

Penerapan Particle Swarm Optimization untuk Penentuan Parameter Regularisasi pada Kernel Regularized Discriminant Analysis

Monica Widiarsi^{1,2}, Army Justitia¹, Agus Zainal Arifin¹

¹ Laboratorium Vision and Image Processing, Jurusan Teknik Informatika
Institut Teknologi Sepuluh Nopember (ITS), Surabaya

² Universitas Surabaya (Ubaya)

monicawidiarsi@yahoo.com, army_justitia@yahoo.com, agusza@cs.its.ac.id

Abstrak

Metode Kernel Regularized Discriminant Analysis (KRDA) mampu mengatasi permasalahan singularitas dan overfitting yang disebabkan data dengan sampel sedikit berdimensi tinggi. Kinerja KRDA dipengaruhi oleh nilai parameter regularisasi yang digunakan. Paper ini bertujuan mengoptimalkan nilai parameter regularisasi dengan menerapkan Particle Swarm Optimization (PSO) pada KRDA. Metode PSO digunakan untuk memilih parameter regularisasi optimal berdasarkan nilai akurasi klasifikasi pada KRDA. Uji coba menggunakan 4 dataset yaitu iris, wine, dan ionosphere dari UCI, serta dataset svmguide2 dari aplikasi bioinformatika. Hasil uji coba menunjukkan metode PSO-KRDA dapat menentukan nilai parameter regularisasi optimal secara otomatis. Akurasi klasifikasi yang didapatkan dari uji coba menunjukkan PSO-KRDA lebih baik daripada KRDA dan SVM.

Kata kunci : metode kernel, parameter regularisasi, KRDA, PSO

1. Pendahuluan

Analisis diskriminan merupakan salah satu metode yang sering digunakan pada klasifikasi. Namun, metode klasik analisa diskriminan, *Linear Discriminant Analysis* (LDA), tidak dapat mengatasi permasalahan singularitas ketika data memiliki dimensi tinggi dengan jumlah sampel yang sedikit [1]. PCA+LDA diusulkan dengan mengurangi dimensi data dikurangi terlebih dahulu menggunakan PCA, kemudian baru diterapkan LDA [2]. Metode RLDA diusulkan dapat mengatasi permasalahan singularitas. RLDA menambahkan pengali dari matriks identitas untuk matriks *total scatter*, yang dapat menstabilkan estimasi matriks kovarian. Informasi tambahan tersebut disebut parameter regularisasi (σ_R) dengan $\sigma_R > 0$ [3].

Fitur yang didapatkan dari LDA merupakan kombinasi linier dari data input asli, sehingga tidak dapat menangkap struktur tidak linier pada kumpulan data yang kompleks [4]. LDA diperluas untuk kasus tidak linier dengan menggunakan metode kernel. Metode *Kernel Uncorrelated Discriminant Analysis*

(KUDA) dapat digunakan untuk mengekstrak fitur yang tidak saling berkorelasi di dalam ruang dimensi tereduksi. Namun ketika matriks kernel pada KUDA *non singular*, maka KUDA memproyeksikan semua sampel dalam kelas yang sama ke satu vektor umum di ruang dimensi tereduksi. Jika jumlah sampel tiap kelas sangat besar, maka akan menimbulkan masalah *overfitting*, yang menghasilkan analisa diskriminan menjadi kurang optimal [5].

Untuk mengatasi permasalahan *overfitting* tersebut, Shuiwang Ji dan Jieping Ye (2008) mengusulkan metode *Kernel Regularized Discriminant Analysis* (KRDA) [5]. Namun, kinerja KRDA sangat bergantung pada nilai parameter regularisasi [5]. Akurasi dapat berubah-ubah sesuai dengan nilai parameter regularisasi yang digunakan. Oleh karena itu pemilihan nilai parameter regularisasi yang tepat sangat penting agar didapatkan nilai akurasi klasifikasi yang baik.

Paper ini bertujuan mengoptimalkan nilai parameter regularisasi dengan menerapkan PSO pada KRDA. PSO digunakan untuk menentukan nilai parameter regularisasi optimal berdasarkan akurasi klasifikasi pada KRDA. Diharapkan dengan penerapan PSO pada KRDA, dapat membantu penentuan nilai parameter regularisasi yang menghasilkan akurasi klasifikasi terbaik secara otomatis. Pada uji coba dibandingkan akurasi klasifikasi PSO-KRDA dengan KRDA dan SVM, dan didapatkan akurasi klasifikasi PSO-KRDA lebih baik daripada KRDA dan SVM.

Paper ini disusun dengan struktur sebagai berikut : bagian 2 membahas mengenai LDA. Bahasan mendetail mengenai KRDA dijelaskan pada bagian 3. Bagian 4 menjelaskan mengenai algoritma PSO untuk menentukan parameter regularisasi KRDA. Uji coba dan analisa dibahas pada bagian 5. Kesimpulan dituliskan pada bagian terakhir paper ini.

2. Linear Discriminant Analysis (LDA)

LDA merupakan metode untuk mereduksi dimensi data. LDA memproyeksikan data dimensi tinggi ke subruang dimensi yang lebih rendah dengan memaksimalkan pemisahan data dari kelas-kelas yang berbeda dan meminimalkan penyebaran data dari kelas yang sama secara bersamaan.