# Laporan Program Cidar

Cidar (Citra Darah) merupakan program yang berfungsi mengklasifikasi serta menghitung sel darah putih manusia. Program ini menggunakan prinsip pengolahan citra dan kecerdasan buatan. Berikut adalah spesifikasi program Cidar

1. Bahasa Pemrograman: Python
2. Pustaka program:
   1. Scikit-Learn
   2. Tkinter
   3. Tensorflow
   4. OpenCV
   5. Numpy
   6. Pickle
3. Kompatibilitas: Segala sistem operasi, selama terinstal Python + Pustaka Program

Program Cidar dibangun melalui dua tahapan, yaitu tahap pelatihan kecerdasan buatan dan implementasi program.

1. Pelatihan Kecerdasan Buatan

Pada tahap ini, dibuat program yang digunakan untuk mengekstraksi fitur citra dan melatih kecerdasan buatan berdasarkan data fitur yang sudah tersedia. Ekstraksi fitur didefinisikan sebagai proses mengambil parameter terukur dari sebuah citra. Misal, dari citra sel darah putih program akan mengambil fitur luas sitoplasma dalam satuan piksel citra. Proses ini mengekstrak 29 jenis fitur, dengan informasi berupa warna, geometri, dan tekstur. Adapun citra yang diambil fiturnya merupakan citra sel darah putih terdiri atas basofil, eosinofil, limfosit, monosit, dan netrofil. Tiap jenis terdiri atas 100 citra dengan pewarnaan Wright dan Giemsa. Citra sel darah putih sudah dipotong (*crop*) secara manual.

Setelah melalui proses ekstraksi fitur, fitur tersebut dikumpulkan dalam satu berkas tabel berformat Comma Separated Value (CSV), lalu dijadikan bahan dalam proses pelatihan kecerdasan buatan. Model kecerdasan buatan yang digunakan adalah Decision Tree, K Nearest Neighbour, Naïve Bayes, Deep Neural Network, Random Forest, dan SVM. Proses pelatihan kecerdasan buatan memiliki beberapa langkah, di antaranya

1. Pemisahan Data

Data tabel dipisah menjadi data latih dan data uji dengan perbandingan 80:20. Data latih digunakan sebagai bahan untuk melatih kecerdasan buatan, sedangkan data uji digunakan untuk mengetes model kecerdasan buatan yang sudah dilatih.

1. Penyekalaan Fitur

Tahap ini berfungsi untuk menyekalakan fitur citra berdasarkan persebaran pada data latih.

1. Reduksi Dimensi

Reduksi dimensi merupakan tahap untuk mengurangi jumlah fitur yang digunakan dengan mengganti fitur yang ada dengan fitur yang dibuat oleh algoritma tertentu. Algoritma tersebut bisa berupa Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), dan KPCA (Kernel PCA). Masing-masing algoritma dibandingkan performanya dengan yang tidak menggunakan reduksi dimensi.

1. Pelatihan Kecerdasan Buatan

Kecerdasan buatan dilatih menggunakan data latih dan *hyperparameter* bawaan. *Hyperparameter* merupakan parameter dalam algoritma kecerdasan buatan yang menentukan proses pelatihannya. Pelatihan kecerdasan buatan dilakukan sebanyak 10 kali *cross validation*. *Cross validation* adalah proses membagi data latih menjadi data latih dan data validasi secara acak, di mana hasil pelatihan menggunakan data latih akan dibandingkan performanya dengan data validasi. Melalui cross validation, akan diketahui rerata dan standar deviasi dari performa kecerdasan buatan secara umum. Model kecerdasan kemudian diuji dengan data uji untuk melihat performa uji program ketika dihadapkan dengan data baru.

1. Optimasi dengan *Grid Search*

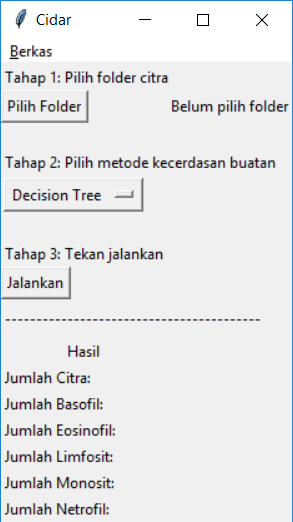
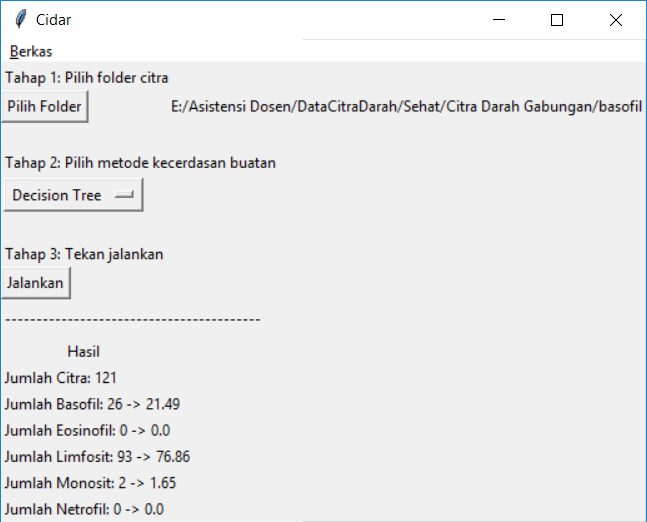
Model kecerdasan buatan akan dilatih ulang dengan mencari *hyperparameter* optimal dengan metode *grid search*. Metode ini membuat algoritma kecerdasan buatan untuk melatih dirinya menggunakan variasi-variasi *hyperparameter* yang tersedia. Kombinasi *hyperparameter* yang memberikan performa terbaik akan digunakan dalam tahap pengujian menggunakan data uji.

1. Penyimpanan hasil kecerdasan buatan

Hasil proses pelatihan kecerdasan buatan adalah model dan skor performa. Model kecerdasan buatan akan disimpan dalam berkas untuk program Cidar, sedangkan skor performa berupa akurasi akan disimpan dalam berkas *spreadsheet*.

1. Implementasi Program

Di tahap ini, program Cidar akan dibuat dengan mengimplementasikan algoritma ekstraksi fitur dan model kecerdasan buatan yang sudah dibuat di tahap sebelumnya ke dalam program. Pada tahap ini pula, dirancang tampilan program serta algoritma pembacaan citra secara banyak. Program Cidar ini mampu membaca banyak citra dalam satu folder, kemudian menghitung jenis sel darah putih yang dikenali dalam folder tersebut. Adapun tampilan program Cidar ditunjukkan pada gambar 1.

Gambar 1. Tampilan Program Cidar Awal (Kiri) dan Saat Berjalan (Kanan)

Hasil akurasi kecerdasan buatan secara detail dapat diakses pada berkas spreadsheet “Hasil\_MachineL\_DataDarah\_100\_auto”.