

ABSTRAK

IMPLEMENTASI KECERDASAN BUATAN (*ARTIFICIAL INTELLIGENCE*) PADA KANKER PROSTAT

Oleh

RUSLIYAWATI

Penyakit kanker merupakan salah satu penyebab kematian utama di seluruh dunia. Salah satu jenis kanker yang menjadi penyebab kematian pada populasi pria adalah kanker prostat. Penyakit ini hanya terdapat pada pria karena pada wanita tidak memiliki kelenjar prostat. Ada beberapa faktor penyebab kanker prostat antara lain pola gaya hidup merokok dan aktivitas seksual. Kesalahan dalam pemberian keputusan faktor penyebab, tingkat risiko, dan kesesuaian pengobatan yang diberikan pada pasien kanker prostat sangat menentukan proses kesembuhan pasien dan menjadi masalah tersendiri dalam pelayanan kesehatan. Oleh karena itu perlu dilakukan penelitian untuk mengembangkan sebuah model sistem pakar medis yang dapat digunakan dalam mengklasifikasi kesesuaian prediksi kanker prostat berdasarkan faktor penyebab dan rekam medis pasien. Variabel masukan penelitian yaitu usia, *antigen spesifik prostate* (PSA), *volume prostate* (PV), dan *% free PSA* (*% fPSA*), skor gleason, durasi merokok, kebiasaan merokok, frekuensi seksual, dan usia pasien. Variabel dependen adalah risiko kanker prostat (PCR). Penelitian ini menggunakan *fuzzy logic* dengan metode FIS Mamdani, *Profile Matching*, dan *Naïve Bayes*. Penelitian yang telah dikembangkan saat ini juga hanya berfokus pada pengembangan aplikasi untuk menyelesaikan permasalahan atau membantu dalam mendiagnosa stadium kanker prostat. Belum ditemukan penelitian yang mengembangkan model teknik pemilihan fitur dan mengevaluasinya untuk memperoleh fitur-fitur yang relevan dengan kelas sehingga mampu meningkatkan akurasi metode *machine learning*.

Penelitian ini bertujuan untuk melakukan studi literatur dan analisis bibliometrik mengenai tren penggunaan kecerdasan buatan (*Artificial Intelligence/AI*) dalam prediksi kanker prostat, dengan pendekatan *Systematic Literature Review* (SLR) yang divisualisasikan menggunakan *VOSviewer*. Selain itu, penelitian ini mengevaluasi teknik seleksi fitur seperti *Information Gain* dan *Gain Ratio* untuk menghasilkan fitur yang relevan guna meningkatkan kinerja model *machine learning*. Evaluasi dan perbandingan dilakukan terhadap lima metode klasifikasi, yaitu *Naïve Bayes*, *Support Vector Machine (SVM)*, *Decision Tree*, *Random Forest*, dan *K-Nearest Neighbors (KNN)*, berdasarkan metrik kinerja meliputi akurasi, sensitivitas, spesifisitas, dan *Area Under the ROC Curve* (AUC-ROC). Untuk mengatasi masalah ketidakseimbangan kelas dalam data, model *Synthetic Minority Oversampling Technique* (SMOTE) juga dikembangkan sebagai bagian dari pendekatan ini.

Hasil penelitian pertama menunjukkan bahwa model *fuzzy* yang dikembangkan memiliki potensi sebagai alat bantu bagi dokter spesialis dalam

mendiagnosis dan menganalisis kemungkinan kanker prostat. Nilai PCR yang dihasilkan berada dalam rentang prediksi klinis dokter ahli, sehingga dapat dijadikan pertimbangan dalam pengambilan keputusan, terutama terkait perlunya tindakan biopsi. Sistem pakar *fuzzy* Mamdani yang dibangun dengan empat parameter *input* klinis, usia, PSA, PV, dan %fPSA mampu memprediksi risiko kanker prostat dalam bentuk nilai PCR berbasis persentase. Sistem ini menunjukkan performa yang baik, dengan akurasi 88%, sensitivitas 88,37%, spesifisitas 87,72%, dan presisi 84,44%.

Hasil penelitian kedua memberikan kontribusi melalui analisis bibliometrik dan tinjauan sistematis terhadap tren penerapan kecerdasan buatan (AI) dalam diagnosis kanker prostat. Hasil analisis menunjukkan bahwa integrasi AI dalam sistem medis terus berkembang, terutama pada aspek pengobatan dan penerapan algoritma seperti *DNN*, *HNN*, *FCN*, *QT for C++*, dan *ANN*. Literatur yang dianalisis mencakup 37 artikel terpilih yang dipublikasikan antara tahun 2015 hingga 2022, yang diperoleh dari basis data *PubMed*, *Embase*, dan *Web of Science*. Analisis dilakukan menggunakan Microsoft Excel 2019 dan GraphPad Prism 8. Temuan menunjukkan peningkatan signifikan dalam jumlah publikasi terkait AI pada kanker prostat, dengan dominasi penggunaan data citra medis (62%) dibandingkan data numerik.

Hasil penelitian ketiga mengenai evaluasi terhadap lima algoritma *machine learning* menunjukkan bahwa algoritma SVM memiliki akurasi tertinggi (88%) dan performa terbaik dalam *precision* dan *recall*, meskipun dengan waktu komputasi paling lama (0,09 detik), sedangkan *Naïve Bayes* unggul dalam efisiensi waktu (0,01 detik) dengan akurasi yang tetap kompetitif (81%). Penerapan seleksi fitur menggunakan *Information Gain* dan *Gain Ratio* terbukti meningkatkan akurasi sebagian besar algoritma hingga rata-rata 2% secara signifikan ($p\text{-value} < 0,05$), kecuali SVM yang tetap stabil meskipun tanpa reduksi fitur. Temuan ini menunjukkan pentingnya pemilihan algoritma dan teknik seleksi fitur yang sesuai dengan kebutuhan aplikasi; kombinasi antara algoritma efisien seperti *Naïve Bayes* atau *AdaBoost* dengan seleksi fitur yang tepat dapat menghasilkan sistem diagnosis kanker prostat yang akurat, cepat, dan dapat diandalkan dalam praktik klinis.

Keyword: kanker prostat, *artificial intelligence*, *machine learning*, *fuzzy logic*

ABSTRACT

IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE METHODS IN PROSTATE CANCER DIAGNOSIS

By

RUSLIYAWATI

Cancer remains one of the leading causes of death worldwide. Among men, cancer is a major contribution to cancer-related mortality. This disease only occurs in men because women do not have a prostate gland. Several factors are associated with the development of prostate cancer, including smoking lifestyle patterns and sexual activity. Mistakes in making decisions about causal factors, risk levels, and the appropriateness of treatment given to prostate cancer patients greatly determine the patient's healing process and become a separate problem in health services. Therefore, it is necessary to develop a medical expert system model that can classify the appropriateness of prostate cancer predictions based on causal factors and patient medical records. The input variables for the study were age, prostate specific antigen (PSA), prostate volume (PV), and % free PSA (% fPSA), Gleason score, smoking duration, smoking habits, sexual frequency, and patient age. The dependent variable is the risk of prostate cancer (PCR). This study uses fuzzy logic with the FIS Mamdani, Profile Matching, and Naïve Bayes methods. Previous research has largely focused on application development to assist in diagnosing prostate cancer stages, with little attention given to feature selection techniques aimed at improving the performance of machine learning algorithms.

This study aims to conduct a literature study and bibliometric analysis on the trend of using artificial intelligence (AI) in prostate cancer prediction, with a Systematic Literature Review (SLR) approach visualized using VOSviewer. In addition, this study evaluates feature selection techniques such as Information Gain and Gain Ratio to produce relevant features to improve the performance of machine learning models. Evaluation and comparison were carried out on five classification methods, namely Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, and K-Nearest Neighbors (KNN), based on performance metrics including accuracy, sensitivity, specificity, and Area Under the ROC Curve (AUC-ROC). To address the issue of class imbalance in the dataset, the Synthetic Minority Oversampling Technique (SMOTE) was also implemented.

The results of the first study specify that the developed fuzzy expert system can serve as a decision-support tool for specialist physicians in diagnosing and investigating the likelihood of prostate cancer. The predicted PCR values align with the clinical prediction ranges of expert doctors and can assist in decision making, particularly concerning the necessity of a biopsy. The Mamdani fuzzy expert system, built with four clinical input parameters. They are age, PSA, PV, and %f PSA which effectively predicts prostate cancer risk as a percentage-based PCR value. The

system demonstrated strong performance, with an accuracy of 88%, sensitivity of 88.37%, specificity of 87.72%, and precision of 84.44%.

The second study contributes to the field through a bibliometric analysis and systematic review of AI applications in prostate cancer diagnosis. The findings show that the integration of AI in medical systems is expanding, especially in treatment optimization and the use of advanced algorithms such as DNN, HNN, FCN, QT for C++, and ANN. The literature review included 37 selected articles published between 2015 and 2022, retrieved from databases such as PubMed, Embase, and Web of Science. Analyses were conducted using Microsoft Excel 2019 and GraphPad Prism 8. Results indicate a significant rise in publications related to AI in prostate cancer, with a dominant use of medical imaging data (62%) over numerical data.

The third study estimated the performance of five machine learning algorithms, revealing that Random Forest achieved the highest accuracy at 92%, followed by SVM (90%), KNN (86%), and Naïve Bayes (83%). However, these algorithms are often considered "black boxes," offering limited interpretability for medical professionals. While fuzzy systems excel in interpretability, machine learning models are superior in accurateness and computational efficiency. Thus, integrating fuzzy logic with machine learning holds promise for emerging a diagnostic system that is both accurate and interpretable.

Keywords: prostate cancer, artificial intelligence, machine learning, fuzzy logic