

BAB I PENDAHULUAN

1.1 Latar Belakang

Di Indonesia telah banyak terdapat perusahaan yang menyediakan layanan teknologi finansial. Saat ini banyak masyarakat lebih memilih melakukan pembayaran dengan teknologi finansial atau e-wallet. Aplikasi dompet digital yang mempunyai banyak peminat di Indonesia adalah Aplikasi Dana. Dana berhasil memperoleh peringkat teratas *startup digital payment* yang paling banyak diketahui oleh Masyarakat Indonesia dengan persentase Dana sebesar 99% [1]. Aplikasi Dana yang menyediakan berbagai layanan keuangan seperti pembayaran tagihan, transfer uang, pembayaran berbasis Qris, dan lain sebagainya. Tetapi, dibalik keberhasilan aplikasi dana itu sendiri tentu saja memiliki beberapa tingkat kepuasan tersendiri bagi pengguna sehingga sering sekali pada ulasan aplikasi terdapat ulasan positif, negatif, dan netral yang tidak relevan dengan rating yang di berikan pada *Google Play Store*.

Google Play Store adalah wadah yang menyediakan aplikasi yang di tawarkan resmi yang di kembangkan oleh Google yang memungkinkan penggunaan dalam mencari dan mengunduh aplikasi untuk sistem *Android*. Data ulasan aplikasi Dana akan di peroleh dengan menggunakan teknik *Web Scraping* yaitu proses mengekstraksi data dari halaman website. Data ulasan aplikasi Dana pada *Google Play store* akan dilakukan *scraping* data menggunakan *API Google-Play-Scraping*. *API Google-Play-Scraper* adalah API untuk mengekstraksi data informasi yang dimiliki oleh Aplikasi Dana dan ulasan aplikasi dari *Google Play Store* dengan lebih memudahkan tanpa ketergantungan eksternal [2]. Data hasil *scraping* selanjutnya akan di lakukan *text preprocessing*.

Text preprocessing merupakan proses membersihkan teks yang dilakukan agar data dapat memenuhi persyaratan untuk dieksekusi. *Text Preprocessing* untuk dilakukan dikarenakan keadaan dalam teks dapat mempengaruhi hasil akurasi, data set yang di olah nantinya akan melewati berbagai proses salah satunya labelin data, *tokenizing*, *stopword removal*, *stemming*, dan sebagainya. Dataset yang telah melalui proses *text preprocessing* dapat menjadi sumber data untuk dianalisa dan diklasifikasi menggunakan analisis sentimen. Dikarenakan datanya dapat dikatakan data yang siap untuk di olah untuk mengetahui hasil dari akhir data.

Analisis sentiment atau opinion mining yang merupakan perpaduan dari *data mining* dan *text mining*, sesuatu teknik untuk menganalisa pendapat, sentiment, evaluasi, sikap, penilaian

produk atau topik. Dengan analisis sentimen dapat dipahami apakah data yang didapatkan bersifat positif, negatif, dan netral. Pada *google play store* DANA pada tahun 2023 sudah menyentuh angka 100 juta orang sudah melakukan unduhan. Pada penelitian ini akan menganalisis aplikasi DANA dikarenakan e-wallet ini sudah banyak digunakan. Dengan melakukan analisis sentimen mengenai sentimen, dan emosi terhadap suatu produk untuk dapat melihat feedback dari user terhadap suatu layanan e-wallet dana diperlukan analisis sentimen terhadap suatu ulasan user yang masuk pada *google play store* yang terlihat tidak relevan dengan rating yang di dapatkan.

Penggumpulan data diambil menggunakan *Web Scrapping* untuk mendapatkan data yang di inginkan. Setelah itu dengan Algoritma Support Vector Machine(SVM) *Rbf Kernel* dan *Random Forest* untuk melakukan preprocessing mengklasifikasikan data. Pada penelitian yang di lakukan Abitdavy Athallah dan rekan- rekan sebelumnya oleh tentang sentimen analisis tentang aplikasi Dana pada ulasan Google playstore menggunakan *Support Vector Machine (SVM)* di peroleh bahwa akurasi sebesar 87,58% presisi sebesar 91,20% dan recall sebesar 90,21%[3]. Selain itu, Fajar Romadoni dan rekan- rekan juga melakukan penelitian pada text mining untuk analisis sentimen pelanggan terhadap layanan uang elektronik menggunakan algoritma *Support Vector Machine* mendapatkan nilai akurasi sebesar 98,7%[4] Dan untuk penelitian menggunakan random forest yang di lakukan oleh Fanka Angelina Larasati tentang analisis sentiment ulasan aplikasi Dana dengan metode *Random forest* juga didapatkan hasil bahwa precision 84%, recall 84%, F1-Score 84% dan accuracy 84% [2].

Meskipun sudah banyak peneliti yang melakukan penelitian mengenai analisis sentiment mengenai dana dengan algoritma *SVM* dan *Random Forest*. Namun, belum ada penelitian yang menggabungkan 2 algoritma ini pada satu penelitian dan melihat perbandingannya. Oleh karena itu, penelitian ini dilakukan untuk melihat perbandingan dari kedua algoritma tersebut untuk melihat dataset dari ulasan pengguna Dana.

Algoritma yang digunakan bertujuan mengklasifikasi sentimen para pengguna terhadap layanan uang elektronik Dana melalui teknik *text mining* dengan data yang di dapatkan dari hasil *scraping* diharapkan setelah mengetahui hasil analisis sentiment dengan menggunakan bahasa pemrograman yang dipilih *python* karena dalam penggunaan dan pengaplikasiannya mudah untuk mengetahui hasil pada kelas positif, negatif dan netral serta menjadi evaluasi meningkatkan kualitas pelayanan pada aplikasi Dana serta sebagai salah satu produk dalam bidang teknologi finansial.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang di sampaikan, maka didapatkan rumusan masalah sebagai berikut :

1. Berapa jumlah baris data dari dataset yang digunakan dan atribut apa saja dipakai ?
2. Metode validasi apa yang digunakan dan parameter apa yang dipilih?
3. Berapa hasil evaluasi yang didapatkan dari metode dan parameter evaluasi yang dipilih?
4. Metode apa saja yang digunakan sebagai pembanding dari evaluasi hasil penelitian ini ?

1.3 Batasan Masalah

Permasalahan dalam penelitian ini akan dibatasi dengan beberapa hal yang tercantum pada *point-point* berikut :

1. Penelitian hanya fokus penelitian terhadap Aplikasi Dana
2. Penelitian ini hanya menggunakan Data yang di *scraping* sendiri pada aplikasi Dana menggunakan metode *web scrapping*.
3. Data yang di gunakan merupakan kumpulan data dari tahun 2018 sampai dengan 2023.
4. Jumlah data set yang digunakan sebanyak 25.000 data ulasan.
5. Algoritma yang di fokuskan dalam penelitian ini hanya *SVM Rbf Kernel* dan *Random forest* Adapun tambahan hanya sebagai pembanding saja.
6. Data yang di olah telah di labelin oleh itu fokus hanya pada ulasan positif , negatif, dan netral yang telah di beri range datanya

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui analisis sentimen data ulasan yang didapatkan dengan hasil *scraping* dengan menggunakan metode *SVM Rbf Kernel* dan *Random Forest*. Selain itu, terdapat beberapa tujuan sebagai berikut :

1. Mengetahui sentiment positif , negatif, dan netral pada data ulasan pengguna Dana.
2. Mengetahui hasil dari algoritma dari setiap metode validasinya
3. Mengetahui hasil akurasi presisi dan recall pada algoritma yang digunakan.

1.5 Manfaat Penelitian

Manfaat dari penelitian karya ilmiah ini adalah :

1. Secara praktis

Hasil penelitian ini dapat membantu mengevaluasi pengguna dari aplikasi dana dalam menilai aplikasi dengan adanya ulasan pengguna pada *Google Play Store*.

2. Secara teoritis

Penelitian ini merupakan sebuah eksplorasi teori-teori yang selama ini di pelajari , serta menambah wawasan , ilmu pengetahuan , dan pengalaman terhadap analisis sentimen *text mining*.Serta mengetahui berapa keberhasilan algoritma *SVM Rbf Kernel* dan metode *Random Forest* pada pengklasifikasian analisis sentimen berupa opini dari pengguna Dana.

1.6 Sistematika Penulisan

Dalam mempermudah pembaca dan mengikuti apa yang akan dipaparkan penelitian ini , maka penulis telah melakukan penyusunan sistematika penulisan yang sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini terdiri dari latar belakang , perumusan masalah, batasan masalah,tujuan penelitian ,manfaat penelitian, dan sistematika penulisan yang digunakan dalam menyusun penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini akan membahas tentang kajian studi literatur dari penelitian dan dasar teori dari penelitian sebelumnya yang berhubungan dengan topik penelitian .

BAB III METODE PENELITIAN

Bab ini membahas tentang objek penelitian ada juga rancangan alur penelitian yang akan di buat untuk penelitian ini serta berisi tentang informasi alat dan bahan dalam mengerjakan penelitian ini .

BAB IV HASIL DAN PEMBAHASAN

Dalam bab ini berisi tentang hasil dari penelitian yang menggunakan data yang di ambi dan berisi pemaparan hasil-hasil dari tahapan penelitian yang telah dilakukan, serta perbedaan dari kedua algoritma yang digunakan juga dibahas pada bab ini.

BAB V PENUTUP

Bab ini penulis menjelaskan tentang kesimpulan hasil dari penelitian yang telah dilakukan serta adanya bagian saran yang berisi untuk memperbaiki adanya kekurangan pada penelitian

