	<p>Objek : Opini public terhadap marketplace melalui layanan costumer service Permasalahan: Bagaimana mengubah opini pengguna yang massif terhadap sebuah e-commerce?</p>	Lexicon Naïve Bayes	Hasil penelitian menunjukkan bahwa kombinasi metode leksikon TextBlob dan metode klasifikasi Naïve Bayes Multinomial mengungguli tiga model lainnya	=
10	<p>Judul : Natural Language Processing on</p>	<p>Objek : Review produk marketplace</p>	Naïve Bayes KNN	Skenario pengujian mendapatkan akurasi rata-rata	<

	<p>Marketplace Product Review Sentiment Analysis [11]</p> <p>Penulis : A. N. Rohman, R. Luviana Musyarofah, E. Utami and S. Raharjo</p> <p>Tahun : 2020</p> <p>Penerbit : International Conference on Cybernetics and Intelligent System (ICORIS)</p>	<p>Permasalahan : Bagaimana meningkatkan hasil analisis sentiment terhadap e-commerce?</p>		<p>Naive Bayes sebesar 52,4% pada dataset Unigram dan akurasi rata-rata K-NN sebesar 79,4% pada dataset Bigram. Pemanfaatan NLP, khususnya normalisasi kata, dapat meningkatkan akurasi sebesar 10% untuk Naive Bayes dan 4% untuk K-NN.</p>	
11	<p>Judul : Sentiment Analysis for Popular e-traveling Sites in Indonesia using Naive Bayes [12]</p> <p>Penulis : T. Sutabri, S. J. Putra, M. R. Effendi, M. N. Gunawan and D.</p>	<p>Objek : Rating produk e-commerce</p> <p>Permasalahan : Pengguna layanan pemesanan dan/atau pembelian online sering menghadapi kesulitan dalam menentukan objek atau layanan yang dipilih yang paling dekat dengan kriteria</p>	<p>Naïve Bayes Probabilistic Value of Corpus</p>	<p>Hasil uji coba menunjukkan tingkat keberhasilan metode dalam menganalisis sentimen sebesar 89%. Hasil analisis sentimen digunakan sebagai acuan dalam perhitungan nilai rating.</p>	<

	<p>Napitupulu</p> <p>Tahun : 2018</p> <p>Penerbit : International Conference on Cyber and IT Service Management (CITSM)</p>	<p>pengguna potensial.</p> <p>Fitur penilaian atau peringkat yang digunakan kadang-kadang tidak objektif dan tidak sesuai dengan konten ulasan yang diberikan oleh para pengulas</p>			
12.	<p>Judul : Sentiment Analysis of Indonesia's National Health Insurance Mobile Application using Naïve Bayes Algorithm [16]</p> <p>Penulis : Natalia Syafitri Kustanto; Intan Nurma Yulita; Indra Sarathan</p> <p>Tahun : 2021</p> <p>Penerbit : International Conference on Artificial</p>	<p>Objek : Sistem asuransi kesehatan</p> <p>Permasalahan : Bagaimana persepsi pengguna terhadap aplikasi Mobile JKN berdasarkan rating?</p>	Naïve Bayes	<p>Akurasi yang diperoleh dari model adalah 93%. Hasil tersebut menunjukkan bahwa versi Mobile JKN 3 dan di atasnya cenderung menerima ulasan negatif daripada positif dari pengguna. Ulasan positif tidak merujuk pada fitur-fitur spesifik aplikasi. Sementara itu, dalam ulasan negatif, banyak pengguna mengeluh tentang</p>	<

	Intelligence and Big Data Analytics			pembaruan aplikasi yang terlalu sering dan masalah saat masuk dan mendaftar yang terkait dengan verifikasi nomor telepon	
13.	<p>Judul : A Comparative Analysis of Sentiment Classification Based on Deep and Traditional Ensemble Machine Learning Models [20]</p> <p>Penulis : Natalia Syafitri Kustanto; Intan Nurma Yulita; Indra Sarathan</p> <p>Tahun : 2021</p> <p>Penerbit : International Conference on Science & Contemporary</p>	<p>Objek : Klasifikasi sentiment analysis</p> <p>Permasalahan : Bagaimana mengevaluasi performa model ensemble deep learning terkini dan model ensemble tradisional untuk mendapatkan akurasi tertinggi dalam klasifikasi sentimen biner?</p>	<p>Voting Ensemble, Bagging Ensemble, Boosting Ensemble, 7-L CNN, GRU, 7-L CNN + GRU + Globe Embedding, 7-L CNN + Long Short-Term memory (LSTM) + Attention Layer</p>	<p>Model ensemble deep learning umumnya memiliki performa yang lebih baik daripada model ensemble tradisional dalam kebanyakan kasus. Pada kedua dataset tersebut, model ensemble deep learning, yaitu 7-L CNN + GRU + Globe dan 7-L CNN + LSTM + Attention Layer, mencapai akurasi tertinggi dengan masing-masing mencapai 94,19% dan 96,37% untuk dataset ulasan</p>	>

	Technologies (ICSCT)			produk dan restoran	
--	-------------------------	--	--	------------------------	--

2.2 Metode yang Diusulkan

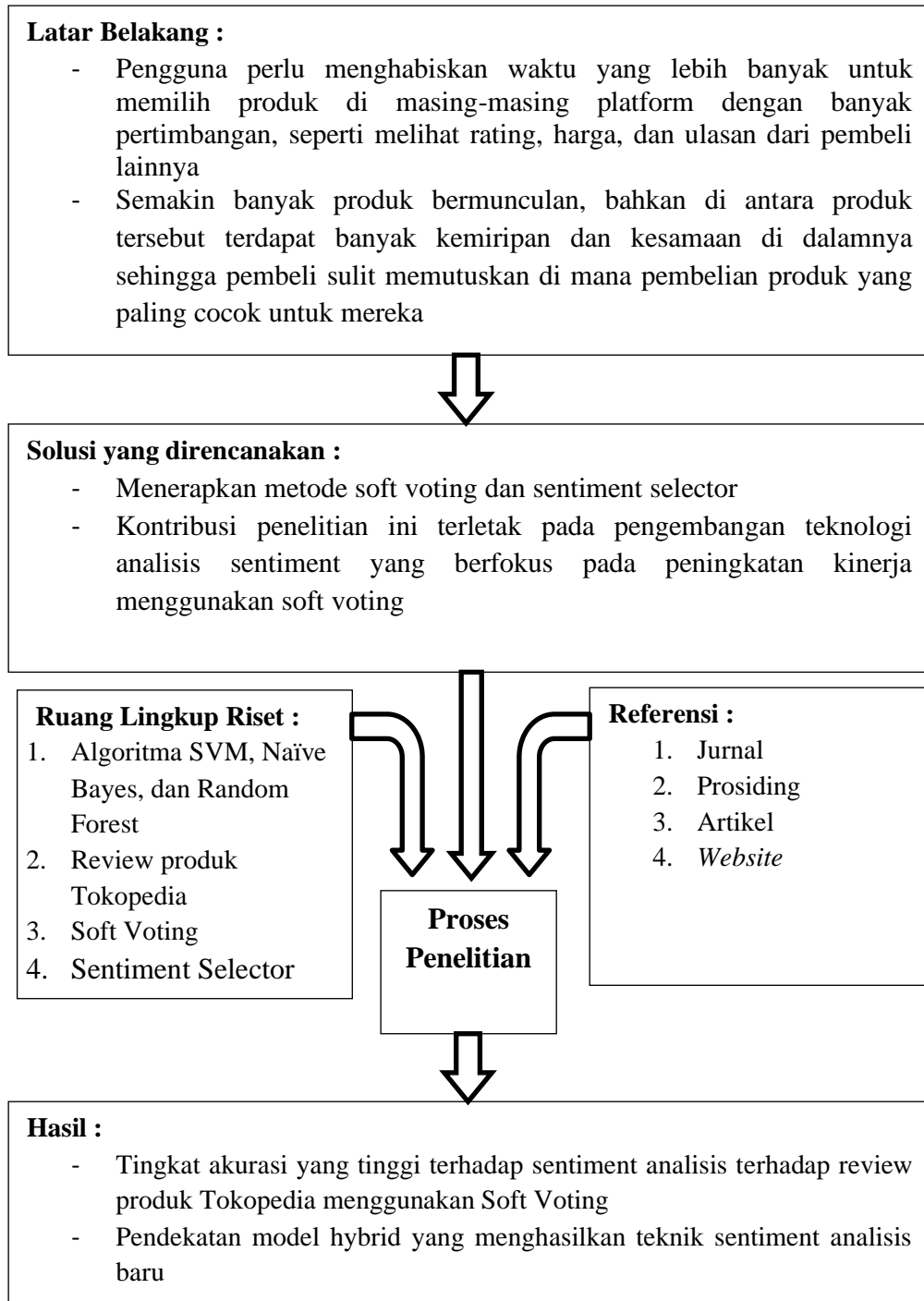
Berdasarkan tabel State of The Art metode yang diusulkan adalah metode klasifikasi hybrid menggunakan model sentiment selector dan model soft voting untuk melakukan proses klasifikasi sentimen pada produk E-Commerce. Algoritma yang digunakan untuk model yang diusulkan adalah Random Forest, dan Support Vector Machine (SVM). Sehingga hasil sentiment analisis lebih akurat. Bahasa pemrograman yang digunakan pada penelitian ini adalah bahasa pemrograman Python. Berdasarkan literatur yang diperoleh Bahasa pemrograman Python lebih baik dibanding Aplikasi Weka dan Matlab karena penggunaannya lebih fleksibel dan merupakan bahasa pemrograman Open Source. Python juga merupakan bahasa pemrograman yang populer digunakan pada data mining.

2.3 Target Hasil Penelitian

Berdasarkan rancangan sistem yang dibuat, penelitian ini diharapkan mampu melakukan peningkatan akurasi terhadap proses klasifikasi secara maksimal yang terdapat dalam seluruh proses sentiment analisis terhadap ulasan pengguna pada E-Commerce. Penelitian ini diharapkan mampu mereduksi lebih dari 60% dari parameter yang ada. Selanjutnya dengan menggunakan parameter terbaik hasil seleksi yang dilakukan diharapkan mampu menunjukkan waktu komputasi yang lebih optimal dan hasil klasifikasi yang maksimal menggunakan metode klasifikasi terbaik dengan hasil persentase klasifikasi diatas 90%.

2.4 Kerangka Pikir

Kerangka pikir dapat dilihat pada gambar di bawah yang menjelaskan



mengenai alur penelitian yang akan dilakukan

Gambar 1. Kerangka Pikir