

BAB I

PENDAHULUAN

1.1 Latar Belakang

Stunting adalah masalah penelitian nasional yang masih trending sebagai riset nasional tahun 2020 hingga 2024, Sumber (*Ministry of Research and Technology of the republic of Indonesia, 2020*). Prevalensi balita stunting di Kalimantan Timur tahun 2022 adalah 23,9%. Stunting menjadi masalah kesehatan yang mendesak, karena dampaknya yang berkelanjutan dapat mengganggu perkembangan kognitif, produktivitas, dan kualitas hidup anak di masa dewasa. Oleh karena itu, penggunaan teknologi informasi dan metode klasifikasi data mining menjadi penting dalam upaya untuk memberikan diagnosis dini dan intervensi yang tepat guna, sehingga dapat mencegah dampak yang lebih serius pada pertumbuhan dan perkembangan anak. Salah satu metode *machine learning* dalam klasifikasi atau prediksi stunting adalah SVM (*Khan, 2021*), karena dalam situasi di mana ada perbedaan kelas yang jelas dan ruang berdimensi besar, SVM cocok untuk masalah dengan banyak fitur sehingga performa metode tidak menurun drastis ketika jumlah fitur melebihi jumlah sampel. Pada penelitian terdahulu SVM+PSO menghasilkan akurasi 78% (*Andriyani et al., 2023*). C4.5 menghasilkan akurasi 61,82% (*Yunus et al., 2023*)(*Ula et al., 2022*)(*Anggriawan & Nugroho, 2023*) dan *Naïve Bayes* menghasilkan akurasi 75% dengan data yang berdimensi rendah (*Arumi et al., 2023*)(*V. Herliansyah et al., 2021*).

Adapun masalah *high dimension* berdasarkan penelitian dari (*Huang, 2020; Migoñ, 2021; Rajabi, 2023; Shao, 2021; Wang et al., 2021; Xdqj et al., 2020; Yun et al., 2023*) menurut penelitian tersebut, metode *Relief* adalah algoritma yang baik dalam menangani data berdimensi tinggi. Menurut (*Raj & Mohanasundaram, 2020*) pemilihan fitur menggunakan algoritma *Relief* yaitu *MultiSURF* dapat memperkirakan atribut yang sesuai dari kumpulan data berdimensi tinggi namun banyaknya fitur berlebihan dapat menurunkan kinerja *MultiSURF*. Untuk memilih fitur yang berkualitas dari dataset yang besar, dalam penelitiannya menyarankan, untuk menggunakan algoritma pembobotan fitur inovatif yang disebut *Boundary Margin Relief* (BMR). Konsep utama BMR adalah memprediksi bobot fitur melalui pengukuran *hyperplane* lokal, yang umumnya dilakukan dalam teknik *I-Relief*. Bobot fitur dalam metode tersebut sangat efektif dalam mengeliminasi fitur yang berlebihan. Terbukti dapat meningkatkan akurasi dengan eksperimennya pada data *Leukemia* metode BMR menghasilkan akurasi sebesar 91,2%. Oleh karena itu, penulis ingin menerapkan metode tersebut dalam penelitian data stunting di Kota Samarinda guna membantu dalam proses seleksi fitur.

Selain itu, masalah ketidakseimbangan kelas juga menambah tantangan dalam klasifikasi karena dapat menyebabkan tumpang tindih antara label atau kelas (*Hussein et al., 2023*) sehingga penelitian ini bertujuan untuk meningkatkan kinerja SVM dengan menangani data yang tidak seimbang, menggunakan seleksi fitur dengan metode *Boundary Margin Relief* (BMR) untuk data berdimensi tinggi, dan metode *Simulated Annealing* (SA) untuk meningkatkan akurasi model SVM. SA Terbukti dapat meningkatkan akurasi SVM, berdasarkan hasil eksperimen dari (*Mahareek et al., 2021*), dalam penelitiannya, menguji metode SVM (MLP karnel) dan SA dengan data *Portuguese course* menghasilkan akurasi sebesar 78.35% tanpa SA setelah di optimasi dengan SA menjadi 90.72%. sehingga penulis ingin menerapkan metode *Simulated Annealing* (SA) sebagai metode optimasi untuk meningkatkan akurasi SVM.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada, masalah penelitian ini dirumuskan sebagai berikut:

- a. Apasaja fitur yang mempengaruhi kinerja BMR dan *Support Vector Machine* (SVM)?
- b. Apakah Metode *Simulated Annealing* (SA) dan *Boundary Margin Relief* (BMR) apabila diterapkan terhadap algoritma klasifikasi *Support Vector Machine* mengalami perbaikan performa pada nilai akurasi model tersebut?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

- a. Mengetahui fitur yang mempengaruhi kinerja BMR dan *Support Vector Machine* (SVM)
- b. Mengetahui seberapa meningkat akurasi dengan *confusion matrix* pada metode SVM apabila SA dan BMR diterapkan untuk menangani *High Dimensional* dan *Imbalanced* pada data Stunting Kota Samarinda

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat bagi beberapa pihak yaitu:

- a. Penulis
Hasil penelitian sebagai kontribusi berupa model optimasi klasifikasi baru dalam mengatasi masalah *High Dimensional* dan *Imbalanced* data Stunting Kota Samarinda
- b. Pembaca
Mengembangkan wawasan mengenai model klasifikasi yang dapat mengoptimalkan kinerja SVM dengan metode SA dan kombinasi metode seleksi fitur BMR untuk mengatasi *High Dimension* dan *Imbalanced* pada data stunting.

1.5 Batasan Masalah

Penulis membatasi lingkup penelitian sebagai berikut untuk memastikan bahwa topik yang dibahas tetap terfokus dan tidak terlalu meluas.

- a. Data yang digunakan adalah data Stunting yang diambil dari Dinas Kesehatan Kota Samarinda tahun 2023.
- b. Menggunakan karnel RBF dengan nilai $cost = 10$ dan $gamma = 5$
- c. Menggunakan SMOTE untuk menangani data *imbalance* pada tahap *pre-processing* data
- d. Menggunakan *Boundary Margin Relief* (BMR) sebagai seleksi untuk memilih fitur-fitur yang berpengaruh tinggi dalam klasifikasi model SVM.
- e. Menggunakan *Simulated Annealing* (SA) untuk meningkatkan kinerja algoritma BMR dan SVM.
- f. Menggunakan *Cross-Validation* k-Fold 10 sebagai pembagi data uji dan data latih