

BAB 2 LANDASAN TEORI

2.1 Precision Medicine

Pengobatan presisi adalah sebuah konsep modern yang telah digunakan untuk menggambarkan pengobatan medis yang akurat yang disesuaikan dengan karakteristik individu setiap pasien, dimana penyakit pasien dianalisis menurut data molekular, genomik, dan model sistem biologi untuk menentukan penyakit pasien pada tingkat molekular dan memilih pengobatan yang tepat. Pengobatan presisi menjanjikan untuk mengkombinasikan data individual mengenai kerentanan genetik terhadap penyakit, data biomarker tentang risiko dan respons penyakit, juga data fisiologis dan perilaku dari sensor dan database baru untuk menciptakan sebuah model yang dapat memprediksi risiko dan respons dengan tingkat akurasi yang tinggi. Semakin banyak ilmuwan yang menggunakan istilah '*precision medicine*' karena memanfaatkan genomik untuk mengkarakterisasi penyakit untuk terapi individual. dengan menggunakan informasi genetik dan database lainnya dari sebuah perspektif *big data*, diabetes dan penyakit lainnya diklasifikasikan menjadi subgrup penyakit, masing-masing dengan pengobatan terbaiknya [26].

2.2 Penyakit Ginjal Kronis (PGK)

Penyakit gagal ginjal adalah kondisi ketika ginjal kehilangan kemampuannya untuk menyaring cairan, toksin, maupun sisa-sisa hasil metabolisme. Penyakit gagal ginjal dapat dibedakan menjadi 2 jenis, yaitu gangguan ginjal akut atau *Acute Kidney Injury* (AKI) dan Penyakit Ginjal Kronis (PGK). Seseorang dikatakan mengalami AKI apabila fungsi kerja ginjalnya sudah menurun atau mengalami kerusakan berat namun baru terjadi selama beberapa jam atau beberapa hari. Sedangkan PGK merupakan kondisi saat terjadinya penurunan progresif fungsi ginjal yang sudah terjadi selama beberapa bulan atau tahun bahkan permanen. Penurunan progresif yang dimaksud adalah terjadinya kerusakan ginjal dan penurunan *Glomerular Filtration Rate* (GFR) yang kurang dari 60mL/min/1.73 m² selama minimal 3 bulan [27].

Beberapa cara yang dapat dilakukan untuk mendeteksi seberapa genting kerusakan fungsi ginjal, salah satunya adalah pemeriksaan kadar albumin darah dalam tubuh manusia (*blood albumin amount* atau *serum albumin*). Perubahan

kadar albumin darah, yang biasanya menurun, dapat diakibatkan oleh kurangnya protein, kalori maupun vitamin yang dikonsumsi oleh tubuh. Albumin sendiri memiliki peran untuk membantu pencegahan bocornya air dari darah ke jaringan lainnya. Akan terjadi hipoalbuminemia, kurangnya kadar albumin dalam tubuh, apabila kadar albumin darah kurang dari 3.5 g/dL. Apabila penurunan kadar albumin terjadi, maka akibatnya akan terjadi pergeseran cairan dalam ruang vaskuler, sehingga terdapat penumpukan cairan yang dapat memperparah pasien gagal ginjal. [28].

2.3 Genotipe dan Fenotipe

Genotipe adalah kumpulan lengkap dari gen dalam DNA yang dimiliki oleh suatu individu. Genotipe didefinisikan oleh keunikan DNA yang dimiliki setiap individu, merujuk pada 2 alel yang diwariskan seseorang untuk gen tertentu. Genotipe juga berperan dalam menentukan perbedaan karakteristik, penampilan, dan perilaku tertentu pada suatu individu. Sedangkan fenotipe adalah ekspresi genotipe seseorang yang dapat diamati karena mengacu pada penampilan fisik atau karakteristik suatu individu. Fenotipe dibentuk oleh genotipe yang juga dipengaruhi oleh faktor lingkungan. Fenotipe dapat diamati melalui berat badan, warna mata, dan golongan darah suatu individu [29].

2.4 Single Nucleotide Polymorphism (SNP)

Single Nucleotide Polymorphism (SNP) adalah variasi urutan DNA yang dihasilkan dari perubahan nukleotida tunggal dalam genom dan merupakan sumber penting dari variabilitas gen manusia [13]. Sebagian besar perbedaan genetik pada manusia dipengaruhi oleh adanya SNP yang terjadi pada genom manusia, dan seringkali dihubungkan dengan adanya perbedaan dalam predisposisi suatu penyakit tertentu ataupun respons tubuh terhadap penggunaan obat [30]. SNP terdiri atas Adenina (A), Sitosina (S), Guanina (G), dan Timina (T) pada rangkaian DNA yang diamati dengan rangkaian DNA referensi [11]. SNP merupakan sumber polimorfisme genetik yang paling lengkap. Terdapat tiga bentuk variasi SNP yaitu *transition*, *transversion*, dan indel (*insertion* dan *deletion*). *Transition* adalah perubahan C menjadi T, atau G menjadi A, dan sebaliknya. *Transversion* adalah perubahan C menjadi G, A menjadi T, C menjadi A, atau T menjadi G, dan sebaliknya. Indel dapat berupa penyisipan atau penghapusan nukleotida tunggal

[31].

2.5 Elastic Net

Tujuan utama dari regresi *Elastic Net* adalah untuk menemukan koefisien yang meminimalkan jumlah kuadrat kesalahan dengan menerapkan penalti pada koefisien tersebut. *Elastic Net* menggabungkan pendekatan L1 ($\sum_{j=1}^p |\beta_j|$) dan L2 ($\sum_{j=1}^p (\beta_j)^2$) (*Lasso* dan *Ridge*). Sehingga *Elastic Net* melakukan proses pemerataan yang lebih efisien. *Elastic Net* mengatasi kekurangan *Lasso Regression* apabila terdapat sekelompok peubah yang sangat berkorelasi, maka *Lasso Regression* cenderung memilih satu peubah dari kelompok dan mengabaikan yang lain. *Elastic Net* menangani kekurangan tersebut dengan menambahkan bagian kuadrat *penalty* $\|\beta\|^2$ yang apabila digunakan menjadi *Ridge Regression* [32]. Secara umum, *Elastic Net* dirumuskan sebagai berikut [33]:

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda_2 \sum_{j=1}^p (\beta_j)^2 + \lambda_1 \sum_{j=1}^p |\beta_j| \dots\dots\dots(1)$$

$$\text{dengan } \alpha = \frac{\lambda_1}{\lambda_1 + \lambda_2}, 0 \leq \alpha \leq 1$$

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda [(1 - \alpha) \sum_{j=1}^p (\beta_j)^2 + \alpha \sum_{j=1}^p |\beta_j|] \dots\dots\dots(2)$$

Dengan α adalah parameter gabungan. Jika $\alpha = 0$ maka akan menjadi fungsi *Ridge Regression*, sedangkan jika $\alpha = 1$ maka akan menjadi fungsi *Lasso Regression*. λ merupakan konstanta penalti dengan nilai default = 1. Jika $\lambda = 0$, maka fungsi akan sepenuhnya menjadi *Ordinary Least Square* (OLS). β_j merupakan koefisien regresi (slope atau kemiringan), β_0 merupakan konstanta (intersep). x merupakan nilai data (variabel independen), dan y merupakan nilai target (variabel dependen).

Elastic Net mengadaptasikan fitur pada *Ridge Regression* yang mampu menekan koefisien fitur yang tidak signifikan mendekati nol, dengan fitur pada *Lasso Regression* yang mampu menghilangkan fitur yang tidak signifikan dengan menjadikan koefisiennya sama dengan nol, sehingga membuat *Elastic Net* dapat memprediksi sekaligus melakukan seleksi fitur dengan baik.

2.6 Mean Absolute Error

Mean Absolute Error (MAE) digunakan untuk melakukan evaluasi performa model *Elastic Net* yang telah dihasilkan melalui penelitian ini. Jika nilai MAE

mendekati angka nol, maka model akan dianggap semakin baik. Adapun berikut persamaan yang digunakan untuk menghitung nilai MAE [34].

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| \dots \dots \dots (3)$$

- n = jumlah data
- f_i = nilai hasil prediksi ke-i
- y_i = nilai sebenarnya

2.7 Explained Variance (R^2)

Koefisien determinasi (R^2) sering digunakan pada pengaplikasian statistika sebagai penentu apakah model regresi yang digunakan sudah cukup baik atau tidak. Pemodelan dengan nilai R^2 yang mendekati 1 dianggap sudah cukup baik, sedangkan jika nilai R^2 mendekati 0 dianggap model regresi buruk [35]. Berikut merupakan persamaan untuk menghitung nilai (R^2) [36].

$$R^2 = 1 - \frac{\sum (y_i - y_c)^2}{\sum (y_i - y_r)^2} \dots \dots \dots (4)$$

- y_i = nilai ke-i
- y_c = nilai estimasi dari persamaan regresi
- y_r = nilai rata-rata

