

ABSTRAK

Penyakit lambung merupakan salah satu gangguan pada sistem pencernaan yang banyak dipengaruhi oleh pola hidup tidak sehat, seperti pola makan yang tidak teratur, konsumsi makanan pedas dan asam, kebiasaan merokok, konsumsi kafein, serta tingkat stres yang tinggi. Kondisi tersebut dapat meningkatkan risiko terjadinya gangguan lambung yang apabila tidak ditangani sejak dini dapat berkembang menjadi lebih serius. Oleh karena itu, diperlukan suatu metode yang mampu membantu proses deteksi dini risiko penyakit lambung secara cepat dan akurat. Penelitian ini bertujuan untuk mengimplementasikan algoritma C5.0 dalam mengklasifikasikan risiko penyakit lambung berdasarkan pola hidup pasien ke dalam dua kategori, yaitu parah dan tidak parah. Metode penelitian yang digunakan meliputi proses pengumpulan data, *preprocessing* data, transformasi data, pembentukan label klasifikasi, serta proses klasifikasi menggunakan algoritma C5.0. Selanjutnya dilakukan pembentukan pohon keputusan dan evaluasi model menggunakan *confusion matrix* untuk mengetahui performa model klasifikasi yang dihasilkan. Data penelitian diperoleh melalui penyebaran kuesioner dan data pendukung lainnya yang berkaitan dengan pola hidup pasien. Hasil penelitian menunjukkan bahwa algoritma C5.0 mampu mengklasifikasikan risiko penyakit lambung dengan baik berdasarkan data pola hidup pasien serta menghasilkan model klasifikasi yang mudah dipahami. Sistem yang dibangun juga mampu membantu proses deteksi dini risiko penyakit lambung secara lebih efektif dan efisien. Dengan demikian, algoritma C5.0 dinilai efektif untuk diterapkan dalam klasifikasi risiko penyakit lambung berbasis pola hidup pasien dan dapat digunakan sebagai alat bantu pendukung keputusan dalam bidang kesehatan.

Kata Kunci: Penyakit Lambung, Algoritma C5.0, *Data Mining*, Klasifikasi, Pola Hidup Pasien.

ABSTRACT

Gastric disease is a digestive system disorder that is largely influenced by unhealthy lifestyles, such as irregular eating habits, consumption of spicy and acidic foods, smoking, caffeine consumption, and high stress levels. These conditions can increase the risk of gastric disorders, which, if not addressed early, can progress to more serious conditions. Therefore, a method is needed to facilitate the rapid and accurate early detection of gastric disease risk. This study aims to implement the C5.0 algorithm to classify gastric disease risk based on patient lifestyle into two categories: severe and non-severe. The research methods used included data collection, data preprocessing, data transformation, classification label formation, and the classification process using the C5.0 algorithm. Decision tree development and model evaluation using a confusion matrix were then conducted to determine the performance of the resulting classification model. The research data were obtained through questionnaires and other supporting data related to patient lifestyles. The results showed that the C5.0 algorithm was able to effectively classify gastric disease risk based on patient lifestyle data and produced an easily understood classification model. The developed system also facilitates the early detection of gastric disease risk more effectively and efficiently. Thus, the C5.0 algorithm is considered effective for classification of gastric disease risk based on patient lifestyle and can be used as a decision-support tool in the healthcare sector.

Keywords: *Gastric Disease, C5.0 Algorithm, Data Mining, Classification, Patient Lifestyle.*