



INTISARI
KLASIFIKASI SEL DARAH PUTIH MENGGUNAKAN GAUSSIAN
NAIVE BAYES

Oleh
Ahmad Zaki Hakimi
14/367228/TK/42408

Diajukan kepada Departemen Teknik Nuklir dan Teknik Fisika Fakultas Teknik
Universitas Gadjah Mada pada tanggal 17 Januari 2019
untuk memenuhi sebagian persyaratan untuk memperoleh derajat
Sarjana S-1 Program Studi Teknik Fisika

INTISARI

Informasi jumlah sel darah putih dapat digunakan untuk mengidentifikasi kondisi kesehatan dan berbagai macam jenis penyakit yang menyerang seseorang. Jumlah sel darah putih akan meningkat saat seseorang terkena suatu infeksi sebagai Respon sistem imun tubuh untuk melindungi tubuh. Terdapat instrumen yang dapat melakukan klasifikasi dan perhitungan sel darah putih, namun harga instrumen tersebut relatif mahal sehingga tidak tersedia di setiap pusat kesehatan. Perhitungan sel darah putih dapat dilakukan secara manual oleh seorang ahli, namun akan membutuhkan waktu yang cukup lama. Oleh karena itu dilakukan penelitian untuk mengembangkan mikroskop termodifikasi yang dapat menghitung dan mengklasifikasi sel darah putih dari suatu apusan darah pada suatu preparat. Dalam rangka penelitian tersebut, dilakukan penelitian penerapan *machine learning* dalam melakukan klasifikasi sel darah putih menggunakan algoritma Gaussian naïve Bayes.

Penelitian ini dimulai dengan mengekstrak fitur dari preparat yang berisi apusan darah melalui sebuah komputer. Data akan melalui *preprocessing* berupa standarisasi dan akan dilakukan seleksi fitur lalu diklasifikasi menggunakan algoritma Gaussian naïve Bayes. Algoritma naïve Bayes menghitung probabilitas suatu sampel data fitur tergolong sebagai setiap jenis sel darah putih. Algoritma kemudian mengklasifikasi sampel data fitur sebagai jenis sel darah putih dengan probabilitas tertinggi.

Penerapan algoritma Gaussian naïve Bayes menghasilkan *F1 score* sebesar 91,26%. Untuk mencapai akurasi tersebut, *preprocessing* data berupa standarisasi serta seleksi fitur dengan metode *sequential feature selection* perlu dilakukan. Didapatkan bahwa fitur dari inti sel darah putih sangat penting dalam mengklasifikasi sel darah putih.

Kata kunci: sel darah putih, naïve Bayes, klasifikasi, standarisasi, *sequential feature selection*

Pembimbing Utama : Nopriadi, S.T., M.Sc., Ph. D
Pembimbing Pendamping : Ir. Agus Arif, M.T.



ABSTRACT
WHITE BLOOD CELL CLASSIFICATION USING GAUSSIAN NAÏVE BAYES

by
Ahmad Zaki Hakimi
14/367228/TK/42408

Submitted to the Department of Nuclear Engineering and Engineering Physics
Faculty of Engineering Universitas Gadjah Mada on 17th January 2019
in partial fulfillment of the requirement for the Degree of
Bachelor of Engineering in Physics Engineering

ABSTRACT

The number of white blood cells present in the blood can be used to identify certain health conditions and various types of diseases. The number of white blood cells present in blood circulation increase when someone is exposed to infections as a response from the body's immune system. Present time, there is already an instrument that is able to classify and count white blood cell, but its price is relatively expensive and is not available in many health centers. Calculation of white blood cells can be done manually by a trained expert, but time-consuming and can lead to human error. Therefore, a study was conducted to design a modified microscope that is able to calculate and classify white blood cells from a blood smear on a glass slide. In the framework of this research, the Gaussian naïve Bayes, a machine learning algorithm, is implemented to classify white blood cells.

This research starts with feature extraction of blood smears from glass slides using a computer. The data will be preprocessed using standardization techniques and feature selection will be done then the data will be classified using the Gaussian naïve Bayes algorithm. The algorithm calculates the probability of a sample feature being a certain type of white blood cell then classifies it to the type of blood white blood cell which has the highest probability.

The implementation of the Gaussian naïve Bayes achieved an F1 score of 91,26%. Optimization were added to the algorithm to produce this accuracy which were data preprocessing in the form of standardization and feature selection. It was found that features of the white blood cell nucleus are a major factor in the classification of white blood cells.

Keywords: white blood cells, naïve Bayes, classification, standardization, feature selection

Supervisor : Nopriadi, S.T., M.Sc., Ph. D
Co-supervisor : Ir. Agus Arif, M.T.