

RINGKASAN

Sistem Prediksi Risiko Kanker Serviks Menggunakan Algoritma C4.5 (Studi Kasus: Rekam Medis di RSUP dr. Kariadi Semarang), Siti Nur Aisyah dengan NIM G41221957, Tahun 2025, Manajemen Informasi Kesehatan, Politeknik Negeri Jember, dibawah bimbingan Ibu Ervina Rachmawati, S.ST., MPH.

Kanker serviks merupakan salah satu jenis kanker dengan angka kejadian dan kematian tertinggi pada wanita di Indonesia. Upaya pencegahan dapat dilakukan dengan memprediksi tingkat risiko sejak dini agar intervensi medis dan promotif dapat dilakukan sebelum penyakit berkembang. Berdasarkan data 10 besar kunjungan pasien kanker di RSUP dr. Kariadi Semarang periode Januari hingga Juni 2025, kanker payudara menempati urutan tertinggi, diikuti oleh kanker serviks. Meskipun kanker payudara memiliki jumlah kasus lebih banyak, penelitian ini difokuskan pada kanker serviks karena RSUP dr. Kariadi telah memiliki program deteksi dini kanker payudara berupa *screening* ABUS yang juga berfokus pada pencegahan sejak dini, sementara sistem *screening* mandiri untuk prediksi risiko kanker serviks yang berfokus pada pencegahan sejak dini masih belum tersedia.

Penelitian ini bertujuan untuk membangun sistem prediksi risiko kanker serviks menggunakan algoritma C4.5 berdasarkan rekam medis pasien di RSUP dr. Kariadi Semarang. Sistem ini diharapkan dapat membantu tenaga medis dalam mengidentifikasi tingkat risiko pasien secara cepat dan akurat, sehingga dapat digunakan sebagai dasar dalam pengambilan keputusan klinis. Data yang digunakan terdiri atas 200 sampel rekam medis pasien rawat jalan dan rawat inap, yang mencakup 100 pasien dengan diagnosis kanker serviks dan 100 pasien non-kanker serviks. Jumlah ini dipilih agar distribusi data seimbang dan hasil klasifikasi lebih representatif.

Analisis data dilakukan dengan pendekatan *Knowledge Discovery in Database* (KDD), yang mencakup tahapan data *selection*, data *transformation*, data mining, dan *knowledge*. Tahap data *selection* dilakukan untuk memilih atribut yang relevan terhadap risiko kanker serviks, yaitu jumlah pasangan seksual, jumlah persalinan, riwayat keluarga, dan keputihan abnormal. Tahap data *transformation* dilakukan dengan proses kategorisasi (*discretization*), penyandian (*encoding*), dan

konversi format ke CSV agar sesuai dengan kebutuhan analisis menggunakan perangkat lunak WEKA. Tahap utama, yaitu *data mining*, menggunakan algoritma C4.5 untuk membangun model klasifikasi risiko kanker serviks.

Pengujian model dilakukan dengan rasio pembagian data latih dan data uji sebesar 50:50, 60:40, 70:30, 80:20, dan 90:10. Hasil terbaik diperoleh pada rasio 80:20 dan 90:10 dengan tingkat akurasi mencapai 95%. Atribut keputihan abnormal muncul secara konsisten sebagai node akar dalam pohon keputusan (*decision tree*), menunjukkan bahwa faktor ini merupakan penentu dominan terhadap tingkat risiko. Berdasarkan perhitungan *gain ratio*, atribut keputihan abnormal memiliki nilai tertinggi sebesar 0,556, diikuti oleh riwayat keluarga sebesar 0,181, dan jumlah persalinan sebesar 0,153. Pohon keputusan menghasilkan enam aturan klasifikasi (*IF-THEN rules*) yang memisahkan kategori berisiko tinggi dan rendah secara jelas. Evaluasi model menggunakan *confusion matrix* menunjukkan akurasi 95%, presisi 94,7%, dan sensitivitas 94,7%, yang menandakan model memiliki kemampuan prediksi yang sangat baik.

Pengembangan sistem dilakukan menggunakan metode *Rapid Application Development* (RAD) untuk mempercepat proses pembuatan sistem tanpa mengurangi kualitas hasil. Tahapan RAD meliputi *requirement planning*, *design workshop*, dan pengembangan. Pada tahap *requirement planning* disini mengidentifikasi kebutuhan sistem prediksi risiko kanker serviks berdasarkan permasalahan dan tujuan penelitian dengan mempertimbangkan keterbatasan waktu, kemudahan penggunaan, serta efektivitas pengolahan data. Tahap *design workshop* menghasilkan rancangan antarmuka menggunakan aplikasi *Figma*, sedangkan tahap pengembangan dilakukan dengan membangun sistem berbasis web menggunakan *PHP*, *MySQL*, dan *XAMPP*.

Sistem yang dihasilkan terdiri atas 2 pengguna, yaitu pengguna umum dan admin sistem. Pengguna umum dapat melakukan prediksi dan melihat hasilnya, sedangkan admin memiliki akses tambahan untuk meninjau seluruh riwayat prediksi pasien. Dengan akurasi model sebesar 95% dan hasil pengujian sistem yang berjalan baik, sistem prediksi risiko kanker serviks berbasis algoritma C4.5 ini berpotensi menjadi alat bantu pengambilan keputusan yang efektif bagi tenaga