

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Perkembangan teknologi informasi dan komunikasi khususnya internet di Indonesia sangatlah pesat, sehingga membuat Indonesia menjadi salah satu negara yang memiliki pengguna internet aktif terbanyak dan terus tumbuh tiap tahunnya. Menurut survei yang dilakukan Asosiasi Penyedia Jasa Internet Indonesia pada kuartal kedua tahun 2022, konten yang paling sering diakses masyarakat Indonesia di internet adalah media sosial. Pada Provinsi Bali khususnya, 71,74% konten yang diakses melalui internet adalah media sosial (APJI, 2022). Hal ini menyebabkan terjadinya perubahan kebiasaan di masyarakat, saat ini masyarakat yang biasa mengonsumsi informasi melalui media cetak mulai beralih ke media internet.

Media sosial adalah sebuah media *online* yang membuat para penggunanya bisa dengan mudah mencari informasi, berpartisipasi, berbagi, dan menciptakan isi (konten) pada sebuah *platform* di internet. Beberapa media sosial yang populer digunakan di Indonesia contohnya seperti Facebook, Instagram, dan Twitter.

Pada umumnya masyarakat Indonesia menggunakan media sosial untuk mengekspresikan diri dengan cara membagikan foto, video, atau ceritanya. Pengguna media sosial juga dapat saling berinteraksi satu sama lain dengan cara saling memberikan *like* atau komentar. Semua buah pikiran yang terkumpul

tersebut akan membentuk suatu opini masyarakat di dalam media sosial. Opini ini bisa saja menjadi penilaian terhadap suatu produk, pelayanan, atau keadaan yang sedang terjadi sehingga bisa dimanfaatkan dengan *sentiment analysis* atau *opinion mining* agar menjadi sebuah informasi.

Sentiment analysis (analisis sentimen) adalah bidang studi yang menganalisis opini, sentimen, emosi, penilaian, dan sikap seseorang terhadap suatu produk atau sebuah kebijakan (Liu Bing, 2015). Analisis sentimen sendiri telah dimanfaatkan oleh pebisnis atau organisasi tertentu untuk mendapatkan opini dari masyarakat. Beberapa penelitian telah menggunakan informasi analisis sentimen untuk memprediksi opini masyarakat terkait suatu produk atau kebijakan. Adilah pada penelitiannya yang berjudul “*Sentiment Analysis of Online Transportation Service using the Naïve Bayes Methods*”, menggunakan analisis sentimen untuk melihat opini publik terhadap layanan transportasi online Gojek. Hasil penelitian ini dapat dijadikan sebagai rekomendasi untuk meningkatkan kinerja transportasi *online* (Tika Adilah et al., 2020). Naf’an pada penelitian yang berjudul “*Sentiment Analysis of Cyberbullying on Instagram User Comments*”, membuat penelitian yang bertujuan untuk membuat sistem yang dapat mengklasifikasikan komentar apakah mengandung unsur *cyberbullying* di Instagram dengan menggunakan analisis sentimen (Naf’an et al., 2019). Neogi pada penelitian yang berjudul “*Sentiment analysis and classification of Indian farmers’ protest using twitter data*”, Dalam penelitiannya, mengumpulkan data dari Twitter tentang protes petani di India untuk memahami sentimen yang dibagikan publik di tingkat internasional dengan menggunakan analisis sentimen (Neogi et al., 2021).

Salah satu perguruan tinggi negeri di Bali yaitu Universitas Pendidikan Ganesha, juga memanfaatkan potensi teknologi media sosial ini sebagai media informasi dan *branding* untuk perguruan tinggi. Jumlah pengguna yang besar dan penambahan yang selalu konsisten meningkat setiap tahunnya membuat media sosial dapat digunakan menjadi *platform* atau sarana untuk melakukan *branding*, diplomasi, dan penyampaian informasi yang lebih efektif. Saat ini Universitas Pendidikan Ganesha memiliki beberapa akun media sosial seperti Facebook, Twitter, Instagram, dan Youtube. Pertumbuhan interaksi paling aktif terjadi pada akun Instagram (@undiksha.bali) yang saat ini (Mei 2022) memiliki jumlah pengikut sebesar 31,6 ribu orang. Media sosial ini membuat pengguna lain dapat berinteraksi, beropini, dan memberikan penilaian secara terbuka kepada Universitas Pendidikan Ganesha dengan cara memberikan komentar langsung pada akun tersebut.

Dari komentar yang masuk tersebut bisa dimanfaatkan untuk melakukan analisis sentimen. Analisis sentimen penting untuk dilakukan karena mengetahui sentimen yang diberikan publik kepada suatu lembaga, dapat membantu memberikan gambaran penilaian masyarakat apakah memberikan respon yang positif, netral, atau negatif terhadap kinerja atau kebijakan dari lembaga tersebut. Hasil analisis tersebut dapat digunakan sebagai acuan dalam melakukan strategi promosi, pengambilan keputusan, atau peningkatan layanan universitas. Tetapi saat ini fitur bisnis dari media sosial Instagram belum menyediakan analisis sentimen.

Salah satu cara untuk melakukan analisis sentimen adalah dengan menggunakan model klasifikasi. Klasifikasi adalah pemberian kategori yang telah didefinisikan ke data. Mengklasifikasikan data merupakan salah satu cara untuk

mengorganisasikan data, data yang memiliki karakteristik atau ciri yang sama akan dikelompokkan ke dalam kategori yang sama untuk mempermudah pencarian informasi.

Metode umum yang digunakan untuk melakukan klasifikasi adalah dengan menggunakan *machine learning*. Salah satu algoritma klasifikasi yang sering digunakan untuk analisis sentimen adalah *Support Vector Machine* (SVM). SVM melakukan klasifikasi dengan cara mencari *hyperplane* (garis pemisah) terbaik untuk memisahkan data ke dalam beberapa kelas sentimen. Keunggulan dari SVM adalah akurasi yang baik terutama untuk klasifikasi untuk dua kelas (Nirmala Dharmapatni & Merawati, 2020; Luqyana et al., 2018; Haryanto et al., 2018)

Tantangan dalam menggunakan SVM adalah ketika menerapkannya pada data yang tidak linear dan lebih dari dua kelas. Dalam melakukan klasifikasi data yang sifatnya tidak linear, SVM memiliki sebuah fungsi yang bernama fungsi kernel atau *kernel trick*. Konsep dari *kernel trick* ini adalah dengan menggunakan ruang dengan dimensi yang lebih tinggi untuk membuat *hyperplane*. Beberapa fungsi kernel yang digunakan di SVM contohnya seperti kernel *Radial Basis Function* (RBF). Pada kernel tersebut terdapat *hyperparameter* yang nilainya berpengaruh untuk menentukan kinerja klasifikasi yang dihasilkan oleh SVM. Kelemahan dari SVM adalah sulitnya menentukan nilai dari *parameter* kernel yang optimal (Samadzadegan et al., 2010; Zhao et al., 2011; Ilhan, 2013; N & Deka, 2014). Oleh sebab itu, direkomendasikan beberapa algoritma *metaheuristic* atau pencarian untuk mencari nilai parameter SVM yang optimal. Beberapa penelitian sebelumnya telah menggunakan metode ini untuk mendapatkan nilai parameter yang lebih optimal. Ramasamy, pada penelitiannya yang berjudul "*Selection of*

Optimal Hyper-parameter Values of Support Vector Machine for Sentiment Analysis Tasks Using Nature-inspired Optimization Methods” menggunakan algoritma pencarian yaitu *Cuckoo Search Optimization (CSO)*, *Ant Lion Optimizer (ALO)*, dan *Polar Bear Optimization (PBO)* untuk mengoptimasi parameter SVM (Ramasamy et al., 2021). Harafani pada penelitiannya yang berjudul “*Support Vector Machine Parameter Optimization to Improve Liver Disease Estimation with Genetic Algorithm*”, menggunakan *Genetic Algorithm (GA)* untuk mengoptimasi klasifikasi SVM untuk melakukan perkiraan terhadap masalah penyakit hati (Harafani, 2020).

Salah satu algoritma pencarian yang direkomendasikan adalah *Genetic Algorithm*. *Genetic Algorithm (GA)* adalah algoritma pencarian yang terinspirasi dari prinsip genetika dan seleksi alam dari teori evolusi (Whitley, 1993). Secara sederhana algoritma ini konsep utamanya adalah individu-individu yang paling unggul akan bertahan hidup (diteruskan), sedangkan individu-individu yang lemah akan dieliminasi. Pada algoritma genetika, suatu individu dievaluasi berdasarkan suatu fungsi tertentu sebagai ukuran nilai kualitasnya (Dr. Suyanto, 2019). Fungsi ini dikenal sebagai *fitness function*. *Fitness* dalam GA didefinisikan sebagai gambaran kelayakan suatu solusi terhadap suatu permasalahan. Misalnya, dalam kasus untuk mencari nilai optimal dari parameter SVM dengan kernel RBF yang membutuhkan dua nilai parameter yaitu C dan γ . Dibentuklah kromosom yang memiliki dua gen (bagian dalam individu di konsep GA) secara acak yang kemudian akan diiterasi sesuai konsep seleksi di GA dan dievaluasi nilai *fitness*-nya pada setiap perulangan untuk mencari nilai parameter terbaik yang memberikan nilai akurasi paling tinggi.

Dari uraian di atas, peneliti tertarik untuk melakukan penelitian dengan judul “*Sentiment Analysis Menggunakan Support Vector Machine dengan Optimasi Genetic Algorithm (Studi Kasus Respon Masyarakat di Media Sosial Universitas Pendidikan Ganesha)*”. Alasan penggunaan *Support Vector Machine (SVM)* yang dikombinasikan dengan *Genetic Algorithm* adalah untuk menguji seberapa besar pengaruh kombinasi algoritma *metaheuristic* (pencarian) terhadap kinerja dari algoritma *classification* (klasifikasi) untuk analisis sentimen. Optimasi dengan *Genetic Algorithm* dilakukan untuk menemukan nilai parameter terbaik untuk kernel SVM-RBF. Penelitian ini mengambil studi kasus dari media sosial milik Universitas Pendidikan Ganesha. *Dataset* yang digunakan berasal dari kolom komentar di akun Instagram Universitas Pendidikan Ganesha (@undiksha.bali). Informasi yang dihasilkan dari penelitian ini diharapkan dapat memberikan gambaran persepsi masyarakat terhadap Universitas Pendidikan Ganesha di media sosial yang mungkin nantinya bisa berguna sebagai bahan evaluasi untuk peningkatan layanan kampus.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang yang telah dikemukakan di atas, maka permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut.

1. Bagaimana rancangan dan implementasi dari model klasifikasi untuk melakukan analisis sentimen terhadap komentar yang masuk di akun Instagram Universitas Pendidikan Ganesha dengan menggunakan metode *Support Vector Machine* dengan optimasi *Genetic Algorithm*?
2. Bagaimana performa dari *Support Vector Machine* setelah dioptimasi dengan *Genetic Algorithm* dalam melakukan klasifikasi analisis sentimen?

1.3 TUJUAN PENELITIAN

Tujuan dari penelitian ini adalah sebagai berikut.

1. Merancang dan mengimplementasikan model klasifikasi analisis sentimen dengan menggunakan metode *Support Vector Machine* dengan optimasi *Genetic Algorithm* untuk studi kasus komentar di akun *Instagram* Universitas Pendidikan Ganesha.
2. Mengetahui performa dari kombinasi *Support Vector Machine* dan *Genetic Algorithm* dalam melakukan analisis sentimen.

1.4 BATASAN PENELITIAN

Pada pengerjaan penelitian ini diberikan batasan-batasan untuk permasalahan yang diteliti antara lain.

1. Data terakhir diambil pada bulan Mei 2022 dari akun media sosial Instagram Universitas Pendidikan Ganesha (@undiksha.bali).
2. Data yang digunakan adalah komentar yang menggunakan bahasa Indonesia yang dalam kalimatnya mengandung sebuah opini.

1.5 MANFAAT PENELITIAN

Manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut.

1. Menghasilkan model klasifikasi yang dapat digunakan untuk melakukan *text classification* khususnya *sentiment analysis* dengan menggunakan metode *Support Vector Machine* yang di optimisasi algoritma genetika.
2. Mengetahui kinerja dari *Support Vector Machine* yang dikombinasikan dengan algoritma genetika dalam melakukan klasifikasi analisis sentimen.

3. Mengetahui efektivitas dan perilaku algoritma genetika dalam pencarian parameter yang optimum pada klasifikasi *Support Vector Machine*.

