

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kesehatan masyarakat telah menjadi fokus utama dalam menghadapi tantangan kesehatan global. Sebagai upaya untuk mencapai masyarakat yang lebih sehat, program hidup sehat telah menjadi pilar yang penting [1]. Tumor adalah pertumbuhan abnormal yang dapat berkembang di mana saja dalam tubuh manusia, dan dapat diklasifikasikan sebagai jinak atau ganas [2]. Tumor otak memiliki dua jenis tumor primer yang tumbuh secara independen dan tumor sekunder yang berasal dari penyebaran sel kanker dari organ lain ke otak. Tumor otak primer terdiri dari beberapa jenis, dengan Glioblastoma menjadi yang paling umum dan agresif.

Dalam sektor kesehatan, khususnya dalam mendiagnosis penyakit otak seperti Glioma, Pituitari, dan Meningioma, interpretasi citra medis oleh manusia sering kali dibatasi oleh subjektivitas dan ketidakpastian [3]. Tumor otak primer yang paling umum ditemukan adalah glioma, dan *American Association of Neurological Surgeons* (Asosiasi Ahli Bedah Saraf Amerika) menyatakan bahwa tumor ini menyumbang sekitar 78% dari total kasus tumor otak, dengan Glioma diklasifikasikan sebagai tumor yang sangat agresif [4].

Machine Learning dapat dimanfaatkan dalam mengklasifikasikan data penyakit infertilitas untuk meningkatkan produktivitas dan konsistensi diagnostik. Banyak penelitian telah menggunakan teknik klasifikasi untuk meningkatkan konsistensi hasil dan mendorong terciptanya taksonomi baru, sehingga memungkinkan teknologi untuk berkembang [5]. Klasifikasi adalah salah satu model dalam penambangan data yang merupakan teknik yang berguna untuk memprediksi data dan mengestimasi nilai dari berbagai set data. Tujuan model klasifikasi adalah untuk memprediksi nilai variabel yang tidak diketahui berdasarkan variabel yang diberikan [6]. Selain faktor teknis, persepsi dan pemahaman publik terhadap sistem berbasis teknologi juga menjadi aspek penting dalam penerapan sistem informasi dan pengambilan keputusan berbasis data [11].

Machine Learning juga dapat digunakan untuk mengklasifikasikan penyakit tumor. Salah satu metode klasifikasi dalam *Machine Learning* yang dapat digunakan adalah *Support Vector Machine* (SVM). *Support Vector Machine* (SVM) merupakan salah satu algoritma *Machine Learning* yang digunakan untuk permasalahan klasifikasi dan regresi dengan cara membangun *hyperplane* optimal untuk memisahkan data ke dalam kelas-kelas tertentu [7][8]. Secara umum, algoritma *Support Vector Machine* mengungguli metode *K-Nearest Neighbor* (KNN), mencapai akurasi hingga 89,70% [9].

Di era informasi saat ini, volume dan kompleksitas data yang tersedia semakin meningkat. Dalam konteks *Machine Learning*, salah satu tantangan utama adalah pemilihan fitur yang relevan. Pemilihan fitur bertujuan untuk memilih subset atribut yang paling berpengaruh terhadap hasil klasifikasi, sehingga dapat meningkatkan kinerja model sekaligus mengurangi kompleksitas komputasi. Kinerja algoritma SVM sangat bergantung pada kualitas fitur yang digunakan, sehingga pemilihan fitur menjadi tahapan yang penting.

Beberapa metode pemilihan fitur yang umum digunakan bersama SVM antara lain Gain Ratio, Chi-Square, dan ReliefF. Gain Ratio merupakan pengembangan dari Information Gain yang bertujuan untuk mengurangi bias terhadap atribut dengan jumlah nilai yang banyak. Chi-Square merupakan metode statistik yang digunakan untuk mengukur hubungan antara fitur dan kelas target [10]. Sementara itu, ReliefF merupakan metode seleksi fitur yang mampu mengevaluasi kontribusi setiap fitur berdasarkan kedekatan antar data sehingga efektif digunakan pada data yang kompleks [12]. Selain metode seleksi fitur berbasis filter, terdapat metode reduksi dimensi lain seperti *Principal Component Analysis* (PCA), namun metode tersebut tidak digunakan dalam tahap eksperimen pada penelitian ini.

Oleh karena itu, penelitian ini difokuskan pada analisis pengaruh penerapan metode seleksi fitur terhadap peningkatan akurasi algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan data tumor primer.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang tersebut, permasalahan yang diselesaikan pada penelitian kali ini adalah menguji sejauh mana efektivitas metode *feature selection* untuk meningkatkan akurasi algoritma *Support Vector Machine* (SVM) pada klasifikasi data tumor primer.

1.3 Tujuan

Penelitian ini bertujuan untuk menganalisis pengaruh penerapan metode *feature selection* terhadap peningkatan akurasi algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan data tumor primer. Melalui penelitian ini diharapkan dapat diketahui bagaimana pemilihan fitur yang tepat mampu meningkatkan kualitas hasil klasifikasi, mengurangi kompleksitas data, serta meningkatkan efisiensi kinerja model.

Selain itu, penelitian ini bertujuan untuk mengidentifikasi metode *feature selection* yang paling efektif dalam meningkatkan performa algoritma SVM, serta membandingkan kinerja model sebelum dan sesudah penerapan seleksi fitur. Dengan adanya perbandingan tersebut, diharapkan dapat diketahui besarnya peningkatan akurasi yang dihasilkan sehingga hasil penelitian dapat menjadi acuan dalam pengembangan sistem klasifikasi tumor yang lebih akurat dan andal.

1.4 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat dalam pengembangan ilmu pengetahuan di bidang *Machine Learning*, khususnya terkait penerapan metode *feature selection* untuk meningkatkan kinerja algoritma *Support Vector Machine* (SVM) dalam klasifikasi tumor

primer, sehingga dapat menjadi referensi bagi penelitian selanjutnya. Selain itu, hasil penelitian ini diharapkan mampu memberikan manfaat praktis bagi dunia kesehatan dengan meningkatkan akurasi dan keandalan sistem pendukung keputusan medis, membantu tenaga kesehatan dalam proses diagnosis yang lebih cepat dan tepat, serta mengurangi potensi kesalahan dalam penanganan pasien.

1.5 Batasan Penelitian

Batasan dalam penelitian ini meliputi:

1. Penelitian ini membahas penerapan metode *feature selection* untuk meningkatkan akurasi algoritma *Support Vector Machine* (SVM) dalam klasifikasi data tumor primer, dan tidak membahas penyakit lain di luar kategori tumor primer.
2. Algoritma klasifikasi yang digunakan terbatas pada *Support Vector Machine* (SVM) tanpa melakukan perbandingan langsung dengan algoritma lain seperti Random Forest, Naïve Bayes, atau Neural Network.
3. Metode *feature selection* yang dikaji dibatasi pada *Gain Ratio*, *Chi-Square*, dan *ReliefF*, tanpa mencakup metode seleksi fitur lainnya.
4. Data yang digunakan dalam penelitian ini berasal dari dataset sekunder, sehingga tidak melibatkan pengambilan data langsung dari rumah sakit atau pemeriksaan pasien secara langsung.
5. Evaluasi kinerja model dibatasi pada pengukuran akurasi, presisi, recall, dan F1-score, tanpa membahas implementasi sistem dalam bentuk aplikasi klinis atau penggunaan secara real-time.

1.6 Keterbaruan

Berdasarkan kajian terhadap penelitian-penelitian sebelumnya, dapat disimpulkan bahwa penelitian ini tidak bertujuan untuk mengembangkan algoritma baru, melainkan berfokus pada analisis komparatif penerapan metode *feature selection* terhadap kinerja algoritma *Support Vector Machine* (SVM). Kontribusi utama dari penelitian ini terletak pada evaluasi sistematis pengaruh tiga metode *feature selection*, yaitu *Gain Ratio*, *Chi-Square*, dan *ReliefF*, terhadap peningkatan akurasi klasifikasi data tumor primer menggunakan dataset yang sama dan skenario pengujian yang seragam. Dengan pendekatan tersebut, penelitian ini diharapkan dapat memberikan gambaran empiris mengenai metode seleksi fitur yang paling efektif untuk meningkatkan performa SVM pada data medis, khususnya klasifikasi tumor primer.