

Analisis Pengaruh Noise terhadap Performa K-Nearest Neighbors Algorithm dengan Variasi Jarak untuk Klasifikasi Beban Listrik = Analysis of the Effect of Noise on the Performance of the K-Nearest Neighbors Algorithm with Distance Variations in Electrical Load Classification

Aris Surya Yunata, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920544041&lokasi=lokal>

Abstrak

Teknik NILM (Non-Intrusive Load Monitoring) digunakan dalam pemantauan konsumsi energi. Penerapan NILM digunakan untuk energi efisiensi, manajemen energi, dan diagnosa peralatan di rumah tangga, industri atau penyedia energi. Variabel pengukuran yang digunakan yaitu daya aktif dan daya reaktif. Namun, data pengukuran sering kali terpengaruh oleh noise. Berbagai macam metode digunakan dalam NILM. Metode K-NN adalah salah satu metode machine learning yang banyak digunakan untuk klasifikasi beban listrik dalam teknik NILM dengan performa yang baik dan bersaing dengan metode lain yang lebih kompleks. Penelitian ini bertujuan untuk menganalisis pengaruh noise terhadap performa algoritma k-Nearest Neighbors (K-NN) dalam klasifikasi beban listrik. Berbagai tingkat noise secara random diberikan pada data pengukuran yang diperoleh sebesar 1% hingga 20%. Selanjutnya, model K-NN dilatih dan dievaluasi dengan nilai $k = 1$ sampai 25 dan menggunakan 15 variasi jarak. Dalam penelitian ini bahasa pemrograman python digunakan untuk mengevaluasi performa K-NN. Hasil eksperimen menunjukkan bahwa penambahan noise pada data pengukuran secara signifikan mempengaruhi performa algoritma K-NN dalam mengklasifikasikan beban listrik. Pengaruh ini terlihat pada nilai akurasi, presisi dan recall. Performa K-NN menurun hingga 15% yang didapatkan dari perbandingan nilai akurasi untuk data yang diberikan noise 20%. Nilai k yang memberikan akurasi maksimal $k = 25$ untuk data yang diberikan noise. Nilai k yang memberikan presisi dan recall maksimal bernilai $k = 3$ untuk data yang diberikan noise. Dari ke 15 jenis jarak yang dipakai di metode K-NN pada penelitian ini, jarak Clark dan Divergence yang memiliki nilai akurasi diatas rata-rata, jarak Canberra memiliki nilai presisi di atas rata-rata dan jarak Neyman Chi Squared memiliki nilai recall diatas rata-rata. Perbandingan performa antara metode K-NN dengan Random Forest dan Extra Trees Classifier juga telah dilakukan. Dari hasil pengujian dan penelitian didapatkan bahwa dengan metode K-NN memberikan performa yang baik untuk mendisagregasi data yang diberikan noise besar dibandingkan dengan metode Random Forest dan Extra Trees Classifier. metode K-NN memiliki nilai akurasi lebih tinggi dibandingkan dengan metode Random Forest dan Extra Trees Classifier, Selisih yang dihasilkan mencapai 15%.

.....The NILM (Non-Intrusive Load Monitoring) technique is used in monitoring energy consumption. NILM applications include energy efficiency, energy management, and appliance diagnostics in households, industries, or energy providers. Measurement variables used are Real Power and Reactive Power. However, measurement data are often affected by noise. Appliance diagnosis uses various machine learning methods. The K-NN method is one of the widely used machine learning methods for classifying electrical loads with good performance, competing even with more complex methods. Python has become a mainstay in data science. This programming language enables data analysis to perform machine learning algorithms. This study aims to analyze the impact of noise on the performance of the k-Nearest Neighbors (K-NN) algorithm

in classifying electrical loads. Various noise levels, ranging from 1% to 20%, were randomly added to the measurement data obtained. Subsequently, the K-NN model was trained and evaluated with k values ranging from 1 to 25, using 15 distance variations. Experimental results showed that the addition of noise to the measurement data significantly affected the performance of the K-NN algorithm in classifying electrical loads. This impact is reflected in the values of accuracy, precision, and recall. K-NN performance decreased by up to 15%, as indicated by the accuracy comparison for data with 20% noise. The k value providing maximum accuracy was $k = 25$ for both low and high noise data. The k value providing maximum precision and recall was $k = 3$ for both low and high noise data. Among the 15 types of distances used in the K-NN method in this study, Clark and Divergence distances had above-average accuracy values, Canberra distance had above-average precision values, and Neyman Chi-Squared distance had above-average recall values. Testing and research results showed that the K-NN method performs well in disaggregating data with high noise compared to the Random Forest and Extra Trees Classifier methods.