

INTISARI

METODE KLASIFIKASI *RANDOM K NEAREST NEIGHBOR* (RKNN)

Sinung Jiwangga Adi
09/283159/PA/12463

Data dimensi tinggi banyak tersedia di bioinformatika, kemometrika, perbankan dan aplikasi lainnya, sehingga kesuksesan analisis dan pemodelan data ini sangat menantang. *Random K Nearest Neighbor* (RKNN) terdiri dari serangkaian model berbasis *k nearest neighbor*, yang masing-masing diambil subset acaknya dari variabel input. Sebuah analisis teoritis dan empiris dilakukan terhadap kinerja RKNN. Berdasarkan RKNN yang diusulkan, dibangun sebuah metode seleksi fitur. Dalam *meranking* fitur yang penting dibuat kriteria yang dinamakan *support*, yang didefinisikan dan dihitung dari kerangka RKNN. Metode seleksi model dua tahap *backward* dikembangkan menggunakan *support*. Pendekatan RKNN dapat diaplikasikan untuk data respons kualitatif maupun kuantitatif, misalnya masalah klasifikasi dan regresi, serta aplikasinya dalam statistika, perbankan, *machine learning*, pengenalan pola, bioinformatika, dan lain-lain.

Kata kunci: klasifikasi, seleksi fitur, RKNN.

ABSTRACT

CLASSIFICATION USING RANDOM K NEAREST NEIGHBOR (RKNN)

Sinung Jiwangga Adi
09/283159/PA/12463

High dimensional data is widely available in bioinformatics, chemometrics, banking and other applications, so that the successful analysis and modeling of these data is highly challenging. Random K Nearest Neighbor (RKNN) consist an ensemble of base k nearest-neighbor models, each taking a random subset of the input variables. A theoretical and empirical analysis of the performance of the RKNN is performed. Based on the proposed RKNN, a new feature selection method is devised. To rank the importance of the variables, a criterion, named support, is defined and computed on the RKNN framework. A two-stage backward model selection method is developed using supports. The RKNN approach can be applied to both qualitative and quantitative responses, i.e., classification and regression problems, and has applications in statistics, banking, machine learning, pattern recognition and bioinformatics, etc.

Keywords: classification, feature selection, RKNN.