

Pemanfaatan machine learning untuk diagnosis kanker prostat dari spektrum raman pada sampel jaringan biopsi = Utilization of machine learning for prostate cancer diagnosis from raman spectrum on biopsy tissue samples

Aryo Haris Wirakusuma, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920566899&lokasi=lokal>

Abstrak

Kanker prostat merupakan masalah kesehatan yang umum terjadi pada pria di seluruh dunia. Diagnosis dini dan akurat sangat penting untuk pengobatan yang efektif. Spektroskopi Raman (RS) menawarkan teknik yang menjanjikan untuk menganalisis sampel jaringan dan berpotensi membedakan antara jaringan prostat jinak dan ganas. Studi ini menyelidiki kemandirian algoritma Machine Learning dalam mengklasifikasikan kanker prostat menggunakan data RS dari spesimen biopsi. Data Spektrum Raman yang digunakan berasal dari penelitian di Kanada yang dikumpulkan dari tiga kelompok Cohort yakni Centre Hospitalier de l'Université de Montréal (CHUM) sebagai Cohort training, serta University Health Network (UHN) dan Centre Hospitalier de l'Université de Montréal (CHUQc-UL) sebagai Cohort testing. Spektrum ini, yang mewakili komposisi kimia jaringan (Raw Spectra), digunakan untuk training dan model evaluation Machine Learning.

Untuk membantu menganalisis komposisi kimia data RS yang akurat, diperlukan algoritma Machine Learning dalam mengklasifikasikan BPH dan PC. Dua algoritma yang digunakan adalah Support Vector Machine (SVM) dan Extreme Gradient Boosting (XGBoost). Data RS dilakukan training dan testing menggunakan data yang berbeda yang menghasilkan nilai metrik klasifikasi dari dua algoritma yang dibandingkan dalam mengklasifikasikan sampel BPH atau PC. Ternyata, algoritma XGBoost memiliki kemampuan klasifikasi yang kurang unggul daripada SVM, hal ini dibuktikan dengan SVM memiliki rata-rata Akurasi sebesar 83,3%, Sensitivitas sebesar 96,7%, Spesifisitas sebesar 46,4%, F1-Score sebesar 98,1%, dan ROC-AUC sebesar 87,7%. Sementara, XGBoost menunjukkan Akurasi sebesar 78%, Sensitivitas sebesar 80%, Spesifisitas sebesar 75%, F1-Score sebesar 78%, dan ROC-AUC sebesar 85%. Selain itu kedua algoritma juga bisa menentukan titik Feature Importance pada grafik Spektra Raman, yang ditunjukkan dengan beberapa fitur vibrasi molekul untuk BPH dan PC berdasarkan algoritma SVM dan XGBoost, yakni 720 cm⁻¹, 828 cm⁻¹, dan 931 cm⁻¹ sebagai karakteristik jaringan BPH, dan 1.431 cm⁻¹ dan 1.470 cm⁻¹ sebagai jaringan PC.

.....

Prostate cancer is a common health problem in men worldwide. Early and accurate diagnosis is essential for effective treatment. Raman spectroscopy (RS) offers a promising technique to analyze tissue samples and potentially differentiate between benign and malignant prostate tissue. This study investigated the efficacy of a Machine Learning algorithm in classifying prostate cancer using RS data from biopsy specimens. The Raman Spectrum data used were from a Canadian study collected from three cohort groups, namely the Centre Hospitalier de l'Université de Montréal (CHUM) as the training Cohort, and the University Health Network (UHN) and Centre Hospitalier de l'Université de Montréal (CHUQc-UL) as the testing Cohort. This spectrum, which represents the chemical composition of the tissue (Raw Spectra), is used for training and model evaluation Machine Learning.

To help analyze the chemical composition of accurate RS data, a Machine Learning algorithm is needed to classify BPH and PC. The two algorithms used are Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost). The RS data is trained and tested using different data that produces classification metric values from the two algorithms that are compared in classifying BPH or PC samples. It turns out that the XGBoost algorithm has a classification capability that is less superior than SVM, this is evidenced by SVM having an average Accuracy of 83.3%, Sensitivity of 96.7%, Specificity of 46.4%, F1-Score of 98.1%, and ROC-AUC of 87.7%. Meanwhile, XGBoost showed Accuracy of 78%, Sensitivity of 80%, Specificity of 75%, F1-Score of 78%, and ROC-AUC of 85%. In addition, both algorithms can also determine the Feature Importance point on the Raman Spectra graph, which is indicated by several molecular vibration features for BPH and PC based on the SVM and XGBoost algorithms, namely 720 cm⁻¹, 828 cm⁻¹, and 931 cm⁻¹ as characteristics of BPH tissue, and 1.431 cm⁻¹ and 1.470 cm⁻¹ as PC tissue.