

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Leukemia atau kanker darah merupakan penyakit sel darah yang disebabkan oleh berkembangnya sel darah putih (leukosit) secara tidak semestinya, dimana sel baru tersebut tidak dapat matang seperti normalnya (Yayasan Kanker Indonesia (YKI), 2008). Leukemia menyerang bagian tubuh manusia. Bagian tubuh dari penderita akan terinfeksi mengalami gangguan pada sel darah putihnya sehingga menyebabkan sistem imun mengalami gangguan. Penyakit leukimia memiliki dua klasifikasi yaitu klasifikasi granulocytes dan klasifikasi monocytes.

Dibidang kesehatan, mendiagnosis penyakit leukemia merupakan suatu hal yang sulit karena kurangnya peralatan yang dapat mendeteksi penyakit leukimia dan masih didiagnosis secara manual dengan bantuan dokter. Diagnosis manual tersebut dapat mengalami kesalahan yang disebabkan oleh kelalaian manusia. Dari masalah tersebut, diagnosis jenis penyakit leukemia dapat dibantu oleh kecanggihan teknologi yaitu kecerdasan buatan.

Kecerdasan buatan telah dimanfaatkan oleh banyak orang untuk membuat keputusan melalui komputer otomatis. Dalam kasus ini, kecerdasan buatan tersebut akan mengolah data yang berasal dari ciri morfologi gen sel darah putih. Dengan adanya teknologi komputer otomatis, kecerdasan buatan dapat mendiagnosis penyakit leukemia menggunakan teknik klasifikasi seperti metode Naïve Bayes dan metode K-Nearest Neighbor (KNN).

Metode pengklasifikasian data yang dapat digunakan untuk penelitian ini salah satunya yaitu K-Nearest Neighbor (KNN), metode tersebut merupakan algoritma klasifikasi yang menggunakan kedekatan jarak antara data yang satu dengan data yang lainnya (Mustafa & Simpen, 2019). Pada penelitian sebelumnya yang dilakukan oleh Khasanah et.al. berdasarkan ciri warna dan

Muhammad Nur'adli Hasbi Gumay, 2022

Perbandingan Metode Naïve Bayes dan K-Nearest Neighbor Pada Klasifikasi Morfologi Gen Sel Darah Putih

UPN Veteran Jakarta, Fakultas Ilmu Komputer, S-1 Informatika

[www.upnvj.ac.id – www.library.upnvj.ac.id - www.repository.upnvj.ac.id]

bentuk dengan metode K-Nearest Neighbor (KNN) mengenai pengklasifikasian sel darah putih dan menghasilkan nilai akurasi dari segmentasi sel darah putih sebesar 78% dan nilai akurasi yang menggunakan confusion matrix sebesar 64%. Penelitian tersebut telah berhasil melakukan klasifikasi sel darah putih dengan metode K-Nearest Neighbor (KNN). Akan tetapi, dalam penelitian tersebut masih perlu untuk menambahkan ekstraksi ciri lainnya misal seperti berdasarkan tekstur dan morfologi sehingga penelitian ini menggunakan metode K-Nearest Neighbor (KNN) dengan Naïve Bayes.

Dalam penelitian ini, metode lainnya yang digunakan yaitu metode Naïve Bayes. Dalam memperkirakan masa yang akan datang, Naïve Bayes bisa mempelajari data yang akan diolah dari data sebelumnya. Klasifikasi Naïve Bayes adalah cara pemodelan yang mudah dan memiliki noise sedikit (M. M. Kini et al., 2015). Metode Naïve Bayes bekerja sangat baik daripada model classifier lainnya. Hal ini dibuktikan oleh xhemali, hinde stone dalam jurnal penelitiannya yang berjudul “NB vs Decision Tree vs Neural Network in the classification of training” menyatakan bahwa Naïve bayes mempunyai tingkat keakuratan yang lebih baik daripada model classifier yang lain. Metode Naïve Bayes sudah pernah diterapkan untuk sel darah putih yaitu pada penelitian sebelumnya yang dilakukan oleh Shidada dan Hariyanto mengenai identifikasi Acute Lymphoblastic Leukemia pada citra mikroskopis menggunakan algoritma Naïve Bayes menghasilkan nilai akurasi sebesar 80% dari 25 data uji. Namun pada penelitian tersebut, proses segmentasi tidak hanya bisa dilakukan menggunakan thresholding melainkan dapat menggunakan proses segmentasi yang lain.

Berdasarkan penelitian sebelumnya dan permasalahan yang ada, peneliti bermaksud untuk membandingkan metode Naïve Bayes dan metode K-NN untuk klasifikasi ciri morfologi gen sel darah putih. Dengan menggunakan dua metode tersebut, nantinya akan melihat nilai hasil tingkat akurasi dari kedua metode tersebut dan setelah itu dilakukan perbandingan dari hasil akurasi kedua

metode tersebut untuk menentukan nilai optimal dan metode terbaik untuk mengolah data ini.

1.2 Rumusan Masalah

Berdasarkan uraian diatas, maka dapat dideskripsikan rumusan masalah sebagai berikut:

1. Apakah metode Naïve Bayes dan metode K-NN dapat diaplikasikan terhadap *dataset* sel darah putih?
2. Bagaimana perbandingan nilai akurasi pada prediksi dan klasifikasi ciri morfologi gen sel darah putih?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini yaitu untuk memprediksi performa akurasi dari metode KNN dan metode Naive Bayes serta untuk mengetahui metode apa yang tepat dalam klasifikasi sel darah putih.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini yaitu mengetahui hasil dan perbandingan dari algoritma K-NN dan Naïve Bayes dalam mengklasifikasi jenis sel kanker berdasarkan ciri morfologi gen pada sel darah putih dan mengetahui algoritma mana yang tepat.

1.5 Ruang Lingkup

Dalam menyusun Laporan Tugas Akhir, peneliti menentukan ruang lingkup agar pembahasan lebih terarah. Adapun permasalahan yang akan dibahas dalam penelitian ini berikut ini yaitu:

1. Metode yang digunakan pada penelitian ini menggunakan metode K-NN dan metode Naive Bayes.
2. Penelitian ini hanya untuk meneliti penyakit pada *Leukemia* yang memiliki jenis *Acute Lymphoblastic Leukemia* (ALL) dan jenis *Acute Myeloid Leukemia* (AML).
3. Praproses pada penelitian ini berdasarkan ciri dari morfologi gen yang terdapat pada inti sel darah putih yang dapat dilihat dari *dataset* yang sudah di konversi menjadi bentuk *Comma Separated Value Files* (CSV).

1.6 Sistematika Penulisan

Dalam penulisan laporan penelitian ini akan menjelaskan isi menurut sistematika penulisan yang dapat diuraikan sebagai berikut

BAB I PENDAHULUAN

Pada bab 1 akan membahas latar belakang berdasarkan permasalahan yang ada. Selain itu membahas rumusan masalah sesuai dengan latar belakang, menentukan batasan masalah, tujuan dan manfaat penelitian ini.

BAB II TINJAUAN PUSTAKA

Bab ini berisi penjelasan teori mendasar yang digunakan sebagai landasan dalam penelitian ini.

BAB III METODE PENELITIAN

Bab ini menjelaskan mengenai tahapan-tahapan dan segala metode yang akan diterapkan dalam menyusun penelitian untuk mencapai tujuan penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan hasil dan pembahasan dari penelitian yang telah dilakukan berupa data-data percobaan dan hasil analisis dari proses dan metode yang telah dibahas pada Bab III.

BAB V PENUTUP

Bab ini menjelaskan mengenai kesimpulan dan saran dari penelitian yang telah dikerjakan.

DAFTAR PUSTAKA