

BAB I

PENDAHULUAN

1.1 Latar Belakang

Diabetes merupakan penyakit kronis yang dapat terjadi ketika sel beta pankreas tidak dapat memproduksi insulin dalam jumlah cukup atau ketika tubuh tidak dapat secara efektif menggunakan insulin yang dihasilkannya. Insulin adalah hormon yang berfungsi untuk mengatur kadar gula dalam darah (Istriningsih & Solikhati, 2021). Diabetes menjadi salah satu penyakit kronis serius karena memiliki potensi menimbulkan komplikasi yang mengancam jiwa, melumpuhkan, serta menurunkan angka harapan hidup. International Diabetes Federation (IDF) Diabetes Atlas melaporkan bahwa prevalensi global diabetes pada usia 20–79 tahun diperkirakan mencapai 10,5% (536,6 juta orang) pada 2021 dan diproyeksikan meningkat menjadi 12,2% (783,2 juta orang) pada 2045 (Sun *et al.*, 2022). Penyakit diabetes melitus di Kabupaten Temanggung tahun 2023 menduduki peringkat keempat dari 31 jenis penyakit pada RSUD Kabupaten Temanggung, RS PKU Muhammadiyah Temanggung, RSK Ngesti Waluyo, dan RS Gunung Sawo (Diskominfo Kabupaten Temanggung, 2024).

Diabetes dibagi menjadi dua jenis, yaitu diabetes melitus tipe 1 dan tipe 2 dengan lebih dari 95% penderita diabetes menderita diabetes melitus tipe 2 (WHO, 2024). Diabetes melitus tipe 2 merupakan penyakit kronis yang ditandai dengan meningkatnya kadar gula darah sewaktu (GDS) akibat resistensi insulin. Diabetes melitus tipe 2 dapat memengaruhi tekanan aliran darah karena dengan meningkatnya GDS mengakibatkan gangguan fungsi pembuluh darah (disfungsi

endotel) yang pada akhirnya menyebabkan penyempitan pembuluh darah, kekakuan pembuluh darah, dan peningkatan tekanan darah (Jia & Sowers, 2021).

Tingginya tekanan darah pada pasien diabetes melitus tipe 2 selain disebabkan oleh meningkatnya kadar gula darah juga dapat disebabkan oleh faktor usia. Usia merupakan faktor utama dalam meningkatnya tekanan darah karena adanya perubahan fisiologis pada jantung, pembuluh darah, dan hormon seiring proses penuaan (Tindangen *et al.*, 2020). Tingginya tekanan darah dapat mengakibatkan hipertensi yang dapat diukur dengan memeriksa tekanan darah sistolik atau tekanan darah diastolik pada tubuh. Penelitian ini menggunakan tekanan darah sistolik dalam analisisnya. Hipertensi merupakan penyakit kronis yang menjadi penyebab kematian nomor satu di dunia (Kemenkes, 2024).

Seberapa besar pengaruh kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pada pasien diabetes melitus tipe 2 dapat dianalisis dengan metode regresi. Analisis regresi merupakan metode yang bertujuan untuk mengetahui pengaruh suatu variabel prediktor (X) terhadap variabel respon (Y). Analisis regresi dapat dikelompokkan menjadi tiga, yaitu regresi parametrik, regresi nonparametrik, dan regresi semiparametrik (Susanti *et al.*, 2019).

Regresi parametrik merupakan regresi dengan pola data sudah diketahui, misalnya berpola linier, eksponensial, ataupun yang lainnya. Regresi parametrik memiliki beberapa asumsi dalam penggunaannya, di antaranya yaitu *error* pada model harus independen dan berdistribusi normal identik. Asumsi-asumsi yang tidak terpenuhi dapat mengakibatkan model tidak akurat untuk digunakan (Suparti *et al.*, 2018). Regresi nonparametrik dapat dijadikan metode alternatif untuk analisis regresi dengan pola data yang belum diketahui (Rasyid *et al.*, 2025). Bentuk fungsi

yang tidak diketahui dapat diestimasi dengan teknik pemulusan atau penghalusan (*smoothing*). Kelebihan dari regresi nonparametrik yaitu tidak memiliki asumsi yang ketat seperti regresi parametrik (Suparti *et al.*, 2018). Regresi semiparametrik merupakan regresi yang digunakan apabila pola data pada variabel tertentu diketahui bentuknya, sedangkan variabel yang lain tidak diketahui bentuk pola datanya (Hesikumalasari *et al.*, 2022).

Teknik pemulusan atau penghalusan pada regresi nonparametrik terdiri dari kernel, polinomial lokal, spline, dan deret fourier (Chamidah & Lestari, 2022). Regresi kernel adalah metode yang bertujuan untuk menggambarkan hubungan antara variabel prediktor dan variabel respon dengan memanfaatkan bobot fungsi densitas kernel. Kelebihan dari regresi kernel yaitu mampu menyesuaikan pola data dengan bentuk acak dan tingkat konvergensinya relatif cepat (Hardle, 1994). Jenis regresi kernel dibagi menjadi dua berdasarkan banyaknya variabel prediktor, yaitu regresi kernel univariabel dan multivariabel. Regresi kernel univariabel terdiri dari satu variabel prediktor dan satu variabel respon sedangkan regresi kernel multivariabel memiliki dua atau lebih variabel prediktor dan satu variabel respon.

Regresi kernel menggunakan bobot fungsi densitas kernel untuk memperoleh nilai estimasi pada variabel respon. Nilai estimasi dan bobot didapatkan dengan menggunakan estimator fungsi regresi kernel. Estimator tersebut dibagi menjadi tiga, yaitu estimator *Nadaraya-Watson*, *Priestley-Chao*, dan *Gasser-Muller* (Herawati *et al.*, 2022). Estimator yang paling sering digunakan adalah estimator *Nadaraya-Watson*. Estimator tersebut merupakan estimator yang tidak mengasumsikan bentuk fungsi tertentu sehingga dapat fleksibel menyesuaikan diri pada data yang dianalisis. Keunggulan lainnya yaitu dapat memberikan estimasi

yang akurat pada pola data acak serta memiliki konsep yang sederhana dan mudah dipahami (Wulandari *et al.*, 2025).

Estimator densitas kernel merupakan metode untuk mendapatkan estimasi yang akurat pada regresi kernel. Estimator tersebut dihitung menggunakan fungsi kernel dan *bandwidth*. Jenis fungsi kernel yaitu fungsi *uniform*, *triangle*, Epanechnikov, dan Gaussian (Hardle, 1991). Fungsi kernel Gaussian sering digunakan karena perhitungannya yang sederhana dan memiliki kelebihan yaitu terdefinisi bobotnya pada semua bilangan riil (Rahmah *et al.*, 2025). Nilai *bandwidth* sangat berpengaruh terhadap nilai prediksi karena nilai *bandwidth* terlalu kecil dapat mengakibatkan fungsi pendugaan *overfitting* sedangkan *bandwidth* terlalu besar dapat mengakibatkan fungsi pendugaan *underfitting* (Hardle, 1994). *Overfitting* merupakan kondisi ketika model terlalu mengikuti pola data *training*, sehingga memiliki akurasi tinggi pada data *training*, tetapi akurasinya rendah pada data *testing*. *Underfitting* merupakan kondisi ketika model terlalu sederhana dalam menangkap pola data *training*, sehingga memiliki akurasi rendah pada data *training*, sedangkan akurasi pada data *testing* bisa tinggi ataupun rendah tergantung pola data. Oleh karena itu, penentuan *bandwidth* optimal sangat dibutuhkan untuk mendapatkan model terbaik dan estimasi yang lebih akurat.

Pemilihan *bandwidth* optimal dapat menggunakan metode optimasi *Mean Square Error* (MSE), *Cross-Validation* (CV), *Generalized Cross-Validation* (GCV), dan *Unbiased Risk* (UBR) (Prahutama *et al.*, 2018; Rasyid *et al.*, 2025). Optimasi MSE kurang tepat digunakan untuk pemilihan *bandwidth* karena semakin kecil nilai *bandwidth* maka nilai MSE juga akan semakin kecil (Suparti *et al.*, 2018). Optimasi CV memiliki kekurangan yaitu nilai CV dapat tak terdefinisi karena

terdapat data yang pembobotannya nol. Oleh karena itu, metode GCV dan UBR dipilih untuk mengatasi hal tersebut sehingga mendapatkan estimasi yang lebih akurat.

Penelitian yang dilakukan oleh Muhajiriansyah dan Binuko (2023) menganalisis hubungan kadar HbA1c dan kadar gula darah sewaktu terhadap peningkatan tekanan darah pasien diabetes melitus tipe 2 di RS Darmayu Ponorogo, Jawa Timur dengan uji *Chi-Square* mendapatkan hasil bahwa terdapat hubungan antara kadar HbA1c dan kadar gula darah sewaktu terhadap tekanan darah pasien diabetes melitus tipe 2. Penelitian lain dilakukan oleh Naseri *et al.* (2022) mengenai hubungan usia dengan tekanan darah sistolik dan diastolik pada pasien diabetes melitus tipe 2 di Noble OPD Clinic Afghanistan dengan uji korelasi Pearson 2-tailed memberikan hasil bahwa usia dengan tekanan darah sistolik memiliki hubungan positif signifikan sedangkan usia dengan tekanan darah diastolik tidak memiliki hubungan signifikan.

Rasyid *et al.* (2025) membandingkan optimasi CV dan GCV menggunakan regresi kernel multivariabel fungsi *triangle* dalam memodelkan Indeks Pembangunan Manusia berdasarkan kualitas pendidikan di Indonesia 2023 mendapatkan hasil bahwa model terbaik diperoleh dengan optimasi GCV karena nilai R^2 lebih tinggi dan nilai MAPE lebih rendah daripada model dengan optimasi CV. Penelitian lain yang dilakukan oleh Azhar *et al.* (2024) mengenai perbandingan antara metode optimasi UBR dan GCV dalam pemilihan titik knot pada pemodelan faktor-faktor yang memengaruhi Tingkat Pengangguran Terbuka (TPT) di Provinsi Jawa Tengah menggunakan regresi spline truncated multivariabel mendapatkan

hasil bahwa optimasi GCV lebih baik dari optimasi UBR karena nilai R^2 lebih tinggi dan nilai MSE lebih rendah daripada model dengan optimasi UBR.

Berdasarkan uraian yang telah dipaparkan, penelitian ini menganalisis pengaruh kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pada pasien diabetes melitus tipe 2 di RS PKU Muhammadiyah Kabupaten Temanggung 2024. Metode yang digunakan adalah regresi nonparametrik dengan pendekatan kernel menggunakan fungsi kernel Gaussian serta metode optimasi GCV dan UBR. Metode optimasi tersebut berbeda dengan metode optimasi pada penelitian sebelumnya yang dilakukan oleh Rasyid *et al.* (2025) yang membandingkan antara CV dan GCV karena optimasi CV jika diterapkan pada regresi kernel dapat menghasilkan banyak nilai CV yang tak terdefinisi disebabkan oleh data observasi dengan pembobotan nol. Oleh karena itu, penelitian ini membandingkan optimasi GCV dan UBR untuk mengetahui apakah penelitian Azhar *et al.* (2024) yang memberikan hasil GCV lebih unggul daripada UBR pada regresi spline juga berlaku pada regresi kernel.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, dapat dirumuskan permasalahan sebagai berikut:

1. Bagaimana penerapan model regresi kernel multivariabel pada data kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pasien yang mengidap diabetes melitus tipe 2 di RS PKU Muhammadiyah Temanggung tahun 2024 dengan optimasi GCV dan UBR?
2. Apa model terbaik yang dihasilkan dari perbandingan uji kebaikan model regresi kernel multivariabel dengan optimasi GCV dan UBR pada data *training*

kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pasien yang mengidap diabetes melitus tipe 2 di RS PKU Muhammadiyah Temanggung tahun 2024 yang diukur dengan nilai R^2 ?

3. Bagaimana kinerja model terbaik regresi kernel multivariabel pada data *testing* kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pasien yang mengidap diabetes melitus tipe 2 di RS PKU Muhammadiyah Temanggung tahun 2024 yang diukur dengan nilai SMAPE?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini bertujuan agar analisisnya lebih fokus dan tidak meluas dari pembahasan yang dimaksud. Oleh karena itu, penulis membatasi penelitian ini dalam ruang lingkup sebagai berikut:

1. Data dalam penelitian ini yaitu data kadar gula darah sewaktu (GDS), usia, dan tekanan darah sistolik pada pasien diabetes melitus tipe 2 di RS PKU Muhammadiyah Temanggung tahun 2024.
2. Data yang digunakan dibagi menjadi data *training* dan *testing* dengan perbandingan 85:15 yang dibagi secara acak untuk pembentukan modelnya.
3. Analisis yang digunakan adalah regresi nonparametrik kernel multivariabel dengan estimator kernel *Nadaraya-Watson*.
4. Fungsi pembobot menggunakan fungsi kernel Gaussian dengan rentang *bandwidth* 1 yang diujikan adalah 1 sampai 5 dengan kenaikan *bandwidth* adalah 0,1 sedangkan rentang *bandwidth* 2 yang diujikan adalah 0,1 sampai 3 dengan kenaikan *bandwidth* adalah 0,1.
5. Pemilihan *bandwidth* optimal menggunakan metode optimasi *Generalized Cross-Validation* (GCV) dan *Unbiased Risk* (UBR).

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan sebelumnya, tujuan dari penelitian ini sebagai berikut:

1. Memperoleh model regresi kernel multivariabel pada data kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pasien yang mengidap diabetes melitus tipe 2 di RS PKU Muhammadiyah Temanggung tahun 2024 dengan optimasi GCV dan UBR.
2. Memperoleh model terbaik dari perbandingan uji kebaikan model regresi kernel multivariabel dengan optimasi GCV dan UBR pada data *training* kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pasien yang mengidap diabetes melitus tipe 2 di RS PKU Muhammadiyah Temanggung tahun 2024 yang diukur dengan nilai R^2 .
3. Memperoleh kinerja model terbaik regresi kernel multivariabel dengan pada data *testing* kadar gula darah sewaktu dan usia terhadap tekanan darah sistolik pasien yang mengidap diabetes melitus tipe 2 di RS PKU Muhammadiyah Temanggung tahun 2024 yang diukur dengan nilai SMAPE.