

Analisa dan seleksi fitur model k-nearest neighbor untuk mengklasifikasi kualitas tidur berdasarkan dataset pmdata = Features selection analysis from k-nearest neighbor model for sleep quality classification based on pmdata dataset.

Bagaskara Ghanyvian Istiqlal, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20505368&lokasi=lokal>

Abstrak

Kualitas tidur yang baik sangatlah penting untuk berbagai aspek kehidupan seperti kesehatan fisik, kesehatan mental, keselamatan, konsentrasi, performa, penyembuhan, dan lain-lain. Kualitas tidur tidak hanya mencakup aspek fisiologis, tetapi juga memperhatikan aspek mental seperti: kondisi setelah tidur, kepuasan dengan tidur, dan pengaruh pada kehidupan sehari-hari. Penelitian ini mengusulkan penggabungan data objektif yang berasal dari Fitbit dan kuesioner subjektif untuk mengklasifikasi kualitas tidur menggunakan K-Nearest Neighbor. Klasifikasi ini bertujuan untuk mempelajari fitur-fitur yang paling pengaruh dalam kualitas tidur. Data objektif yang berisikan data fisiologis dan aspek tidur terukur oleh Fitbit, serta data subjektif mengenai aspek mental, keduanya dijadikan fitur deskriptif dalam model. Analisa fitur yang paling berpengaruh dilakukan dari dua sudut pandang model, yaitu fitur target kualitas tidur subjektif dan fitur target kualitas objektif. Kedua model dilatih dengan serangkaian data preprocessing yang termasuk didalamnya terdapat seleksi fitur dan ekstraksi fitur. Seleksi fitur berbasis ANOVA F Test akan dibandingkan dengan ekstraksi fitur Principal Component Analysis (PCA) dan Neighborhood Component Analysis (NCA). Seleksi fitur ANOVA F-Test lebih baik dari PCA dan NCA dengan peningkatan skor sebesar 0,06-0,08 pada model objektif, dan 0,01-0,06 pada model subjektif. Skor terbaik terbaik dari model subjektif yaitu 0,52 dengan parameter jumlah fitur = 3 dan k-neighbors = 27. Skor terbaik terbaik dari model objektif yaitu 0,72 dengan parameter jumlah fitur = 7 dan k-neighbors = 4. Pada akhirnya, ditemukan 3 Fitur yang paling berpengaruh dalam klasifikasi subjektif, dan 7 fitur yang paling berpengaruh dalam klasifikasi objektif.

<hr>

Good quality sleep is very important for various aspects of life such as physical health, mental health, safety, concentration, performance, healing, and others. Sleep quality does not only include physiological aspects, but also pay attention to mental aspects such as condition after sleep, satisfaction with sleep, and influence on daily life. This study proposes combining objective data from Fitbit and subjective questionnaires to classify sleep quality using K-Nearest Neighbor. This classification aims to study the features that have the most influence in sleep quality. Objective data containing physiological data and sleep aspects measured by Fitbit, as well as subjective data on mental aspects, are both used as descriptive features in the model. The analysis of the most influential features is carried out from two viewpoints of the model, namely the subjective sleep quality target feature and the objective quality target feature. Both models are trained with a series of preprocessing data which includes feature selection and feature extraction. ANOVA F Test based on feature selection will be compared with feature extraction of Principal Component Analysis (PCA) and Neighborhood Component Analysis (NCA). ANOVA F-Test feature selection is better than PCA and NCA with an increase in scores of 0.06-0.08 in the objective model, and 0.01-0.06 in the subjective model. The best score of the subjective model is 0.52 with the parameter number of features = 3 and k-neighbors = 27.

The best score of the objective model is 0.72 with the parameter number of features = 7 and k-neighbors = 4. In the end, it was found 3 the most influential features in the subjective classification, and 7 the most influential features in the objective classification.