

Pemodelan Prediksi Kadar Gula Darah pada Pasien Diabetes Menggunakan Metode Regresi Linear

Hanan Hanafi¹, Ganjar Alfian^{1,*}, Tri Widodo², Muhammad Syafrudin³

¹Departemen Teknik Elektro dan Informatika, Sekolah Vokasi, Universitas Gadjah Mada, Indonesia;
hananhanafi@mail.ugm.ac.id

²Program Studi Pendidikan Teknologi Informasi, Universitas Teknologi Yogyakarta, Indonesia;
triwido@uty.ac.id

³Department of Artificial Intelligence, Sejong University, Republic of Korea;
udin@sejong.ac.kr

*Korespondensi: ganjar.alfian@ugm.ac.id;

Abstract – Predicting blood glucose levels in the future can help diabetic patients to take preventive action earlier so that they can control their blood glucose levels. This study proposed a blood glucose levels prediction model using linear regression method. Time series data of blood glucose levels from 30 type 1 diabetic patients were used as a single input parameter with the additional utilization of statistical features. The dataset was tested using several algorithms, namely Linear Regression, Lasso Regression, Ridge Regression, eXtreme Gradient Boosting (XGB), and K-Nearest Neighbor (KNN). The prediction models have been tested and showed that the Linear Regression method outperformed other models with average Root Mean Square Error (RMSE) values of 15 patients tested were 5.024, 12.488, and 20.635, Mean Absolute Error (MAE) values were 2.891, 8.272, and 14.926 and Coefficient of Determination (R^2) values were 0.962, 0.741, and 0.39 for the Prediction Horizon (PH) of 5 minutes, 15 minutes, and 30 minutes, respectively. The results of the prediction model in this study were implemented in a website-based information system. In this system, the user can predict blood glucose levels in the future by utilizing a history of blood glucose levels in the previous 30 minutes. Users can also see the visualization of data based on the movement of blood glucose levels with different time interval. This system is expected to help diabetic patients to predict their future blood glucose levels so that they can control their blood glucose levels and avoid bad health conditions in the future.

Keywords – Prediction Model, Blood Glucose, Diabetes, Linear Regression

Intisari – Mengetahui kadar gula darah di masa depan akan dapat membantu penderita diabetes dalam melakukan tindakan preventif lebih awal, sehingga dapat mengontrol kadar gula darah dan penyakit diabetesnya. Penelitian ini mengusulkan pemodelan prediksi kadar gula darah pada pasien diabetes menggunakan metode regresi linear. Dataset yang digunakan adalah data *time series* dari kadar gula darah pada 30 anak penderita diabetes tipe 1. Dataset tersebut digunakan sebagai parameter *input* tunggal dengan tambahan pemanfaatan fitur statistik yang diuji menggunakan beberapa algoritma, yaitu *Linear Regression*, *Lasso Regression*, *Ridge Regression*, *eXtreme Gradient Boosting (XGB)*, dan *K-Nearest Neighbor (KNN)*. Pada tahap evaluasi performa model menunjukkan bahwa metode regresi linear lebih baik dari model prediksi lainnya. Hasilnya menunjukkan untuk *Prediction Horizon (PH)* pada 5 menit, 15 menit, dan 30 menit didapat nilai rata-rata *Root Mean Squared Error (RMSE)* dari 15 pasien yang diuji sebesar 5,024, 12,488, dan 20,635, nilai *Mean Absolute Error (MAE)* sebesar 2,891, 8,272, dan 14,926 serta nilai *Coefficient of Determination (R^2)* sebesar 0,962, 0,741, dan 0,39. Hasil model prediksi pada penelitian ini diimplementasi dan divisualisasikan ke sistem informasi berbasis *website*. Dalam sistem tersebut pengguna dapat memprediksi kadar gula darah di masa depan dengan berdasarkan riwayat kadar gula darah pada waktu 30 menit sebelumnya. Pengguna juga dapat melihat visualisasi data pergerakan kadar gula darah berdasarkan rentang waktu tertentu. Sistem ini diharapkan dapat membantu pasien diabetes untuk memprediksi kadar gula darah pada masa depan, sehingga dapat mengontrol kadar gula darahnya dan menghindari kondisi kesehatan yang buruk pada masa depan.

Kata kunci – Model Prediksi, Gula Darah, Diabetes, Regresi Linear

I. PENDAHULUAN

Diabetes adalah penyakit kronis yang disebabkan oleh ketidakmampuan tubuh untuk memproduksi hormon insulin atau karena penggunaan yang tidak efektif dari produksi insulin [1]. Berdasarkan klasifikasi saat ini, diabetes terbagi menjadi 2 jenis yaitu diabetes tipe 1 dan diabetes tipe 2 [2]. Diabetes ditandai dengan tingginya kadar gula dalam darah dan dapat menyebabkan kerusakan serius pada jantung, pembuluh darah, ginjal, dan saraf. Gejala diabetes ditandai dengan mudah haus, sering buang air kecil, cepat lelah, penglihatan kabur, dan berat badan yang turun. Diabetes

adalah ancaman yang serius bagi kesehatan dunia yang merupakan salah satu dari empat prioritas penyakit tidak menular yang ditindaklanjuti oleh para pemimpin dunia [3]. Diabetes merupakan peringkat ke sembilan penyebab kematian terbanyak secara global pada Tahun 2019, meningkat sebesar 70% sejak Tahun 2000 [4]. Secara global, diperkirakan 422 juta orang dewasa hidup dengan diabetes pada Tahun 2014. Selama dekade terakhir, prevalensi diabetes meningkat lebih besar di negara-negara berpenghasilan rendah dan menengah dibandingkan negara berpenghasilan tinggi. Diabetes menyebabkan 1,5 juta kematian pada tahun 2012 serta tambahan 2,2 juta kematian

dari penyakit kardiovaskular, ginjal kronis, dan tuberkulosis yang berhubungan atau komplikasi dengan diabetes [3].

Diabetes tipe 1 tidak dapat dicegah dengan pengetahuan saat ini, sehingga hal efektif yang dapat dilakukan oleh pasien atau penderita diabetes tipe 1 adalah mencegah diabetes tipe 2 dan komplikasinya seperti berolahraga secara teratur, makan makanan sehat, menghindari rokok, dan mengontrol tekanan darah [3]. Mengetahui prediksi kadar gula darah di masa depan juga dapat membantu seorang pasien atau penderita diabetes untuk mengontrol kadar gula darahnya dan melakukan tindakan preventif lebih awal, sehingga dapat menghindari kondisi yang fatal di masa yang akan datang.

Memprediksi kadar gula darah di masa depan dapat menggunakan metode berbasis *machine learning*. Metode *machine learning* sudah banyak digunakan di berbagai area dan dapat meningkatkan performa sistem, seperti pada area *Internet of Things* [5] dan *Natural Language Processing* [6]. Penelitian sebelumnya menunjukkan bahwa sistem prediksi kadar gula darah pada pasien diabetes dapat dibangun menggunakan metode *machine learning*. Pappada dkk [7] mengusulkan model prediksi kadar gula darah pada pasien diabetes menggunakan *Neural Network* (NN) dengan memasukkan data 10 pasien menggunakan data nilai *Continuous Glucose Monitoring* (CGM), dosis insulin, nilai meteran glukosa, asupan gizi, gaya hidup, dan faktor emosional. Pada penelitiannya menggunakan *Prediction Horizon* (PH) 75 menit yang menghasilkan nilai RMSE 43,9 mg/dL.

Penelitian ini menjelaskan pentingnya penggunaan metode *machine learning*, khususnya model prediksi berbasis regresi linear, dalam meningkatkan efektivitas pemantauan kadar gula darah pada penderita diabetes. Keunggulan penelitian ini terletak pada kemampuan model prediksi untuk memberikan estimasi akurat mengenai tingkat gula darah masa depan. Dengan mengintegrasikan model terlatih (*trained model*) ini ke dalam sistem monitoring gula darah berbasis web, penelitian ini tidak hanya menciptakan solusi yang lebih canggih, tetapi juga praktis. Integrasi tersebut akan memungkinkan pemantauan gula darah secara *real-time*, memberikan informasi yang sangat berharga bagi pasien dan tenaga medis. Kemampuan untuk memprediksi masa depan membuka pintu bagi intervensi yang lebih dini dan personalisasi perawatan, yang pada gilirannya dapat mengoptimalkan manajemen penyakit diabetes.

II. DASAR TEORI

Penelitian sebelumnya telah menghasilkan temuan yang signifikan, menunjukkan bahwa penerapan model prediksi menggunakan algoritma *machine learning* dapat memberikan hasil yang akurat dalam memprediksi kadar gula darah. Dengan memanfaatkan data pasien yang luas dan beragam, seperti riwayat medis, pola makan, dan faktor gaya hidup, metode ini mampu memberikan informasi yang berharga untuk mendukung upaya diagnosis dan pengelolaan diabetes secara lebih efektif. Georga dkk [8] mengusulkan menggunakan model prediksi dengan metode *Support Vector*

Regression (SVR) untuk prediksi gula darah pada pasien diabetes tipe 1. Model yang diusulkan tersebut menggunakan data gula darah, insulin, makanan, serta aktivitas fisik dari 27 pasien diabetes tipe 1. Evaluasi model tersebut menggunakan *Prediction Horizon* (PH) 15 menit, 30 menit, dan 60 menit dengan hasil RMSE masing-masing didapat 5,21 mg/dL, 6,03 mg/dL, dan 7,14 mg/dL.

Penelitian Ben Ali dkk [9] mengusulkan model prediksi kadar gula darah pada pasien diabetes tipe 1 menggunakan *Artificial Neural Network* (ANN) hanya dengan menggunakan masukkan *Continuous Glucose Monitoring* (CGM) sebagai masukkan tunggal. Model tersebut diuji dengan menggunakan data dari 13 pasien diabetes tipe 1. Hasil evaluasi model prediksi tersebut didapat nilai RMSE 6,43 mg/dL, 7,46 mg/dL, dan 8,13 mg/dL, dan 9,03 mg/dL dengan *Prediction Horizon* (PH) 15 menit, 30 menit, 45 menit, dan 60 menit.

Penelitian Hamdi dkk [10] menggunakan model prediksi dengan metode *Support Vector Regression* (SVR) untuk prediksi kadar gula darah pada pasien diabetes tipe 1. Pada penelitiannya model tersebut menggunakan data 12 pasien dengan memasukkan hanya menggunakan data kadar gula darah atau *Continuous Glucose Monitoring* (CGM) sebagai masukkan tunggal model. Hasil evaluasi performa model tersebut didapat nilai RMSE 9,44 mg/dL untuk *Prediction Horizon* (PH) 15 menit, 10,78 untuk *Prediction Horizon* (PH) 30 menit, 11,82 mg/dL untuk *Prediction Horizon* (PH) 45 menit, dan 12,95 mg/dL untuk *Prediction Horizon* (PH) 60 menit.

Penelitian lain dilakukan oleh Martinsson dkk [11] mengusulkan model prediksi kadar gula darah menggunakan *Recurrent Neural Network* dengan metode *Long Short-Term Memory* (LSTM). Model tersebut menggunakan data kadar gula darah atau *Continuous Glucose Monitoring* (CGM) dari 6 pasien. Performa model dievaluasi dengan menggunakan *Prediction Horizon* (PH) 30 menit dan 60 menit. Evaluasi tersebut didapat nilai rerata RMSE 18,87 mg/dL pada prediksi 30 menit dan 31,40 mg/dL pada prediksi 60 menit.

Alfian dkk [12] mengusulkan model pembelajaran mesin untuk memprediksi kadar gula darah pada pasien diabetes menggunakan metode *Artificial Neural Network* (ANN). Dalam penelitiannya tersebut menggunakan data kadar gula darah pada pasien diabetes tipe 1 untuk memprediksi kadar gula darah pada *Prediction Horizon* (PH) 30 menit dan 60 menit selanjutnya. Model prediksi tersebut diuji kepada 12 anak yang mengalami Diabetes tipe 1. Berdasarkan hasil pengujiannya tersebut didapat model ANN mengungguli model lain dengan rata-rata RMSE 2,82 mg/dL, 6,31 mg/dL, 10,65 mg/dL, dan 15,33 mg/dL untuk *Prediction Horizon* (PH) 15 menit, 30 menit, 45 menit, dan 60 menit.

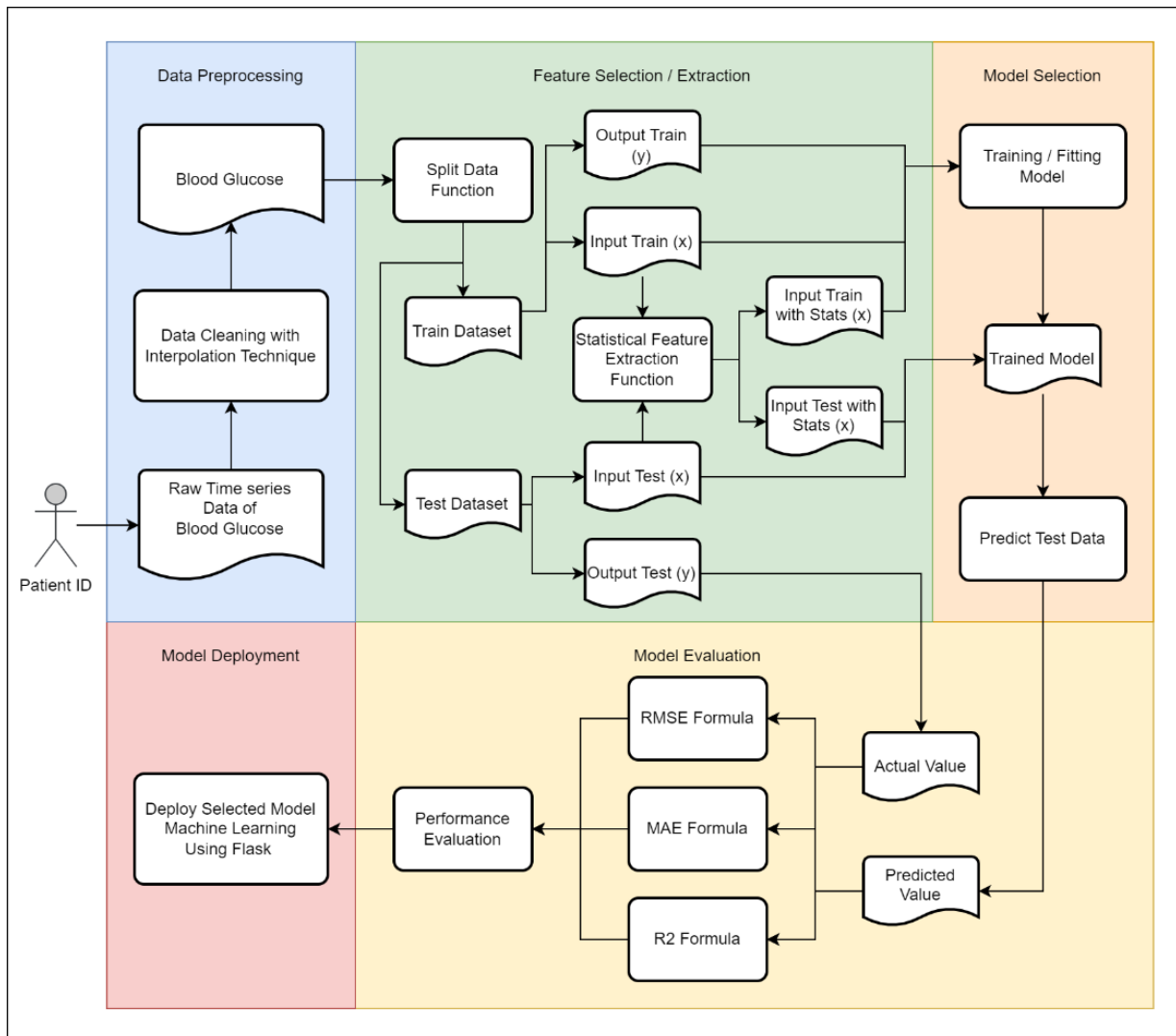
Penelitian ini mengusulkan model prediksi kadar gula darah pada pasien diabetes menggunakan metode *Linear Regression*. Kemudian hasil analisis dan pengembangan

model prediksi pada penelitian ini diimplementasi ke sistem informasi berbasis *website*. Pengguna juga dapat melihat visualisasi data pergerakan kadar gula darah berdasarkan waktu dalam bentuk diagram. Sistem ini diharapkan dapat membantu pasien diabetes untuk memprediksi kadar gula darah pada masa depan, sehingga dapat menghindari hal buruk atau kondisi yang fatal bagi penderita diabetes pada masa depan.

III. METODOLOGI

Pada penelitian ini menggunakan dataset berbentuk *time series* atau historis dari kadar gula darah pada 30 pasien anak diabetes tipe 1. Dataset tersebut berasal dari jurnal penelitian yang berjudul *Evaluation of Counter-regulatory Hormone Responses during Hypoglycemia and the Accuracy of Continuous Glucose Monitors in Children with T1DM* yang diterbitkan oleh Jaeb Center for Health Research (JCHR). Dataset tersebut digunakan sebagai parameter *input* tunggal dengan tambahan pemanfaatan data statistik atau *statistical feature* dalam jarak waktu tertentu (*window*).

Membangun sistem prediksi kadar gula darah pada pasien diabetes menggunakan metode regresi linear pada tahap awalnya adalah membuat model prediksi. Pada tahapan pembuatan model prediksi dilakukan analisis dan perancangan model prediksi. Analisis awal yang dilakukan adalah menganalisis data yang akan digunakan untuk membuat model prediksi dengan melakukan pra-pemrosesan data dan ekstraksi fitur untuk mendapatkan data masukan dan keluaran yang diinginkan. Hasil ekstraksi fitur tersebut digunakan untuk melatih dan menguji algoritma yang dipilih untuk digunakan dalam sistem. Hasil pengujian algoritma tersebut dievaluasi untuk mendapatkan model prediksi dengan tingkat akurasi terbaik. Selanjutnya, hasil model prediksi tersebut akan diimplementasikan ke dalam sistem berbasis *website* yang dapat digunakan untuk memprediksi kadar gula darah sesuai masukan yang dibutuhkan sistem. Tahapan dalam analisis dan perancangan model prediksi yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Skema analisis dan perancangan model prediksi gula darah

Data pre-processing atau pra-pemrosesan data adalah tahap pertama yang diperlukan sebelum *machine learning* diterapkan, karena algoritma belajar dari data dan hasil *machine learning* sangat bergantung pada data yang tepat yang diperlukan untuk penyelesaian masalah tertentu yang disebut *feature* [13]. Pada penelitian ini digunakan teknik interpolasi atau *data interpolation*.

Feature extraction atau *feature selection* adalah proses untuk mendapatkan nilai masukan dan keluaran yang

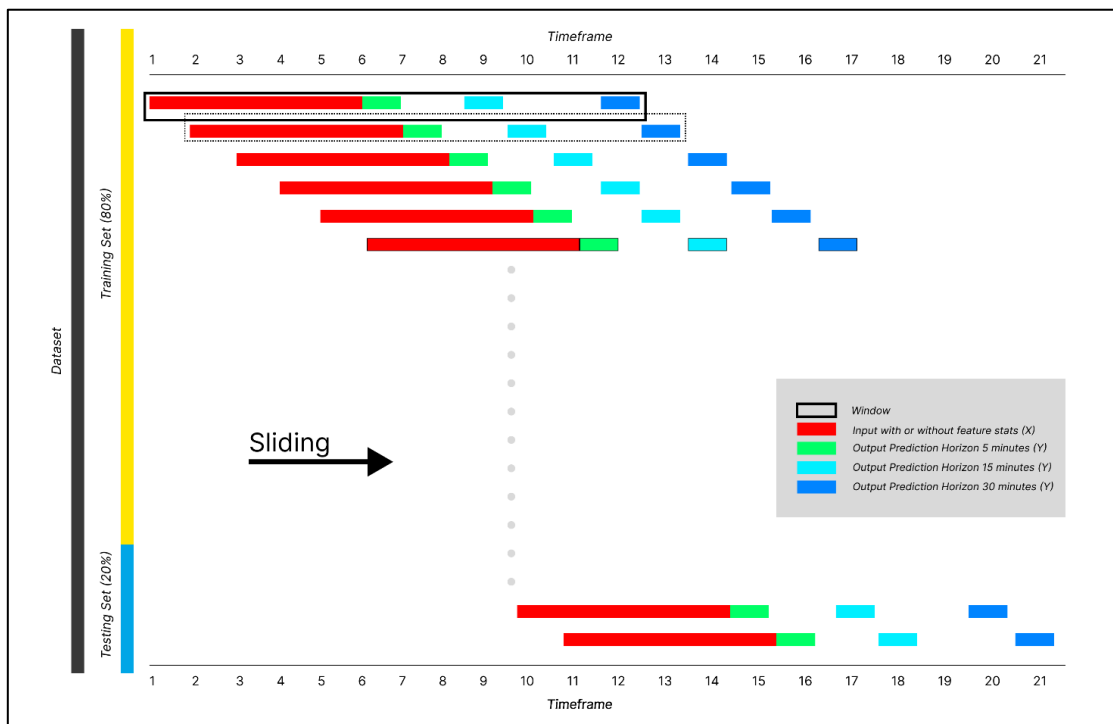
diinginkan untuk digunakan pada algoritma model *machine learning*. Pada tahap ini dilakukan pembagian data menjadi data latih dengan porsi 80% dan data uji dengan porsi 20% dari keseluruhan dataset. Masing-masing dalam data latih dan data uji di dalamnya terdapat data *input* dan *output*. Kemudian dilakukan ekstraksi fitur untuk mendapatkan fitur statistika dari *input* data latih dan data uji. Fitur statistika yang diekstraksi dan digunakan ditunjukkan pada Tabel 1.

Tabel 1. Fitur statistika yang diekstraksi

Nama	Deskripsi
Min	Nilai minimum dari kadar gula darah dalam satu <i>window</i>
Max	Nilai maksimum dari kadar gula darah dalam satu <i>window</i>
Mean	Nilai rata-rata dari kadar gula darah dalam satu <i>window</i>
Std	Nilai standar deviasi dari kadar gula darah dalam satu <i>window</i>
Diff	Nilai perbandingan dari kadar gula darah tertinggi dengan terendah dalam satu <i>window</i>
Median	Nilai tengah dari kadar gula darah dalam satu <i>window</i>
Kurtosis	Nilai keruncingan dari distribusi kadar gula darah dalam satu <i>window</i>
Skew	Nilai asimetris dari distribusi kadar gula darah dalam satu <i>window</i>

Gambar 2 menunjukkan ilustrasi mengubah data *time series* menjadi format *input* (x) dan *output* (y) dengan menggunakan pendekatan teknik *sliding windows*. Pertama, *dataset* dibagi dua menjadi data latih (*training set*) dan data uji (*testing set*) dengan porsi pembagian 80% dan 20%. Kemudian, data latih dan data uji tersebut diubah menjadi format masukan (X) dan keluaran (Y) dengan *Prediction*

Horizon (PH) yang diinginkan, pada penelitian ini digunakan PH 5 menit, 15 menit, dan 30 menit. Selanjutnya, data masukan dilakukan ekstraksi fitur statistik sebagai penambahan parameter masukan yang kemudian akan diuji dengan membandingkan performa model tanpa fitur statistik dengan model yang menggunakan fitur statistik.



Gambar 2. Ilustrasi data *time series* menggunakan pendekatan *sliding windows*

Model selection atau *modeling* adalah proses pemilihan model algoritma *machine learning* yang digunakan dalam melatih dan menguji dataset yang telah diolah sebelumnya. Pada penelitian ini menggunakan metode algoritma *Linear Regression*, *Lasso Regression*, *Ridge Regression*, *Support Vector Regression* (SVR), *eXtreme Gradient Boosting* (XGB), dan *K-Nearest Neighbor* (KNN). Model dari metode algoritma tersebut masing-masing akan dilatih menggunakan data latih yang sudah diproses pada tahap sebelumnya. Kemudian, akan dievaluasi menggunakan data uji pada tahap selanjutnya.

Penggunaan model algoritma pada implementasinya hanya menjalankan fungsi dalam pustaka python dengan memasukkan parameter yang dibutuhkan. Tetapi, dalam fungsi tersebut terdapat persamaan matematika yang dapat dihitung sendiri secara manual. Model algoritma *Linear Regression* dapat dihitung menggunakan persamaan (1-3).

$$Y_i = A + BX \quad (1)$$

$$A = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n(\sum X^2) - (\sum X)^2} \quad (2)$$

$$B = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} \quad (3)$$

Dimana Y_i adalah nilai yang ingin diprediksi (*dependent variable*), X adalah nilai yang digunakan untuk memprediksi (*independent variable*), A adalah nilai *intercept* (nilai rata-rata pada variabel Y_i apabila nilai pada variabel X bernilai 0), dan B adalah nilai slope (nilai seberapa besar kontribusi yang diberikan oleh variabel X terhadap variabel Y).

Penelitian ini menggunakan nilai kadar gula darah pada 30 menit sebelumnya dengan jangka waktu 5 menit, sehingga variabel independen (X) memiliki lebih dari 1 nilai. Kasus tersebut disebut dengan *Multivariate Linear Regression* yang dapat dihitung menggunakan persamaan (4).

$$Y_i = A + B_1X_1 + B_2X_2 + \dots + B_nX_n \quad (4)$$

Dimana B_nX_n adalah nilai koefisien regresi dari independen variabel X , dan n merupakan jumlah banyaknya variabel independen atau titik datanya.

Model evaluation atau evaluasi performa model adalah langkah penting untuk menilai kualitas model prediksi. Beberapa metrik evaluasi yang dapat digunakan adalah *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE), dan *coefficient of determination* (R^2). Semakin kecil nilai MSE, RMSE, MAE, dan MAPE serta semakin besar nilai R -squared, maka semakin baik performa model tersebut. Pada penelitian ini metrik yang digunakan adalah RMSE, MAE, dan R^2 untuk mengevaluasi performa model prediksi.

IV. HASIL DAN PEMBAHASAN

A. Evaluasi Performa Model

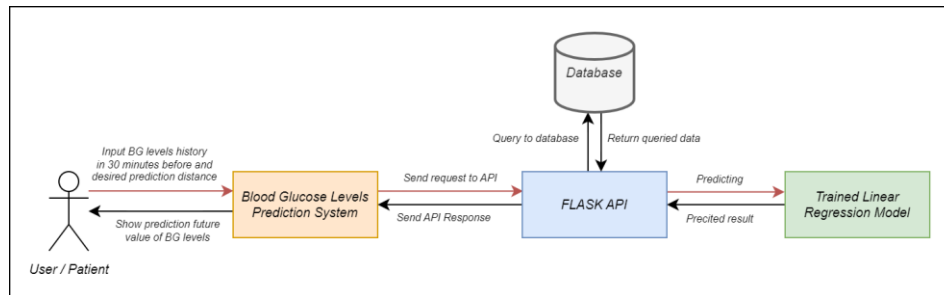
Setelah model prediksi dilatih dengan menggunakan data latih yang telah diolah sebelumnya, selanjutnya adalah melakukan pengujian pada model prediksi tersebut. Pengujian model prediksi ini berfungsi untuk mengevaluasi performa model algoritma *machine learning* yang digunakan untuk mendapatkan model dengan akurasi yang terbaik. Pengujian model prediksi dilakukan dengan melakukan prediksi pada model dengan menggunakan data uji. Data uji tersebut menggunakan 15 data pasien diabetes dari total 30 pasien dalam dataset. Hasil prediksi data uji tersebut dibandingkan dengan data aktual atau sebenarnya menggunakan metrik evaluasi. Metrik evaluasi yang digunakan dalam pengujian model prediksi ini adalah RMSE, MAE, dan R^2 . Performa model prediksi yang diuji dan dievaluasi yaitu model tanpa fitur statistik dengan model yang menggunakan fitur statistik.

Tabel 2 menunjukkan perbandingan metrik evaluasi pada model prediksi tanpa menggunakan fitur statistika, sedangkan Tabel 3 menunjukkan perbandingan metrik evaluasi pada model prediksi dengan menggunakan fitur statistika. Berdasarkan tabel perbandingan metrik evaluasi tersebut, model yang memiliki performa terbaik secara umum adalah regresi linear dengan rata-rata RMSE dan MAE yang paling kecil dan nilai R^2 yang tinggi. Selanjutnya metrik evaluasi performa model dengan fitur statistik dievaluasi untuk melihat apakah dengan menggunakan fitur statistika pada masukan model memiliki pengaruh pada performa model.

B. Sistem Prediksi Kadar Gula Darah Berbasis Website

Secara umum gambaran skema rancangan sistem yang akan dibangun ditunjukkan pada Gambar 3. Pada rancangan sistem tersebut pengguna memasukkan seluruh data yang dibutuhkan dalam sistem berbasis *website*, kemudian masukkan user tersebut akan dikirim sebagai *payload data* ke API yang dibangun menggunakan pustaka python flask, selanjutnya dalam API tersebut data yang diterima akan diproses menggunakan model prediksi yang telah dilatih. Hasil prediksi tersebut dikembalikan dalam bentuk *response data* API ke sistem berbasis *website* dan secara otomatis hasil prediksi tersebut ditampilkan pada halaman *website* sistem.

Halaman *get prediction* merupakan halaman fitur utama dalam sistem ini yang akan menampilkan hasil prediksi kadar gula darah. Pada prosesnya halaman di awal akan tampil masukkan yang harus diisi oleh pengguna, kemudian selanjutnya akan diproses data masukkan tersebut dan akan ditampilkan hasil prediksinya pada halaman selanjutnya. Hasil implementasi rancangan antarmuka halaman *get prediction* dapat dilihat pada Gambar 4. Setelah sistem berhasil memprediksi kadar gula darah selanjutnya, sistem akan menampilkan hasil prediksi berdasarkan waktu yang ditentukan berupa tampilan grafik atau diagram garis yang dapat dilihat pada Gambar 4.



Gambar 3. Ilustrasi skema rancangan sistem yang dibangun

Tabel 2. Perbandingan metrik evaluasi performa model tanpa menggunakan fitur statistik

Model	Metrik	Prediction Horizon (Menit)		
		5	15	30
Linear Regression	RMSE (mg/dL)	5.046±1.617	12.574±3.107	20.312±5.936
	MAE	2.875±1.386	8.212±2.204	14.62±4.469
	R ²	0.962±0.07	0.72±0.622	0.463±0.896
Lasso Regression	RMSE (mg/dL)	5.101±1.272	12.398±3.141	20.021±6.272
	MAE	3.032±1.114	8.242±2.131	14.557±4.722
	R ²	0.954±0.101	0.715±0.658	0.48±0.871
Ridge Regression	RMSE (mg/dL)	5.047±1.617	12.574±3.107	20.312±5.936
	MAE	2.876±1.387	8.212±2.204	14.62±4.469
	R ²	0.962±0.07	0.72±0.622	0.464±0.895
Support Vector Regression (SVR)	RMSE (mg/dL)	6.308±1.856	14.358±2.345	23.053±6.502
	MAE	4.247±1.97	10.513±2.357	17.618±5.319
	R ²	0.935±0.124	0.663±0.716	0.262±1.383
Extreme Gradient Boosting (XGB)	RMSE (mg/dL)	10.335±4.083	16.037±5.438	22.297±8.135
	MAE	6.99±2.689	11.757±4.152	17.093±6.383
	R ²	0.871±0.224	0.746±0.318	0.584±0.337
K-Nearest Neighbor (KNN)	RMSE (mg/dL)	6.473±1.967	15.047±2.425	24.584±7.033
	MAE	4.357±1.873	10.938±2.198	18.761±5.696
	R ²	0.946±0.076	0.666±0.638	0.236±1.271

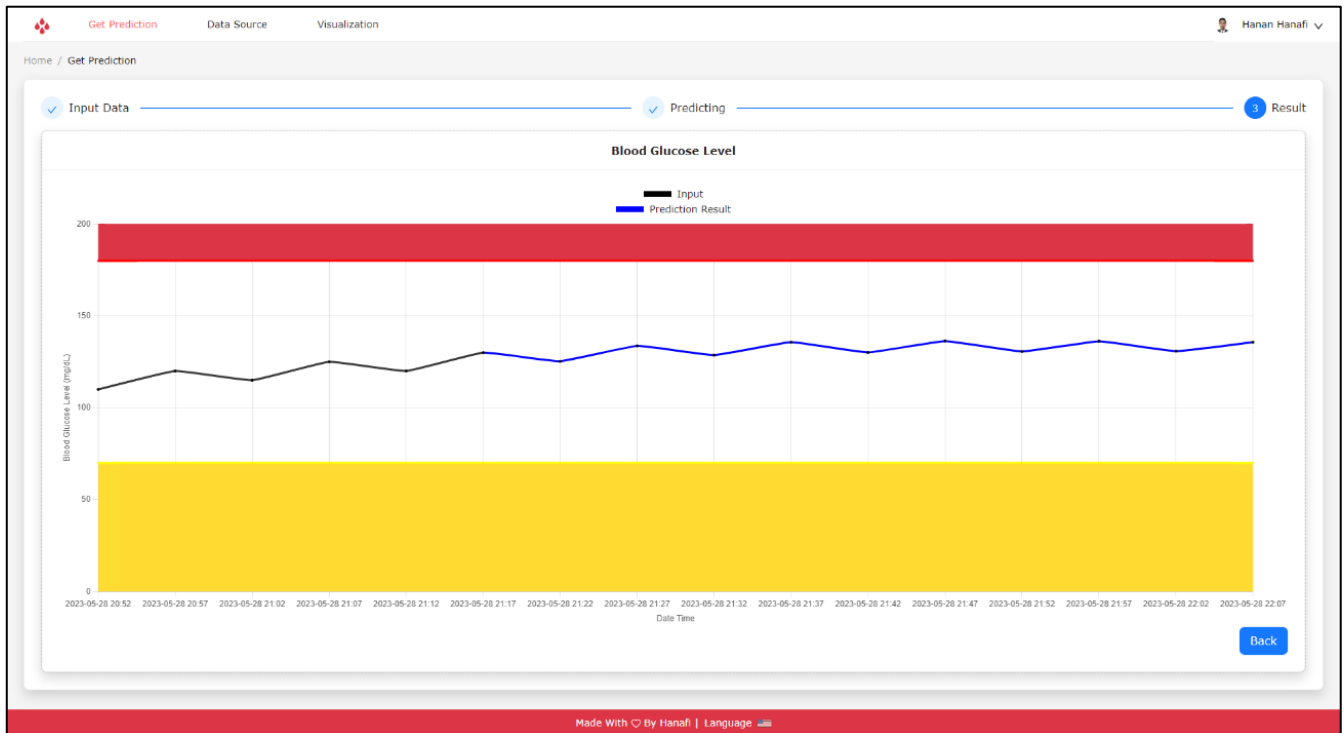
Tabel 3. Perbandingan metrik evaluasi performa model dengan menggunakan fitur statistik

Model	Metrik	Prediction Horizon (Menit)		
		5	15	30
Linear Regression	RMSE (mg/dL)	5.024±1.758	12.488±3.249	20.635±6.162
	MAE	2.891±1.449	8.272±2.368	14.926±4.71
	R ²	0.962±0.069	0.741±0.554	0.39±1.162
Lasso Regression	RMSE (mg/dL)	5.223±1.379	12.476±3.165	20.284±6.505
	MAE	3.089±1.187	8.333±2.195	14.752±4.94
	R ²	0.952±0.104	0.729±0.599	0.441±0.991
Ridge Regression	RMSE (mg/dL)	5.024±1.757	12.483±3.245	20.603±6.15
	MAE	2.89±1.448	8.256±2.346	14.905±4.691
	R ²	0.962±0.069	0.741±0.555	0.402±1.119
Support Vector Regression (SVR)	RMSE (mg/dL)	6.201±1.729	13.976±2.524	22.294±6.377

Model	Metrik	Prediction Horizon (Menit)		
		5	15	30
Extreme Gradient Boosting (XGB)	MAE	4.181±1.824	10.107±2.247	17.078±5.143
	R ²	0.938±0.119	0.687±0.662	0.318±1.307
	RMSE (mg/dL)	12.321±4.012	17.779±5.9	23.794±8.527
	MAE	9.134±3.026	13.531±4.646	18.552±6.617
K-Nearest Neighbor (KNN)	R ²	0.828±0.271	0.707±0.311	0.547±0.307
	RMSE (mg/dL)	6.465±1.528	14.105±2.23	23.519±6.523
	MAE	4.389±1.566	10.176±1.822	18.002±5.197
	R ²	0.943±0.091	0.715±0.554	0.351±0.993

Pada Gambar 4 juga terdapat informasi nilai minimum dan maksimum dari kadar gula darah. Kadar gula darah yang terlalu tinggi atau terlalu rendah dapat menimbulkan masalah kesehatan seperti komplikasi penyakit hingga

risiko kematian. Kadar gula darah yang terlalu tinggi atau *hyperglycemia* adalah kondisi dimana kadar gula darah melebihi 180 mg/dL, sedangkan kadar gula darah yang terlalu rendah atau *hypoglycemia* adalah kondisi dimana kadar gula darah kurang dari 70 mg/dL [14].



Gambar 4. Implementasi halaman *get prediction* bagian hasil prediksi

V. SIMPULAN

Berdasarkan hasil evaluasi performa model prediksi dan pengujian sistem prediksi yang telah dilakukan, didapat kesimpulan bahwa secara keseluruhan metode *Linear Regression* memiliki performa lebih baik daripada *Ridge Regression*, *Lasso Regression*, *Support Vector Regression* (SVR), *eXtreme Gradient Boosting* (XGB), dan *K-Nearest Neighbor* (KNN). Hasil evaluasi juga menunjukkan penambahan fitur statistik dapat meningkatkan performa model regresi linear dengan hasil pada *Prediction Horizon* (PH) 5 menit, 15 menit, dan 30 menit didapat nilai rata-rata

RMSE dari pengujian data 15 pasien adalah 5,024, 12,488, dan 20,635, nilai MAE adalah 2,891, 8,272, dan 14,926 serta nilai R² adalah 0,962, 8,272, dan 14,926. Kemudian, model prediksi regresi linear tersebut berhasil diimplementasi dan divisualisasikan ke dalam sistem berbasis *website*, sehingga sistem dapat digunakan oleh pengguna atau pasien diabetes untuk memprediksi kadar gula darah pada masa depan.

REFERENSI

- [1] Kementerian Kesehatan Republik Indonesia, "Pola Hidup Sehat dan Deteksi Dini Bantu Kontrol Gula Darah Pada Penderita Diabetes," 2022. <https://www.kemkes.go.id/article/view/21111600001/blood-sugar-control-for-dm-patient-through-healthy-lifestyle-and-early-detection.html>
- [2] World Health Organization, *Classification of diabetes mellitus*. Geneva: World Health Organization, 2019. [Online]. Available: <https://apps.who.int/iris/handle/10665/325182>
- [3] World Health Organization, *Global report on diabetes*. Geneva: World Health Organization, 2016. [Online]. Available: <https://apps.who.int/iris/handle/10665/204871>
- [4] World Health Organization, *Improving diabetes outcomes for all, a hundred years on from the discovery of insulin: report of the Global diabetes summit*. Geneva: World Health Organization, 2021. [Online]. Available: <https://apps.who.int/iris/handle/10665/350930>
- [5] Yuris Mulya Saputra, Ganjar Alfian, and Muhammad Qois Huzyan Octava, "Perancangan Federated Learning Berbasis Homomorphic Encryption untuk Perangkat Internet of Things," *J. Internet Softw. Eng.*, vol. 4, no. 1, May 2023, doi: 10.22146/jise.v4i1.6378.
- [6] Rochana Prih Hastuti, Vellya Riona, and Margareta Hardiyanti, "Content Retrieval dengan Fasttext Word Embedding pada Learning Management System Olimpiade," *J. Internet Softw. Eng.*, vol. 4, no. 1, May 2023, doi: 10.22146/jise.v4i1.6766.
- [7] S. M. Pappada *et al.*, "Neural Network-Based Real-Time Prediction of Glucose in Patients with Insulin-Dependent Diabetes," *Diabetes Technol. Ther.*, vol. 13, no. 2, pp. 135–141, Feb. 2011, doi: 10.1089/dia.2010.0104.
- [8] E. I. Georga *et al.*, "Multivariate Prediction of Subcutaneous Glucose Concentration in Type 1 Diabetes Patients Based on Support Vector Regression," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 1, pp. 71–81, 2013, doi: 10.1109/TITB.2012.2219876.
- [9] J. Ben Ali, T. Hamdi, N. Fnaiech, V. Di Costanzo, F. Fnaiech, and J.-M. Ginoux, "Continuous blood glucose level prediction of Type 1 Diabetes based on Artificial Neural Network," *Biocybern. Biomed. Eng.*, vol. 38, no. 4, pp. 828–840, Jan. 2018, doi: 10.1016/j.bbe.2018.06.005.
- [10] T. Hamdi, J. Ben Ali, V. Di Costanzo, F. Fnaiech, E. Moreau, and J.-M. Ginoux, "Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm," *Biocybern. Biomed. Eng.*, vol. 38, no. 2, pp. 362–372, Jan. 2018, doi: 10.1016/j.bbe.2018.02.005.
- [11] J. Martinsson, A. Schliep, B. Eliasson, and O. Mogren, "Blood Glucose Prediction with Variance Estimation Using Recurrent Neural Networks," *J. Healthc. Inform. Res.*, vol. 4, no. 1, pp. 1–18, Mar. 2020, doi: 10.1007/s41666-019-00059-y.
- [12] G. Alfian *et al.*, "Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features," *Biocybern. Biomed. Eng.*, vol. 40, no. 4, pp. 1586–1599, Oct. 2020, doi: 10.1016/j.bbe.2020.10.004.
- [13] A. Holzinger, "Big Data Calls for Machine Learning," in *Encyclopedia of Biomedical Engineering*, R. Narayan, Ed., Oxford: Elsevier, 2019, pp. 258–264. doi: 10.1016/B978-0-12-801238-3.10877-3.
- [14] American Diabetes Association, "Hypoglycemia (Low Blood Glucose)." <https://diabetes.org/healthy-living/medication-treatments/blood-glucose-testing-and-control/hypoglycemia>