

KLASIFIKASI SEL DARAH PUTIH LIMFOBLAS DAN NONLIMFOBLAS PADA PASIEN ACUTE LYMPHOBLASTIC LEUKEMIA TIPE L1 DENGAN MENGGUNAKAN METODE *RANDOM FOREST*

Oleh

Reza Irvano Wirawan

15/378793/TK/42735

Diajukan kepada Departemen Teknik Nuklir dan Teknik Fisika Fakultas Teknik
Universitas Gadjah Mada pada tanggal 20 Desember 2019
untuk memenuhi sebagian persyaratan untuk memperoleh derajat
Sarjana Program Studi Teknik Fisika

INTISARI

Acute Lymphoblastic Leukemia (ALL) adalah salah satu jenis kanker darah yang disebabkan oleh berlebihnya produksi sel darah putih muda dan beredar di sistem peredaran darah. Berdasarkan sistem klasifikasi FAB, ALL dibagi ke dalam 3 tipe yaitu tipe L1, L2, dan L3 dengan tipe L1 yang paling banyak diidap oleh pasien ALL. Pengidentifikasian ALL selama ini dilakukan secara manual oleh hematolog atau secara otomatis dengan piranti *Automatic Hematology Analyzer*. Namun piranti tersebut berharga sangat mahal sehingga tidak semua rumah sakit memilikinya. Oleh karena itu dilakukan penelitian dengan bantuan mikroskop termodifikasi dan *machine learning* untuk mengklasifikasi sel darah putih.

Penelitian ini berfokus pada klasifikasi sel darah putih pada ALL tipe L1 ke dalam kelas limfoblas dan nonlimfoblas dengan menggunakan metode *random forest*. Data yang digunakan berjumlah 1737 citra sel darah putih dengan menggunakan 30 fitur. Klasifikasi dilakukan dengan menggunakan *tuning* parameter agar diperoleh model dengan akurasi data terbaik.

Model klasifikasi terbaik menghasilkan akurasi data uji sebesar 91,57% dengan kombinasi 30 fitur, jumlah pohon sebanyak 1800 dan rumus jumlah maksimal fitur *sqrt*. Adapun model terbaik pada data *oversampling* menghasilkan akurasi data uji sebesar 89,46% dan didapat dengan kombinasi 30 fitur, jumlah pohon sebanyak 1100, dan rumus jumlah maksimal fitur *sqrt*. *Oversampling* data menggunakan SMOTE berhasil meningkatkan nilai *recall* pada kelas limfoblas sebesar 6,3%.

Kata kunci: *Acute lymphoblastic leukemia, random forest, klasifikasi, tuning, akurasi.*

Pembimbing Utama : Nopriadi, S.T., M.Sc., Ph.D.

Pembimbing Pendamping : Ir. Agus Arif, M.T.

WHITE BLOOD CELL CLASSIFICATION FOR LYMPHOBLAST AND NON-LYMPHOBLAST OF ACUTE LYMPHOBLASTIC LEUKEMIA TYPE L1 PATIENT USING RANDOM FOREST

by

Reza Irvano Wirawan

15/378793/TK/42735

Submitted to the Departement of Nuclear Engineering and Engineering Physics
Faculty of Engineering Universitas Gadjah Mada on December 20th, 2019
in partial fulfillment of the requirement for the Degree of
Bachelor of Engineering in Engineering Physics

ABSTRACT

Acute Lymphoblastic Leukemia (ALL) is a type of blood cancer caused by excess production of immature white blood cells and circulates in blood circulation. Based on the FAB classification system, ALL is divided into 3 types, namely type L1, L2, and L3 with type L1 which is the most affected by ALL patients. Identification of ALL can be done manually by the hematologist or automatically with the Automatic Hematology Analyzer tool. However, this device is very expensive so not every labs or hospitals have it. Therefore, conducted research with a modified microscope and machine learning to classify white blood cells.

This study focuses on the classification of white blood cells in ALL type L1 into lymphoblast and non-lymphoblast classes using random forest method. The data used were 1737 images of white blood cells using 30 features. The classification was performed by using parameters tuning to get the model with the best data accuracy.

The best model produces an accuracy of test data of 91.57% with a combination of 30 features, 1800 trees, and using sqrt as the maximal number of features. Meanwhile, the best model in oversampling data produces 89.46% of test data accuracy and is obtained with a combination of 30 features, 1100 trees, and using sqrt as the maximal number of features. Data oversampling using SMOTE increase the model recall up to 6.3%.

Keywords: Acute lymphoblastic leukemia, random forest, classification, tuning, accuracy.

Supervisor : Nopriadi, S.T., M.Sc., Ph.D.

Co-supervisor : Ir. Agus Arif, M.T.