Universitas Indonesia Library >> UI - Skripsi Membership

Implementasi Metode Support Vector Regression dan Light Gradient Boosting Machine dalam Memprediksi Usia Biologis pada Data Pemeriksaan Medis = Implementation of Support Vector Regression and Light Gradient Boosting Machine Methods for Predicting Biological Age from Medical Examination Data

Kayla Calista Ayal, author

Deskripsi Lengkap: https://lib.ui.ac.id/detail?id=9999920551894&lokasi=lokal

---

Abstrak

Usia biologis mengukur penuaan individu berdasarkan kondisi fisik dan fungsi organ. Meskipun banyak penelitian telah dilakukan untuk memprediksi usia biologis dengan berbagai metode, penerapan metode machine learning masih memiliki ruang untuk penelitian lebih lanjut. Penelitian ini mengimplementasikan dua metode machine learning dengan pendekatan yang berbeda, yaitu metode Support Vector Regression (SVR) dan Light Gradient Boosting Machine (LGBM) dalam memprediksi usia biologis menggunakan data pemeriksaan medis Kementerian Kesehatan tahun 2011 yang mencakup 5960 subjek dan 41 fitur. Proses preprocessing meliputi penyaringan usia kronologis > 30 tahun, pemisahan data berdasarkan jenis kelamin, penanganan missing values dan outlier, serta data encoding. Feature selection menggunakan koefisien korelasi Spearman menghasilkan 8 fitur berbeda untuk setiap jenis kelamin. Data dibagi dengan 90% untuk pelatihan dan 10% untuk pengujian, serta dilakukan tuning hyperparameter menggunakan GridSearchCV. Penelitian ini menggunakan metrik RMSE dan adjusted R-squared, yang dipilih berdasarkan relevansinya dengan tujuan penelitian. Hasil menunjukkan LGBM lebih unggul dari SVR dengan RMSE 7,2064 tahun dan adjusted R-squared 33,36% pada pria, serta RMSE 7,1475 tahun dan adjusted R-squared 22,16% pada wanita. Analisis korelasi menunjukkan hubungan yang wajar antara usia biologis dan usia kronologis serta korelasi yang cukup antara usia biologis dengan biomarker tekanan sistolik dan status hipertensi pada pria, serta tekanan sistolik dan kolesterol pada wanita. Analisis korelasi menunjukkan hubungan signifikan antara usia biologis dengan usia kronologis dan beberapa biomarker. Secara keseluruhan, LGBM lebih efektif dalam memprediksi usia biologis dibandingkan SVR. Hasil dari penelitian diharapkan dapat diaplikasikan dalam kehidupan sehari-hari, seperti membantu perusahaan asuransi menilai kelayakan klaim berdasarkan prediksi usia biologis, serta mendukung keputusan di bidang kesehatan preventif.
.....Biological age measures an individual's aging based on physical condition and organ function. Although numerous studies have been conducted to predict biological age using various methods, there is still room for further research in the application of machine learning techniques. This study implements two machine learning methods with different approaches, namely Support Vector Regression (SVR) and Light Gradient Boosting Machine (LGBM), to predict biological age using medical examination data from the Ministry of Health in 2011, covering 5960 subjects and 41 features. The preprocessing steps include filtering chronological age > 30 years, segregating data by gender, handling missing values and outliers, and data encoding. Feature selection using Spearman correlation coefficients resulted in 8 different features for each gender. The data was split into 90% for training and 10% for testing, with hyperparameter tuning performed using GridSearchCV. This study used RMSE and adjusted R-squared metrics, selected based on their relevance to the research objectives. The results show that LGBM outperformed SVR with an RMSE of 7.2064 years and an adjusted R-squared of 33.36% for men, and an RMSE of 7.1475 years and an adjusted

R-squared of 22.16% for women. Correlation analysis revealed a significant relationship between biological age and chronological age, as well as a reasonable correlation between biological age and biomarkers such as systolic blood pressure and hypertension status in men, and systolic blood pressure and cholesterol in women. Overall, LGBM proved to be more effective in predicting biological age compared to SVR. The results of this study are expected to be applicable in everyday life, such as assisting insurance companies in evaluating claim eligibility based on biological age predictions, as well as supporting decision-making in preventive healthcare.