

Prediksi area riset dengan pendekatan kernel based ensemble = forecasting research areas using kernel based ensemble approach

Agus Widodo, author

Deskripsi Lengkap: <http://lib.ui.ac.id/detail?id=20404464&lokasi=lokal>

Abstrak

Saat ini penentuan area riset masih banyak bergantung kepada pendapat para ahli. Meskipun ahli tersebut memiliki pengetahuan yang mendalam di bidangnya, akan tetapi tidak semua area riset yang emerging dapat diketahui oleh ahli tersebut mengingat cepatnya perkembangan sumber-sumber informasi tentang ilmu pengetahuan dan teknologi. Namun demikian, analisis data yang berjumlah besar memerlukan waktu yang lama dan bisa jadi subyektif jika menggunakan cara manual. Beberapa penelitian sebelumnya telah menggunakan teknik kuantitatif dengan menghitung trend berdasarkan jumlah kata kunci dari suatu topik riset dan memprediksi trend tersebut untuk masa yang akan datang. Untuk prediksi trend dari data time series, saat ini pendekatan machine learning mulai banyak dikaji disamping pendekatan statistik yang sebelumnya lazim digunakan.

Sementara itu, pendekatan ensemble yang menggabungkan hasil prediksi, teknik prediksi atau representasi data diyakini dapat meningkatkan akurasi prediksi. Multiple Kernel Learning (MKL) merupakan suatu teknik ensemble melalui penggabungan kernel yang menggunakan teknik machine learning, yakni Support Vector Machine (SVM), sebagai classifier atau prediktor. Dalam penelitian sebelumnya, MKL telah dimanfaatkan untuk menggabungkan fitur, yang biasa disebut sebagai data integration, dalam bidang image processing tetapi masih menggunakan single kernel. Dalam penelitian ini, MKL dimanfaatkan untuk menggabungkan fitur data time series yang berupa sliding windows dan diterapkan pada multiple kernel. Disamping itu, penelitian ini juga mengajukan penggunaan data historis sebagai pengganti training dataset untuk memilih model prediksi yang sesuai dengan karakteristik time series karena setiap model prediksi memiliki kelebihan dan keterbatasan dalam memprediksi data time series yang jenisnya cukup beragam.

<hr>

Currently, the determination of the research area is still largely dependent on the opinion of experts. Although experts have in-depth knowledge in the field, but not all areas of emerging research can be known by the experts given the rapid development of sources of information regarding science and technology. However, the analysis of large amounts of data would take a quite long time and the result could be subjective if a manual method is employed. Several previous studies have used quantitative techniques to calculate trends based on the number of keywords on research topics and forecast their future trends. For the trend forecasting of time series data, currently, machine learning approaches have been extensively studied in addition to the previous statistical approaches which are commonly used.

Meanwhile, an ensemble approach that may combine the prediction results, prediction techniques or data representations has the capability to increase the prediction accuracy. Multiple Kernel Learning (MKL) is one of such ensemble methods that optimizes the combination of kernels through the use of machine learning technique, such as Support Vector Machine (SVM), as a classifier or predictor. In previous studies,

MKL has been used to combine features, which is commonly referred to as the data integration approach, in the field of image processing but is still implemented on a single kernel. In this study, MKL is used to combine the features of time series data in the form of sliding windows and tested on multiple kernels. In addition, this study also proposes the use of historical data as a substitute for the training dataset to select the prediction technique based on the characteristics of time series considering the diverse kind of time series data such that no single prediction technique can be used for all types of data.