

BAB I

PENDAHULUAN

1.1 Latar Belakang

Wabah Covid-19 yang berlangsung pada akhir tahun 2019 mengharuskan pemerintah untuk bertindak cepat. Keputusan yang diambil oleh banyak negara salah satunya adalah penerapan kebijakan pembatasan sosial yang di Indonesia disebut sebagai Pembatasan Sosial Berskala Besar (PSBB) dan dimuat dalam PP Nomor 21 Tahun 2020 [1]. Kebijakan tersebut mengakibatkan banyak perusahaan wajib menerapkan prinsip *work from home* atau metode bekerja dari rumah. Meskipun tujuan dari kebijakan tersebut adalah untuk memutus rantai Covid-19 di Indonesia, tetapi mengakibatkan banyak karyawan harus di-PHK karena pendapatan perusahaan yang juga berkurang secara drastis. Tercatat bahwa tingkat pengangguran di Indonesia selama Pandemi Covid-19 meningkat dari 4,9% menjadi 7% atau setara dengan 9,7 juta orang [2]. Sebenarnya, masyarakat yang terkena dampak dari Covid-19 tersebut tidak hanya terbatas pada 9,7 juta orang tersebut. Besar kemungkinan bahwa masyarakat yang masih memiliki pekerjaan juga mendapatkan pengurangan pendapatan atau pemasukan bagi mereka yang memiliki bisnisnya sendiri. Peristiwa tersebut mendorong masyarakat untuk mencari cara lain agar mendapatkan penghasilan tambahan. Salah satu cara yang paling digemari saat ini khususnya di Indonesia adalah dengan melakukan investasi.

Investasi adalah sebuah kegiatan menanamkan modal dengan tujuan untuk mendapatkan keuntungan [3]. Kegiatan investasi semakin gencar dilakukan oleh masyarakat tanpa terkecuali masyarakat Indonesia. Kementerian Investasi/Badan Koordinasi Penanaman Modal (BKPM) mencatat bahwa jumlah investasi di Indonesia selama periode Januari – Maret tahun 2022 mencapai Rp 282,4 Triliun. Jumlah tersebut meningkat sebesar 28,5% jika dibandingkan dengan tahun 2021 pada periode yang sama [4]. Hal tersebut menandakan bahwa jumlah investasi masyarakat Indonesia setelah wabah Covid-19 mereda meningkat dengan pesat.

Masyarakat zaman milenial ini melakukan investasi dengan menggunakan aplikasi. Salah satu perusahaan yang menyajikan layanan investasi di Indonesia adalah Bibit. Bibit adalah kepanjangan dari PT. Bibit Tumbuh Bersama yang merupakan sebuah perusahaan *start-up* dan bergerak di bidang layanan jual beli *marketplace* reksa dana. Bibit diluncurkan pada tahun 2019 sebagai bagian dari *startup* investasi Stockbit yang sudah didirikan sejak tahun 2013. Hingga saat ini, Bibit sudah menjadi aplikasi yang diunduh lebih dari 5 juta kali dengan total *rating* 4.8 dari 5 sebagai nilai maksimalnya dan total *review* pengguna sebanyak kurang lebih 202 ribu ulasan. Berbeda dengan Stockbit yang memiliki produk investasi saham, Bibit mengkhususkan produknya pada investasi jenis reksa dana. Beberapa produk reksa dana yang ditawarkan oleh Bibit adalah Reksa dana Pasar Uang, Obligasi, dan Saham yang semuanya dikelola oleh manajer investasi. Pengelolaan dana pada produk reksa dana diatur oleh manajer investasi. Oleh karena itu, produk reksa dana sangat digemari oleh masyarakat Indonesia khususnya kalangan remaja yang belum memiliki pengetahuan yang banyak mengenai dunia investasi.

Alasan pemilihan aplikasi Bibit didasari oleh beberapa hal. Pertama adalah popularitas dari aplikasi itu sendiri. Menurut sebuah survei yang dilakukan oleh Katadata Insight Center (KIC), sebanyak 71,9% responden pada survei tersebut telah menggunakan Bibit sebagai aplikasi investasi yang mereka gunakan [5]. Oleh karena itu, Bibit sudah menjadi salah satu aplikasi yang paling populer digunakan masyarakat Indonesia untuk melakukan investasi khususnya dalam bentuk reksa dana. Kedua, adanya fitur utama yang ditawarkan oleh aplikasi Bibit, yaitu Robo Advisor. Robo Advisor akan membuatkan profil risiko sesuai dengan tujuan awal investasi dari pengguna sehingga jika pengguna menggunakan fitur tersebut, pengguna hanya perlu terus berinvestasi dan pembelian reksa dana akan dilakukan secara otomatis oleh Robo Advisor [6].

Sebagai sebuah aplikasi yang dibuat oleh manusia, tentu saja aplikasi Bibit memiliki keunggulan dan kelemahannya tersendiri. Salah satu cara terbaik untuk mengetahui kelemahan dan kelebihan dari sebuah aplikasi tersebut adalah dengan melihat ulasan atau *rating* dari pengguna. Pada aplikasi *mobile*, salah satu cara untuk melihat *rating* tersebut adalah dengan menggunakan Google Play Store.

Namun, terkadang melihat dari *review* atau *rating* pengguna terasa kurang efisien apabila data yang dimiliki sangat banyak karena calon pengguna pada umumnya hanya melihat *rating* secara keseluruhan tanpa mempertimbangkan lebih lanjut mengenai isi dari ulasan itu sendiri. Oleh karena itu, data ulasan dan *rating* tersebut memerlukan proses lebih lanjut yaitu dengan metode analisis sentimen. Pemilihan Google Play Store sebagai sumber data didasari oleh beberapa hal. Pertama, data yang diperoleh jauh lebih banyak karena seluruh ulasan terhadap aplikasi Bibit sudah terkumpul dalam satu tempat. Selanjutnya, ulasan terhadap aplikasi Bibit diperoleh dari masyarakat yang sudah menggunakan aplikasi Bibit itu sendiri sehingga ulasan yang diberikan juga lebih akurat jika dibandingkan dengan sumber data dari sosial media. Berikutnya, sumber data dari Google Play Store terbuka dan dapat diambil secara gratis tanpa syarat-syarat tertentu.

Analisis sentimen sendiri dapat diartikan sebagai sebuah proses mengeksplorasi data yang disimpan di internet untuk mengidentifikasi dan mengkategorisasi pandangan-pandangan yang diekspresikan dalam bentuk teks [7]. Hasil dari analisis sentimen dapat digunakan untuk melihat gambaran mengenai pendapat pengguna terhadap topik tertentu misalnya film, produk, aplikasi, makanan, sistem, dan lain-lain. Pada aplikasi Bibit, analisis sentimen yang dilakukan ini bertujuan untuk melihat pendapat pengguna aplikasi Bibit apakah itu positif atau negatif. Melalui metode analisis sentimen yang dilakukan terhadap aplikasi Bibit maka perkembangan ulasan aplikasi dari tahun ke tahun berikutnya dapat dianalisis lebih lanjut. Misalnya apabila perbandingan jumlah *rating* positif dan negatif pada tahun 2021 lebih banyak dari tahun 2020 maka dapat disimpulkan bahwa pembaruan-pembaruan aplikasi yang dilakukan pada jangka waktu tersebut sukses membuat aplikasi Bibit menjadi lebih diminati oleh masyarakat. Selain itu, dengan melakukan analisis sentimen maka kredibilitas aplikasi ini dapat terlihat berdasarkan dari ulasan-ulasan pengguna sebelumnya sehingga calon pengguna baru yang ingin memakai aplikasi ini dapat meminimalisir risiko bahwa aplikasi yang digunakannya adalah aplikasi penipuan.

Algoritma yang dipilih untuk melakukan analisis sentimen kali ini adalah algoritma *Naïve Bayes (NB)*, *Decision Tree (DT)*, *Support Vector Machine (SVM)*,

dan *Random Forest (RF)*. Keempat algoritma ini sudah menjadi algoritma yang sangat populer dan digunakan di kalangan peneliti untuk melakukan berbagai macam klasifikasi termasuk di dalamnya adalah analisis sentimen. Cara kerja algoritma *DT* adalah dengan membuat sebuah kerangka seperti akar pohon yang disebut sebagai pohon keputusan dengan mempertimbangkan berbagai faktor yang ada [8]. Di sisi lain, algoritma *NB* memiliki cara kerja dengan melakukan prediksi di masa mendatang berdasarkan dengan data dari masa lalu [9]. *SVM* bekerja dengan cara melihat data dengan titik ekstrim atau *outlier* dari kelas yang berbeda dan kemudian membentuk garis batas atau pembagi untuk memisahkan kelas-kelas yang ada [10]. *RF* termasuk ke dalam *ensemble learning* yaitu sebuah algoritma *machine learning* yang bertujuan untuk meningkatkan akurasi dengan menggabungkan beberapa model untuk menyelesaikan suatu masalah [10]. Algoritma *RF* bekerja dengan cara membuat beberapa pohon keputusan seperti *DT* namun dalam jumlah yang nantinya akan digabung untuk melihat pembagian mayoritasnya.

Sebelumnya tema mengenai analisis sentimen juga sudah banyak dilakukan pada penelitian-penelitian lain. Terdapat 10 artikel jurnal yang diambil sebagai contoh karena memiliki kemiripan tema dengan penelitian kali ini. Penelitian yang pertama membahas mengenai sentimen masyarakat Kanada terhadap kebijakan *social distancing* menggunakan algoritma *SVM*. Model tersebut mendapatkan akurasi sebesar 87% [11]. Penelitian kedua membahas mengenai sentimen masyarakat terhadap kota Bali. Algoritma yang digunakan adalah *SVM*, *K-Nearest Neighbor (KNN)*, *NB*, dan *DT* dengan hasil terbaik diperoleh algoritma *SVM* dengan nilai *AUC* 0,805 [12].

Penelitian ketiga menganalisis sentimen pada aplikasi Bukalapak dengan menggunakan algoritma *SVM* dan *NB*. Akurasi tertinggi diperoleh algoritma *SVM* dengan parameter *max feature TF-IDF* 25% [13]. Penelitian keempat membahas mengenai sentimen penonton terhadap tayangan televisi Indonesian Lawyers Club dengan menggunakan algoritma *KNN*, *NB*, dan *DT*. Akurasi tertinggi diperoleh algoritma *KNN* sebesar 76,94% namun performa tersebut tidak konsisten melainkan berbeda-beda setiap tahunnya [14].

Penelitian kelima membahas analisis sentimen yaitu ulasan terhadap layanan hotel dengan menggunakan *feature selection*. Algoritma yang digunakan adalah *NB*, *DT*, dan *SVM*. Hasil yang diperoleh menyatakan bahwa kombinasi antara *DT* dengan *feature selection* mendapatkan akurasi terbaik yaitu 88,54% [15]. Penelitian keenam membahas mengenai perbandingan performa algoritma *Multinomial NB* dengan *DT* yang ditambah dengan AdaBoost pada ulasan aplikasi PeduliLindungi. Hasil menyatakan bahwa pada kedua algoritma, AdaBoost mampu meningkatkan akurasi model [16].

Penelitian ketujuh melakukan analisis sentimen terhadap protes petani India dengan menggunakan data Twitter. Algoritma yang digunakan pada penelitian tersebut adalah *NB*, *SVM*, *DT*, dan *RF* dengan penambahan *feature extraction TF-IDF* dan *Bag of Words*. Kombinasi antara *feature extraction Bag of Words* dan algoritma *RF* menunjukkan performa terbaik diantara algoritma lainnya [17]. Penelitian kedelapan membahas penggunaan *Information Gain* pada analisis sentimen ulasan film dengan menggunakan algoritma *SVM*, *NB*, dan *KNN*. Pada *dataset* Cornell akurasi *SVM* dengan *Information Gain* memperoleh akurasi 85,65%. Sementara itu, pada *dataset* Stanford akurasi *SVM* dengan *Information Gain* memperoleh akurasi 86,62% [18].

Penelitian kesembilan membahas mengenai prediksi harga pasar saham dengan menggunakan algoritma *SVM*. Diperoleh hasil akurasi yang baik sebesar 89,93% [19]. Penelitian kesepuluh melakukan penelitian terhadap sentimen pengguna Allo Bank menggunakan algoritma *SVM* dan *NB*. Algoritma terbaik adalah *SVM* dengan akurasi sebesar 94,29% [20].

Penelitian kali ini memiliki perbedaan apabila dibandingkan dengan penelitian-penelitian sebelumnya karena beberapa hal. Pertama, objek yang diteliti pada penelitian kali ini adalah aplikasi Bibit sebagai sebuah aplikasi investasi yang cukup populer di kalangan masyarakat Indonesia. Kedua, jangka waktu yang dilakukan untuk penelitian berbeda sehingga dataset dan hasil yang diperoleh juga pasti memiliki perbedaan. Perbedaan ketiga terdapat pada teknik yang dilakukan mengingat penelitian ini akan menambahkan teknik *bootstrapping* untuk

menyeimbangkan kelas data dengan sentimen positif dan negatif pada data *training*. Terakhir, kebanyakan sumber data yang diambil untuk melakukan analisis sentimen masih menggunakan Twitter dan baru beberapa saja yang menggunakan ulasan dari Google Play Store sehingga keberadaan penelitian ini diharapkan dapat menambah referensi penelitian dengan sumber data dari Google Play Store.

1.2 Rumusan Masalah

Berdasarkan penjelasan dari latar belakang sebelumnya, berikut ini adalah beberapa rumusan masalah pada penelitian kali ini:

1. Bagaimana penilaian masyarakat terhadap aplikasi Bibit secara umum?
2. Bagaimana perkembangan sentimen masyarakat terhadap aplikasi Bibit pada tahun 2020 sampai 2022?
3. Bagaimana performa dan perbandingan dari hasil analisis sentimen antara algoritma *DT*, *NB*, *SVM*, dan *RF*?
4. Apakah teknik *bootstrapping* dan optimasi yang dilakukan terhadap model akhir menggunakan *GridSearchCV* dapat meningkatkan akurasi?

1.3 Batasan Masalah

Berikut ini adalah batasan-batasan masalah yang akan ditetapkan agar menjaga penelitian tetap fokus pada inti masalah:

1. Data diperoleh melalui *rating* pada aplikasi Google Play Store yang ditujukan pada aplikasi Bibit.
2. Data yang akan digunakan hanya ulasan yang menggunakan Bahasa Indonesia.
3. Data dikumpulkan dari tahun 2020 hingga 2022.
4. Proses analisis sentimen akan dilakukan dengan menggunakan bahasa pemrograman Python.
5. Algoritma yang akan digunakan hanya *NB*, *DT*, *SVM*, dan *RF*.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan yang dimiliki berdasarkan rumusan masalah yang sudah dipaparkan sebelumnya adalah sebagai berikut:

1. Menggambarkan penilaian masyarakat terhadap aplikasi Bibit.
2. Melihat perkembangan sentimen masyarakat terhadap aplikasi Bibit mulai dari tahun 2020 hingga 2022.
3. Membandingkan performa dan akurasi untuk melakukan analisis sentimen dari algoritma *NB*, *DT*, *SVM*, dan *RF*.
4. Menguji kemampuan teknik *bootstrapping* dan *GridSearchCV* dalam mengoptimasi model.

1.4.2 Manfaat Penelitian

Manfaat yang didapatkan melalui penelitian ini antara lain:

1. Menyajikan informasi terkait sentimen masyarakat terhadap aplikasi Bibit di Google Play Store.
2. Memberikan gambaran mengenai penggunaan algoritma *NB*, *DT*, *SVM*, dan *RF* melalui analisis sentimen pada ulasan aplikasi Bibit.
3. Mengetahui apakah teknik *bootstrap* dan *GridSearchCV* dapat meningkatkan akurasi model.
4. Menjadi referensi bagi penelitian-penelitian dengan topik analisis sentimen di masa mendatang.

1.5 Sistematika Penulisan

1. BAB I PENDAHULUAN

Bab ini berisi latar belakang yang menginformasikan gambaran umum mengenai penelitian ini, rumusan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan yang akan memberikan gambaran mengenai konten dari laporan skripsi ini.

2. BAB II LANDASAN TEORI

Bagian ini berisikan mengenai teori yang berkaitan dengan topik penelitian ini, *framework-framework* pendukung penelitian, *tools* yang digunakan untuk menunjang penelitian, serta penelitian-penelitian yang sudah dilakukan sebelumnya.

3. BAB III METODOLOGI PENELITIAN

Bab ini akan membahas gambaran penelitian, metode yang digunakan dalam bentuk diagram alur, teknik pada proses pengumpulan data, variabel penelitian, dan teknik yang digunakan untuk menganalisis data.

4. BAB IV ANALISIS DAN HASIL PENELITIAN

Bab ini berisikan proses dan hasil penelitian sesuai dengan alur *framework* yang digunakan.

5. BAB V SIMPULAN DAN SARAN

Bab ini berisi simpulan akhir dari penelitian ini beserta dengan saran yang dapat dipertimbangkan untuk penelitian berikutnya.

