

ELM paralel adaptif dengan fitur konvolusional untuk aliran data besar = Adaptive parallel ELM with convolutional features for big stream data

Arif Budiman, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20453952&lokasi=lokal>

Abstrak

ABSTRAK

Tantangan aliran data besar membutuhkan mesin pembelajaran khusus. Ragam, variabilitas dan kompleksitas berkaitan dengan masalah pergeseran konsep CD. Jumlah dan kecepatan berkaitan dengan masalah skalabilitas. Kami mengusulkan pendekatan integrasi jaringan syaraf konvolusi CNN dengan mesin pembelajaran ekstrem ELM yang menggunakan banyak CNNELM secara paralel. Solusi CD pendekatan pertama, CNNELM adaptif ACNNELM-1 menggunakan ELM tunggal dengan banyak CNN dengan menerapkan Adaptif Online Sequential ELM. Pendekatan kedua ACNNELM-2 menggunakan paduan penggabungan matriks dari banyak CNNELM. Solusi skalabilitas, Distributed averaging DA CNNELM bekerja dengan konsep MapReduce. CNNELM mulai dengan cetakan bobot yang sama kemudian dilatih secara asinkronus menggunakan partisi dari data pelatihan. Hasil akhir didapat dengan merata-ratakan bobot kernel dan ELM output. Ini menghemat waktu pelatihan dibandingkan satu CNNELM dengan pelatihan keseluruhan data. Kami mempelajari metoda pelatihan propagasi balik untuk memperbaiki akurasi dengan iterasi. Kami melakukan verifikasi dengan data extended MNIST, Not-MNIST dan CIFAR10. Kami mensimulasikan pergeseran virtual, pergeseran nyata, dan pergeseran hibrid. Pelatihan DA membagi data pelatihan menjadi beberapa himpunan partisi lebih kecil. Perangkat yang dipakai adalah Deep Learning toolbox dengan CPU parallel, dan Matconvnet dengan GPU. Kelemahan metoda ini memerlukan pemilihan penambahan parameter pembelajaran dan distribusi data pelatihan.

ABSTRACT

Big stream data challenges need special machine learning. Variety, variability and complexity are related with concept drift CD problem. Volume and velocity are related with scalability problem. We proposed integration approach Convolutional Neural Network CNN with Extreme Learning Machine ELM that used multi parallel CNNELM. For CD, the 1st approach, the Adaptive CNNELM ACNNELM 1 used single ELM with multi CNN by employing Adaptive Online Sequential ELM. The 2nd approach ACNNELM 2 used matrices concatenation ensemble from multi CNNELM. For scalability solution, the distributed averaging DA CNNELM worked with MapReduce concept. CNNELM started with the same weight template afterward trained asynchronously using the partition of training data. Final result obtained by averaging all kernel and ELM output weights. This saved training time instead of single CNNELM trained by the whole data. We studied the backpropagation method to improve the accuracy through iterations. We verified using extended MNIST, not MNIST and CIFAR10 data set. We simulated virtual drift, real drift, and hybrid drift. The DA training divided the training data set to be some smaller partition set. The tools are Deep Learning toolbox with CPU parallel enhancement, and Matconvnet with GPU. The drawbacks need additional learning parameters and the distribution of training data selection.