



1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
3. Plagiarisme juga dilarang untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Saat ini perkembangan teknologi digital telah mengubah proses pencarian kerja menjadi lebih praktis melalui aplikasi berbasis mobile. Salah satu aplikasi yang viral di Indonesia dan banyak digunakan ialah JobStreet, yang menyediakan layanan pencarian lowongan, pengiriman lamaran, pengelolaan profil, hingga rekomendasi pekerjaan. Berdasarkan Google Play Store, aplikasi JobStreet memiliki rating sekitar 4,6, lebih dari 426 ribu ulasan, dan telah diunduh lebih dari 10 juta kali. Banyaknya ulasan tersebut menunjukkan tingginya aktivitas pengguna serta menghasilkan data opini berbentuk teks tidak terstruktur yang dapat dianalisis menggunakan text mining dan Natural Language Processing (NLP). Analisis ulasan pengguna penting dilakukan untuk mengetahui tingkat kepuasan, kritik, dan kendala yang dialami pengguna secara lebih sistematis. Oleh karena itu, penelitian ini berfokus pada ulasan pengguna aplikasi JobStreet dengan output berupa sentimen positif dan negatif terhadap pengalaman penggunaan aplikasi[1].

Dikarenakan ulasan pengguna yang begitu banyak di Google Play Store menyebabkan informasi di dalamnya sulit dianalisis secara manual karena setiap ulasan memuat pengalaman, kritik, saran, maupun keluhan terkait performa aplikasi. Pada aplikasi JobStreet, ulasan tidak hanya berisi kepuasan pengguna terhadap kemudahan mencari pekerjaan, tetapi juga mencakup kendala fitur, proses login, pembaruan profil, hingga kesulitan saat melamar kerja. Selain itu, perbedaan jumlah ulasan positif dan negatif dapat menyebabkan ketidakseimbangan data dalam proses klasifikasi sentimen. Kondisi ini berpengaruh terhadap performa



model karena algoritma cenderung lebih fokus mempelajari kelas mayoritas dibandingkan kelas minoritas. Penelitian terdahulu oleh Marzuqi dkk. menganalisis sentimen ulasan JobStreet dengan membandingkan algoritma Naïve Bayes, Support Vector Machine, Logistic Regression, dan Random Forest. Penelitian tersebut menggunakan 5.000 ulasan, menerapkan preprocessing dan TF-IDF, serta menghasilkan akurasi tertinggi pada algoritma Naïve Bayes sebesar 85,19%, sedangkan Random Forest memperoleh akurasi 82,60%[1][2].

Meskipun penelitian sebelumnya telah membahas analisis sentimen ulasan JobStreet, masih terdapat peluang pengembangan pada penanganan ketidakseimbangan data dan penggunaan algoritma klasifikasi yang lebih optimal. Penelitian Marzuqi dkk belum menerapkan SMOTE untuk mengatasi ketidakseimbangan data maupun membandingkan performa XGBoost dan Random Forest setelah proses penyeimbangan data dilakukan. Padahal, beberapa penelitian menunjukkan bahwa kombinasi SMOTE dan XGBoost dapat meningkatkan performa klasifikasi, terutama pada kelas minoritas. Oleh karena itu, penelitian ini menerapkan SMOTE dengan algoritma XGBoost dan Random Forest untuk meningkatkan performa klasifikasi sentimen ulasan aplikasi JobStreet[2][3][4].

Penelitian ini penting karena hasil sentimen dapat menjadi acuan informasi untuk developer jobstreet untuk mengevaluasi kualitas layanan aplikasi berdasarkan pengalaman pengguna secara langsung. Melalui klasifikasi sentimen, pengembang dapat mengetahui fitur yang paling banyak mendapat apresiasi maupun keluhan, seperti proses login, pencarian lowongan, pengiriman lamaran, dan rekomendasi pekerjaan. Selain manfaat praktis tersebut, penelitian ini juga diharapkan berkontribusi secara akademik dalam penerapan metode penanganan

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Plagiarisme juga dilarang untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

4. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



data tidak seimbang pada analisis sentimen berbasis machine learning. Penerapan SMOTE bersama algoritma XGBoost dan Random Forest diharapkan menghasilkan model pengelompokan yang lebih optimal dan seimbang pada setiap golongan sentimen. Selain itu, Berdasarkan observasi awal terhadap 1.000 ulasan aplikasi JobStreet, ditemukan bahwa distribusi sentimen positif lebih dominan dibanding sentimen negatif sehingga berpotensi menyebabkan ketidakseimbangan kelas pada proses klasifikasi. Dengan demikian, penerapan SMOTE bersama algoritma XGBoost dan Random Forest diharapkan mampu menghasilkan klasifikasi sentimen yang lebih seimbang dan akurat[3][4].

Pemilihan aplikasi JobStreet sebagai objek penelitian didasarkan pada perannya sebagai platform pencarian kerja yang digunakan secara luas di Indonesia maupun Asia. Google Play Store dipilih sebagai sumber data karena menyediakan ulasan pengguna dalam jumlah besar, bersifat terbuka, dan memuat pengalaman langsung pengguna terhadap aplikasi. Besarnya jumlah ulasan tersebut menjadikan data relevan untuk dianalisis menggunakan pendekatan text mining dan machine learning karena mampu merepresentasikan pengalaman pengguna secara langsung terhadap kualitas layanan aplikasi. Berdasarkan hal tersebut, penelitian berjudul “Analisis Sentimen Ulasan Aplikasi JobStreet pada Google Play Store Menggunakan SMOTE dengan Algoritma XGBoost dan Random Forest” dipilih untuk mengidentifikasi kecenderungan sentimen pengguna, menangani ketidakseimbangan data, serta mengevaluasi/membandingkan performa algoritma XGBoost dan Random Forest dalam klasifikasi sentimen[1][2].

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Plagiarisme juga dilarang tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



1.2 Rumusan Masalah

Bersandarkan latar belakang tersebut, disimpulkan melahirkan beberapa pertanyaan yaitu:

1. Bagaimana proses analisis sentimen pada ulasan aplikasi JobStreet di Google Play Store?
2. Bagaimana proses penerapan teknik SMOTE dalam mengatasi ketidakseimbangan data pada proses analisis sentimen ulasan aplikasi JobStreet?
3. Bagaimana hasil performa algoritma XGBoost dan Random Forest setelah di SMOTE dalam melakukan klasifikasi sentimen pada ulasan aplikasi JobStreet berdasarkan accuary, precision, recall dan f1-score?

1.3 Batasan Penelitian

Agar penelitian terarah dan tidak melebar, ruang lingkup dibatasi sebagai berikut:

1. Jumlah data yang digunakan dalam penelitian ini dibatasi 9000 ulasan.
2. Proses pengambilan data dilakukan menggunakan teknik scraping tanpa melibatkan data pribadi pengguna.
3. Analisis hanya ada kelas positif dan negatif.
4. Penelitian ini tidak membahas pengembangan aplikasi atau sistem berbasis antarmuka (UI/UX), melainkan hanya berfokus pada analisis data teks.



1.4 Tujuan Penelitian

Adapun Tujuan penelitian ini dirumuskan berdasarkan permasalahan yang telah ditetapkan agar tetap selaras dan mudah dipahami dalam alur penelitian berikut:

1. Menganalisis sentimen ulasan aplikasi JobStreet di Google Play Store melalui tahapan *preprocessing*.
2. Menerapkan teknik oversampling SMOTE untuk mengatasi ketidakseimbangan data pada ulasan aplikasi Jobstreet.
3. Mengevaluasi algoritma XGBoost dan *Random Forest* dalam klasifikasi sentimen ulasan aplikasi JobStreet.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat yang jelas dan terukur, baik secara teoritis maupun praktis:

1. Memberikan kontribusi terhadap pengembangan ilmu pengetahuan di bidang *data mining*, *text mining*, dan *machine learning*, khususnya dalam analisis sentimen menggunakan algoritma XGBoost dan *Random Forest*.
2. Menjadi referensi bagi peneliti selanjutnya dalam mengembangkan penelitian analisis sentimen dengan penerapan teknik oversampling SMOTE pada data ulasan aplikasi digital.
3. Memberikan informasi mengenai performa algoritma XGBoost dan *Random Forest* dalam klasifikasi sentimen sebagai bahan pertimbangan dalam pemilihan metode *machine learning* pada penelitian serupa.



1.6 Sistematika Penelitian

Penulisan penelitian ini disusun secara sistematis untuk memastikan alur pembahasan yang kuat dan runtut, terdapat 5 bab utama yang saling terhubung:

BAB I PENDAHULUAN

Menampilkan latar belakang penelitian yang berkaitan dengan perkembangan platform pencarian kerja berbasis digital, pentingnya peran ulasan pengguna dalam mengevaluasi kualitas layanan aplikasi, serta permasalahan yang muncul dalam proses analisis sentimen pada review aplikasi JobStreet di Google Play Store. Selain itu, pada bab ini juga dijelaskan rumusan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan yang digunakan sebagai pedoman dalam penyusunan skripsi.

BAB II TINJAUAN LITERATUR

Menguraikan berbagai teori dan konsep yang menjadi landasan penelitian, meliputi analisis sentimen, text mining, tahapan preprocessing data teks, metode pembobotan TF-IDF, teknik oversampling SMOTE, algoritma XGBoost, algoritma Random Forest, serta metode evaluasi model klasifikasi. Selain itu, bab ini juga memuat kajian terhadap penelitian terdahulu yang relevan sebagai referensi pendukung dan dasar dalam penyusunan kerangka pemikiran penelitian.

BAB III METODE PENELITIAN

Cara/tahapan metode penelitian yang diterapkan, dimulai dari proses pengumpulan data ulasan aplikasi JobStreet pada Google Play Store, preprocessing data teks, proses pelabelan sentimen, transformasi data menggunakan metode TF-IDF, pembagian dataset, penerapan teknik SMOTE untuk menangani ketidakseimbangan data, pelatihan model menggunakan algoritma XGBoost dan

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
3. Plagiarisme juga dilarang untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Random Forest, hingga tahap evaluasi performa model menggunakan confusion matrix, accuracy, precision, recall, dan F1-score.

BAB IV HASIL PENELITIAN

Bab ini menyajikan hasil penelitian beserta pembahasannya berdasarkan seluruh tahapan yang telah dilakukan. Pembahasan mencakup hasil pengumpulan data ulasan, tahapan preprocessing, distribusi data sebelum dan sesudah penerapan SMOTE, hasil pengujian model menggunakan algoritma XGBoost dan Random Forest, evaluasi performa model, serta analisis perbandingan kinerja kedua algoritma dalam melakukan klasifikasi sentimen pada ulasan aplikasi JobStreet.

BAB V PENUTUP

berupa kesimpulan hasil penelitian berdasarkan rumusan masalah dan tujuan penelitian yang telah ditetapkan. Dan juga saran untuk penelitian selanjutnya agar hasil yang diperoleh dapat menjadi lebih optimal serta memiliki cakupan penerapan yang lebih luas.

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Plagiarisme juga dilarang tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.