

INTISARI

Pengembangan Metode Reduksi Data dan Inisialisasi Vektor Referensi pada Algoritma *Learning Vector Quantization*

Oleh

Pande Nyoman Ariyuda Semadi

11/324356/PPA/03679

Algoritma LVQ merupakan algoritma pembelajaran terawasi yang digunakan untuk klasifikasi statistik dan pengenalan. Lapisan kompetitif pada LVQ mempelajari vektor-vektor masukan dan mengklasifikasikannya ke dalam kelas yang sesuai. Inisialisasi vektor referensi merupakan salah satu hal yang harus diperhatikan untuk dapat menghasilkan model klasifikasi yang baik. Pemberian nilai awal yang baik akan membuat jaringan LVQ berjalan lebih optimal. Jumlah data yang dilibatkan dalam proses pembelajaran dapat dikurangi dengan menggunakan metode reduksi. Pengurangan data ini dilakukan dengan tetap berusaha mempertahankan nilai akurasi yang sudah ada.

Penelitian ini mengusulkan metode reduksi dengan memanfaatkan kedekatan geometris data dan klasterisasi untuk inisialisasi vektor referensi. Dengan pengaturan tertentu, penggabungan dua metode tersebut memberikan keuntungan dalam pengurangan jumlah data yang perlu dilibatkan dalam proses pelatihan LVQ. Hal ini dapat diasosiasikan dengan keuntungan di bidang storage. Akurasi klasifikasi dapat dipertahankan dan justru dapat ditingkatkan dalam beberapa kasus. Data yang perlu dilibatkan dalam proses pelatihan dapat diturunkan hingga hanya 45,05% pada dataset *diabetes*, 70,45% pada dataset *ionosphere* dan 41,95% pada dataset *breast cancer*.

Kata kunci : LVQ, reduksi data, inisialisasi vektor referensi

ABSTRACT

Developing Data Reduction Method and Reference Vector Initialization in Learning Vector Quantization Algorithm

by

Pande Nyoman Ariyuda Semadi

11/324356/PPA/03679

LVQ is a supervised learning algorithm used for statistical classification and recognition. The competitive layer in LVQ studies the input vectors and classify them into the correct class. Reference vector initialization is one of the important things to consider in regards to achieving a proper model of classification. A proper initial value will give a more optimum result for the LVQ network. The amount of involved data in the learning process can be reduced by using a reduction method. The data reduction is carried out by still maintaining the existing accuracy value.

This research proposed reduction method by utilizing geometrical proximity of the data and clustering technique for reference vector initialization. By certain adjustment, the combination of those method can reduce amount of involved data in the learning process by still maintaining the existing accuracy value. In several cases, it can increase classification accuracy. The amount of involved data in the learning process can be reduced until 45,05% for diabetes dataset, 70,45% for ionosphere dataset and 41,95% for breast cancer dataset.

Keywords : LVQ, data reduction, reference vector initialization