

PENDAHULUAN

A. Latar Belakang

Beriringan dengan perkembangan zaman yang memasuki zaman data, jumlah informasi yang mencul kepermukaan publik meningkat secara drastis. Informasi-informasi tersebut dipilah oleh peneliti dari berbagai bidang tanpa terkecuali dibidang bioinformatika. Salah satu bentuk data pada bidang ini adalah dataset microarray kanker. Beberapa dari dataset tersebut merupakan dataset yang bisa digunakan untuk pengklasifikasian jenis kanker. Dataset tersebut kemudian dilatih dengan berbagai macam model pembelajaran mesin, salah satunya adalah *Support Vector Machine* (SVM), yang digolongkan kedalam kelompok Supervised Learning. Dimana Algoritma ini akan mengklasifikasikan data input baru ke kelas yang sesuai dengan target kelas berdasarkan hasil *learning* di tahap sebelumnya. Algoritma SVM ini dikenal dengan komputasinya yang berat bagi CPU[1].

Di sisi lain perkembangan teknologi perangkat keras untuk mengatasi permasalahan ini juga berkembang pesat. Ditandai dengan munculnya CUDA (Compute Unified Device Array) yang dikembangkan oleh perusahaan NVIDIA. CUDA merupakan sebuah API (*Application Programming Interface*) yang digunakan memperkerjakan kartu grafis sehingga ia mampu menjalankan perintah serupa yang bisa dijalankan pada perangkat CPU. Dibandingkan dengan CPU, GPU mempunyai satu kelebihan, yaitu mampu untuk menjalankan satu perintah dalam satu kali siklus pada banyak *thread* sekaligus. Keunggulan ini yang kemudian hari oleh para *Software Engineer* untuk mengatasi berbagai permasalahan komputasi yang bisa diselesaikan secara paralel diantaranya operasi aljabar linear. Oleh karena itu dalam tulisan akan dipaparkan bagaimana performa algoritma SVM yang

diimplementasikan pada GPU serta dikombinasikan dengan konsep Paralel SVM dengan dekomposisi SMO.

B. Topik dan Batasannya

Pada penelitian tugas akhir ini akan dibangun suatu program SVM dengan konsep dekomposisi SMO yang kemudian diimplementasikan ke perangkat GPU. Program tersebut akan melatih dirinya terhadap data tersebut agar mampu untuk melakukan klasifikasi menggunakan data uji. Kemudian waktu yang digunakan baik untuk latih maupun uji akan dibandingkan dengan dengan sistem serupa yang dijalankan menggunakan CPU. Sedangkan dataset yang akan digunakan adalah dataset yang diambil dari Elvira Biomedical Dataset Repository[2, 3, 4, 5]. Implementasi PSVM dengan CPU menggunakan library scikit-learn[6] yang berlandaskan libSVM[7], sedangkan GPU menggunakan ThunderSVM. Pengujian dilakukan di lingkungan Google Collaboratory dengan spesifikasi lingkungan:

- CPU: Intel(R) Xeon(R) CPU @ 2.30GHz
- GPU: Tesla K80
- Jumlah core tersedia: 1
- Jumlah thread per core : 2
- Jumlah RAM tersedia : 13 GB

C. Tujuan

Berdasarkan latar belakang yang telah dikemukakan sebelumnya maka tujuan dari penelitan tugas akhir ini dapat dirumuskan sebagai berikut:

- Membuat PSVM dan SVM untuk mengklasifikasi kanker menggunakan data *microarray*.
- Membuat PSVM dan SVM yang diimplementasikan menggunakan GPU.

Bagaimana percepatan (atau perlambatan jika percepatan yang didapatkan berkisar diantara 0 hingga 1) algoritma SVM dengan dekomposisi SMO jika diimplementasikan pada GPU?

D. Organisasi Tulisan

Organisasi penulisan tugas akhir ini akan diurutkan sebagai berikut. Bagian 1 ialah Pendahuluan yang memuat latar belakang, topik, batasan permasalahan, dan tujuan dari penelitian. Bagian 2 adalah penelitian yang telah dilakukan sebagai materi acuan dari metode yang digunakan. Bagian 3 adalah penjelasan tentang sistem yang telah dibangun. Pada bagian 4 pemaparan hasil eksperimen dari sistem. Dan pada bagian 5 ialah kesimpulan dari hasil penelitian yang telah dilakukan.