# BAB I
# PENDAHULUAN

## 1.1. Latar Belakang

Ulasan produk yang dibuat pengguna semakin berperan penting dalam pengambilan keputusan pembelian. Penelitian [1] mencatat bahwa ulasan daring menjadi "semakin tak tergantikan" untuk membantu konsumen membuat keputusan. Penelitian [2] juga menunjukkan konsumen cenderung mencari informasi produk dari ulasan pengguna lain. Survei menunjukkan sekitar 93% konsumen percaya ulasan online memengaruhi pilihan belanja mereka. Dengan kata lain, kebiasaan membaca ulasan sudah menjadi hal umum yang membantu konsumen menilai suatu produk sebelum membeli [2].

Untuk pembaca buku, situs khusus ulasan seperti *Goodreads* sangat populer. *Goodreads* adalah situs berbagi ulasan buku terbesar di dunia, dengan jutaan anggota dan ulasan buku. *Platform* ini berfungsi sebagai jaringan sosial pembaca berbasis buku. Dalam *platform* ini pengguna dapat menilai, merekomendasikan, dan berbagi buku favorit mereka. Dengan *database* ulasan yang luas, *Goodreads* membantu pembaca menemukan buku baru dan mempertimbangkan pembelian berdasarkan pengalaman pembaca lain [1].

Ulasan buku di *platform* seperti *Goodreads* sering kali panjang dan deskriptif. Karena tidak ada format baku, penulis ulasan bebas menjelaskan isi dan opini buku secara rinci. Panjangnya teks ulasan memberikan konteks yang kaya dan detail lebih banyak, memungkinkan pembaca memahami kelebihan dan kekurangan buku secara menyeluruh. Sebagai contoh, penelitian [1] menyatakan volume ulasan yang besar menyediakan konten beragam untuk analisis mendalam. Di sisi lain, kompleksitas bahasa ulasan yang panjang dan adanya ragam gaya penulis menimbulkan tantangan. Ulasan panjang sering bersifat subjektif tinggi, mengandung kosakata Internet, dan dipengaruhi latar budaya penulis. Penelitian [3] menemukan bahwa ulasan buku cenderung bernada positif meski penilaian kualitasnya tidak konsisten, yang menunjukkan ketidakpastian dalam interpretasi sentimen ulasan

Ulasan panjang memberi konteks lebih jelas dan detail tentang isi buku. Penjelasan rincian dalam ulasan yang lebih panjang ini dapat membantu pembaca memperoleh gambaran yang jelas mengenai tema, gaya penulisan, dan argumen buku tersebut. Konten yang kaya ini juga memudahkan analisis aspek tertentu seperti karakter dan alur secara mendalam [3]. Di sisi lain, ulasan panjang memakan waktu lebih lama untuk dibaca dan dipahami. Karena sangat subjektif dan bermuatan opini pribadi, membaca keseluruhan ulasan untuk menangkap ini dari ulasan tersebut bisa menjadi cukup sulit. Selain itu, pada penelitan [4] melaporkan bahwa kualitas ulasan buku online seringkali rendah, meski pengumpulan data secara daring memudahkan mengakses banyak ulasan sekaligus. Kompleksitas bahasa dan ketidakseragaman gaya penulis membuat analisis manual memerlukan usaha dan waktu ekstra. Selain itu, ulasan buku umumnya memiliki bentuk teks tidak terstruktur yang mengandung variasi bahasa, subjektivitas, serta konteks yang kompleks. Oleh karena itu, diperlukan suatu pendekatan otomatis untuk mengekstraksi informasi sentimen dari teks tersebut. Analisis sentimen sebagai bagian dari *Natural Language Processing* (NLP) merupakan pendekatan yang banyak digunakan untuk mengidentifikasi polaritas opini, seperti positif ataupun negatif [5], [6].

Namun demikian, analisis sentimen pada ulasan buku memiliki tingkat kompleksitas yang relatif tinggi. Berbeda dengan ulasan produk yang cenderung singkat dan langsung, ulasan buku sering kali berbentuk narasi panjang yang mengandung deskripsi alur, karakter, serta refleksi personal pembaca. Selain itu, ulasan buku juga dapat mengandung lebih dari satu jenis sentimen dalam satu teks, serta menggunakan gaya bahasa seperti ironi dan metafora [7]. Hal ini menyebabkan model klasifikasi perlu memahami konteks yang lebih luas, sehingga penggunaan satu model saja sering kali belum mampu menghasilkan performa yang optimal.

Sebelumnya, penelitian mengenai analisis sentimen, atau yang disebut sebagai *opinion mining* pernah dilakukan pada penelitian [8], menggunakan data ulasan buku yang berasal dari *Amazon*. Tidak hanya menganalisis sentimen komentar, penelitian ini juga melakukan visualisasi sentimen terhadap data dan opini yang didapatkan dalam data. Sayangnya penelitian ini tidak menjelaskan

metode yang digunakan secara rinci dan hanya menyebutkan IDE yang digunakan yaitu Rstudio IDE 1.1.383. Selain itu, visualisasi yang dilakukan juga tidak menunjukkan evaluasi yang memadai, dan tidak memiliki fitur lain selain menunjukkan visual hasil analisis [8].

Penelitian [9] menggunakan *Naïve bayes* untuk menganalisis sentimen pada ulasan novel populer berbahasa Indonesia dari *Goodreads*. Menurut penelitian ini, Klasifikasi *Naïve bayes* mampu mengklasifikasikan sentimen menjadi kategori positif dan negatif [9]. Namun kesimpulan ini belum didukung dengan evaluasi yang memadai seperti penggunaan *confusion matrix*, *accuracy*, atau skor f1. Selain itu, jumlah data yang digunakan masih sangat terbatas karena buku dan ulasan yang terdapat di *Goodreads* mayoritas menggunakan bahasa Inggris dan hanya sedikit yang menggunakan bahasa Indonesia [9].

Kemudian terdapat juga penelitian [10] yang menggunakan metode Naïve Bayes untuk melakukan klasifikasi data buku dari judul, genre, dan ulasan yang diberikan. Penelitian ini juga memproses hasil penelitian menjadi suatu sistem berbasis web. Hasil dari penggunaan metode naïve bayes dalam penelitian ini mencapai akurasi 72.63%. Meski tingkat akurasi yang dicapai cukup baik, sayangnya, f-*score* yang didapatkan dalam penelitian ini cukup rendah dan hanya mencapai 62.34% [10].

Selain kekurangan yang telah disebutkan, penelitian yang menggunakan model tunggal seperti *Naive Bayes* saja dapat mengasumsikan hubungan linear antar fitur. Akibatnya, model kurang mampu menangkap hubungan non-linear yang kompleks dalam teks, terutama pada ulasan buku yang panjang dan memiliki lebih banyak konteks. Selain itu, pada penelitian mengenai penggunaan *Naive Bayes* untuk melakukan klasifikasi *Big Data,* penulis artikel menyatakan bahwa model *Naive Bayes* memiliki kelemahan dalam mengklasifikasikan sentimen negatif dan netral secara akurat karena ketergantungannya pada distribusi kata [11].

Di sis lain, terdapat metode seperti XGBoost yang memiliki kekuatan dalam mengklasifikasi data dengan hubungan non-linear. Hal ini disimpulkan dari hasil penelitian [12] yang mengklasifikasikan data tanah longsor menggunakan berbagai metode, salah satunya XGBoost. Dalam penelitian ini XGBoost dapat mengklasifikasikan seluruh data dengan label tanah longsor dan mendapatkan

akurasi keseluruhan 82.67%, pada data tanah longsor yang tidak linear. Metode ini diperkenalkan pada penelitian [13] sebagai algoritma *gradient boosting* yang dioptimalkan untuk kecepatan dan performa [13]. Namun demikian, penggunaan satu model saja masih memiliki keterbatasan dalam menangkap keseluruhan karakteristik data yang beragam.

Untuk mengatasi kekurangan dalam penggunaan satu metode ini, dapat digunakan metode *Ensemble Learning*. Dibandingka dengan menggunakan satu algoritma saja, teknik *Ensemble Learning* membuat beragam hipotesis dan mengkombinasikannya untuk menyelesaikan masalah yang diberikan [14]. Dengan kombinasi tersebut, algoritma dapat memiliki kesempatan untuk memanfaatkan kriteria terbaik dari setiap metode dalam *Ensemble* [14]. Salah satu teknik *ensemble* yang banyak digunakan adalah *stacking*, yang memungkinkan kombinasi model dengan karakteristik berbeda melalui mekanisme *Base learner* dan *Meta-Learner* [15]. Pendekatan stacking pertama kali diperkenalkan oleh [15] dan memiliki tujuan awal untuk mendapatkan *generalization accuracy* yang setinggi mungkin. Pada eksperimennya, penelitian ini mendapatkan akurasi sekitar 88.69% pada model terbaiknya. Selain itu, penelitian [16] membandingkan penggunaan model *Stacking ensemble* dan model ensemble learning lainnya, dan hasilnya model *Stacking ensemble* mendapatkan akurasi terbaik mencapai 99.54%.

Dalam pendekatan *Stacking ensemble*, pemilihan model *Base learner* dan *Meta-Learner* merupakan aspek penting yang memengaruhi performa akhir model, model dengan karakteristik yang berbeda akan bisa saling melengkapi dalam *Stacking Ensemble* [15]. Model pertama yang digunakan sebagai *Base learner* adalah *Naive Bayes*. Metode ini banyak dalam klasifikasi teks karena memodelkan distribusi probabilitas kata secara sederhana[17], [18]. *Naive Bayes* bekerja menggunakan asumsi independensi antar fitur, sehingga mampu menangani data berdimensi tinggi dan bersifat *sparse* dengan baik [17]. Dalam melakukan klasifikasi teks menggunakan algoritma tunggal, *Naive Bayes* memiliki performa yang cukup baik ketika digunakan untuk melakukan klasifikasi teks berita yang memiliki lima kelas, dengan akurasi sekitar 92% hingga 98%, mengungguli KNN dengan akurasi 55,8% dan *Decicion Tree* dengan akurasi 82,6% [19]. Namun, ketika digunakan untuk melakukan klasifikasi teks menggunakan data dari buku

populer, akurasi yangg dicapai adalah 72,63% dan F1-*Score* hanya 62,34% [10]. Pada penggunaannya sebagai *Base Learner, Naive Bayes* digunakan sebagai *Base Learner* dalam *Stacking Ensemble* untuk klasifikasi data kanker payudara [20]. Dalam penelitian ini, peneliti menuliskan bahwa *Naive Bayes* dipilih dikarenakn dapat melakukan klasifikasi secara sederhana, dan mendapatkan akurasi klasifikasi 96%, yang selanjutnya digunakan sebagai input *Meta Feature* unruk *Meta Learner* yang digunakan [20].

Model kedua yang digunakan sebagai *Base learner* adalah *XGBoost*, yang merupakan algoritma berbasis *gradient boosting*. Seperti yang telah disebutkan sebelumnya, *XGBoost* dapat digunkan untuk melakukan klasifikasi pada data yang memiliki hubungan non-linear [12]. Berbeda dengan *Naive Bayes* yang bersifat probabilistik dan linear, *XGBoost* membangun model berbasis pohon keputusan secara bertahap untuk memperbaiki kesalahan prediksi sebelumnya, sehingga mampu menangkap interaksi kompleks antar fitur [13]. Pada klasifikasi teks menggunakan data tunggal, *XGBoost* digunakan untuk melakukan klasifikasi pada data cedera liver dikarenakan penggunaan obat-obatan terlarang, dimana diperoleh hasil validasi AUC hingga 0.88 [21]. Meskipun merupakan suatu algoritma *boosting, XGBoost* tetap dapat digunakan dalam *Stacking Ensemble.* Salah satu penelitian menggunakan *XGBoost* sebagai base learner untuk melakukan prediksi *short-term solar PV power*, yang mana peneliti menuliskan bahwa *XGBoost* dipilih karena baik dalam klasifikasi data non linear dan cocok untuk dipasangkan dengan *Meta Learner* yang digunakan yaitu *Ridge Regression* yang mereduksi varian dalam data [22]. Pada artikel ini, peneliti menyatakan bahwa model stcaking dapat bekerja dengan baik dan dapat menolah hipotesis null 95% [22].

Penggunaan kedua model tersebut sebagai *Base learner* didasarkan pada prinsip keberagaman (*diversity*) dalam *ensemble learning*, di mana model dengan karakteristik yang berbeda cenderung menghasilkan kesalahan yang tidak berkorelasi [23]. Dengan demikian, kombinasi *Naive Bayes* yang kuat dalam menangkap distribusi kata dan *XGBoost* yang dapat memodelkan hubungan non-linear diharapkan mampu memberikan representasi prediksi yang lebih akurat.

Selanjutnya, pada tahap *Meta-Learner*, digunakan metode *Maximum Entropy* atau *Logistic regression*. Pemilihan *Maximum Entropy* sebagai *Meta-*

*Learner* didasarkan pada kemampuannya dalam menggabungkan output probabilitas dari beberapa model tanpa mengasumsikan independensi antar fitur [24]. Sebagai model probabilistik, *Maximum Entropy* mampu mempelajari bobot optimal dari prediksi yang dihasilkan oleh masing-masing *Base learner*, sehingga dapat menentukan kontribusi relatif dari setiap model dalam menghasilkan keputusan akhir. Pada penggunaannya sebagai model tunggal, ketika *Maximum Entropy* digunakan dalam klasifikasi ulasan aplikasi dari Google Play, hasil yang didapatkan cukup baik pada akurasi 83% [25]. Dalam Penggunaannya sebagai *Meta Learner* pada *Stacking Ensemble*, penelitian [20] menjelaskan bahwa penggunaan *Maximum Entropy* sebagai *Meta Lerner* adalah untuk menjelaskan hubungan antara variabel dependen biner dan variabel independen tambahan yang berskala interval, ordinal, rasio, maupun nominal. Pada penelitian ini juga didapatkan hasil akurasi deteksi hingga 98% [20]. Selain itu, *Maximum Entropy* merupakan metode yang tidak begitu kompleks dibandingkan model non-linear lainnya, sehingga dapat digunakan sebagai *Meta-Learner* yang berperan sebagai penggabung [26]. Oleh karena itu, penggunaan model linear seperti *Maximum Entropy* menjadi pilihan yang umum karena mampu melakukan generalisasi dengan baik terhadap kombinasi prediksi dari *Base learner*.

Berdasarkan uraian tersebut, tujuan dari penelitian ini adalah untuk membangun dan mengevaluasi model *Stacking ensemble* yang menggabungkan *Naive Bayes* dan *XGBoost* sebagai *Base learner* serta *Maximum Entropy* sebagai *meta learner* dalam analisis sentimen terhadap ulasan buku pada *Platform Goodreads*. Selain itu, penelitian ini juga bertujuan untuk mengimplementasikan model yang diusulkan ke dalam sebuah sistem berbasis *Dashboard* yang memungkinkan pengguna untuk melakukan klasifikasi sentimen ulasan secara otomatis (positif atau negatif) sebagai hasil dari penelitian.

### 1.2. Rumusan Masalah

Berdasar dari latar belakang yang diuraikan sebelumnya, penulis mengajukan pertanyaan dalam penelitian ini yang dapat dijabarkan sebagai berikut:

1. Bagaimanakah proses pengambilan data untuk analisis sentimen menggunakan model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads*?
2. Bagaimana *preprocessing* yang diperlukan sebelum melakukan analisis sentimen menggunakan model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads*?
3. Bagaimana proses klasifikasi teks menggunakan model *Stacking ensemble* dengan Metode *Naive bayes* dan *XGBoost* sebagai *Base learner*, serta *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads*?
4. Bagaimana hasil analisis sentimen menggunakan model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads*?
5. Bagaimana *Dashboard* yang menunjukkan analisis sentimen menggunakan model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads*?

### 1.3. Batasan Masalah

Dalam penelitian ini, diberikan batasan masalah untuk menjaga fokus penelitian. Batasan masalah yang diterapkan pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini merupakan ulasan buku yang diperoleh dari *platform Goodreads*.
2. Data yang digunakan terbatas pada teks ulasan dalam satu bahasa tertentu yaitu bahasa inggris, dan tidak mencakup analisis lintas bahasa.
3. Metode klasifikasi yang digunakan dalam penelitian ini adalah pendekatan *Stacking ensemble* dengan *Naive Bayes* dan *XGBoost* sebagai *Base learner*, serta *Maximum Entropy (Logistic regression)* sebagai *Meta-Learner.*
4. Evaluasi performa model dilakukan menggunakan metrik evaluasi standar klasifikasi, seperti akurasi, presisi, *recall*, dan *F1-score*.
5. Penelitian ini tidak mempertimbangkan aspek lain di luar teks, seperti metadata pengguna, rating numerik, atau informasi tambahan lainnya

## 1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan sebelumnya, adapun tujuan dalam penelitian ini adalah sebagai berikut:

1. Melakukan proses *scrapping* data untuk mengumpulkan data yan dibutuhkan dalam proses analisis sentimen pada data ulasan buku dari *Platform Goodreads*.
2. Menerapkan langkah *preprocessing* yang dibutuhkan sebelum analisis sentimen menggunakan model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads*.
3. Melakukan analisis sentimen model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads*.
4. Melakukan evaluasi hasil analisis sentimen menggunakan model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads.*
5. Membuat *Dashboard* yang menunjukkan analisis sentimen menggunakan metode model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku dari *Platform Goodreads.*

## 1.5. Manfaat Penelitian

Melalui penelitian ini, diharapkan terdapat manfaat bagi pihak-pihak yang membutuhkan, baik secara teoritis maupun praktis, diantaranya:

1. Manfaat Teoritis

   Terdapat beberapa manfaat teoritis yang dapat diberikan penelitian ini bagi dunia akademik. Pertama, penelitian ini dapat mengambil peran dalam mengembangkan pemahaman mengenai metode *Maximum Entropy* ketika berperan sebagau *Meta-Learner* yang digunakan untuk melakukan analisis sentimen pada data yang memiliki ukuran item lebih besar. Selanjutnya, analisis sentimen yang berfokus pada *Review* buku diharapkan dapat memperluas pandangan mengenai bagaimana metode *Maximum Entropy* yang digunakan sebagai *Meta-Learner* dalam menangkap sentimen, pandangan, serta opini pembaca yang dituliskan dalam *Review* buku, sehingga penelitian ini dapat menjadi landasan bagi pengembangan analisis sentimen terhadap tanggapan dan opini pembaca.

2. Manfaat Praktis
    a) Untuk penulis, supaya dapat memperdalam pemahaman dalam bidang analisis data, analisis sentimen, serta interpretasi hasil penelitian, khususnya pada model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku *Goodreads*.
    b) Bagi peneliti selanjutnya, untuk dapat memanfaatkan hasil penelitian ini, baik dalam bidang analisis data maupun analisis sentimen, terutama pada model *Stacking ensemble* dengan *Maximum Entropy* sebagai *Meta-Learner* pada ulasan buku *Goodreads*.