

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi adalah analisis statistika yang digunakan untuk memprediksi nilai variabel respon berdasarkan pada nilai variabel prediktor (Kusnandar, Debatara, Mara, dan Satyahadewi, 2019). Pada umumnya terdapat tiga pendekatan untuk mengestimasi kurva dalam analisis regresi yaitu regresi parametrik, nonparametrik, dan semiparametrik (Utami, 2013). Regresi parametrik diasumsikan bahwa pola pada variabel data sudah diketahui. Regresi nonparametrik diasumsikan bahwa pola pada variabel data tidak diketahui dan kurva regresinya bersifat mulus (*smooth*). Regresi semiparametrik merupakan gabungan dari regresi parametrik dan regresi nonparametrik (Nurhuda, Wasono, dan Nohe, 2022).

Regresi nonparametrik memiliki fleksibilitas yang tinggi karena data diharapkan mampu mencari sendiri bentuk dari kurva regresinya (Nurhuda dkk, 2022). Pada regresi nonparametrik terdapat beberapa estimasi kurva regresi diantaranya estimator kernel, deret ortogonal, deret fourier, *spline*, wavelets (Budiantara, 2005). Salah satu pendekatan regresi nonparametrik yang paling sering digunakan adalah estimator *spline* (Nurhuda dkk, 2022).

Spline adalah salah satu bentuk polinomial yang bersifat tersegmen (Pangestikasari, Rahmawati, dan Ispriyanti, 2018). Kelebihan dari *spline* adalah mampu mengatasi pola data yang naik atau turun secara tajam dengan bantuan titik knot, serta kurva yang dihasilkan relatif lebih mulus (Fathurahman, 2011).

Dalam analisis regresi apabila melibatkan dua variabel respon dan diantara kedua variabel respon tersebut terdapat hubungan, baik secara logika maupun matematis maka disebut regresi birespon (Fernandes, Budiantara, Otok, dan Suhartono, 2014). Apabila regresi birespon memiliki bentuk kurva regresi yang tidak diketahui maka pendekatan yang digunakan adalah regresi nonparametrik birespon *spline*. Pada penelitian ini regresi nonparametrik birespon *spline* diterapkan pada kasus kemiskinan. Kemiskinan adalah suatu keadaan

ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan. Penduduk dikategorikan sebagai penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan di bawah garis kemiskinan. Kemiskinan dapat diukur dengan dua indikator yaitu persentase penduduk miskin dan indeks kedalaman kemiskinan (BPS, 2014). Persentase penduduk miskin merupakan persentase penduduk yang berada di bawah garis kemiskinan, sedangkan indeks kedalaman kemiskinan merupakan ukuran rata-rata kesenjangan pengeluaran dari masing-masing penduduk miskin terhadap garis kemiskinan. Persentase penduduk miskin dapat digunakan untuk mengetahui berapa besar proporsi masyarakat yang masih kesulitan untuk memenuhi hidupnya (Subarna, 2012), sedangkan indeks kedalaman kemiskinan dapat menunjukkan berapa besar rata-rata pengeluaran penduduk miskin di suatu daerah.

Indonesia merupakan salah satu negara berkembang yang masih mengalami masalah kemiskinan. Pada tahun 2020 jumlah penduduk miskin di Indonesia sebanyak 27.550.000 jiwa (BPS, 2021). Salah satu provinsi di Indonesia yang tidak terlepas dari masalah kemiskinan adalah Jawa Timur. Jawa Timur merupakan provinsi dengan jumlah penduduk terbesar kedua di Indonesia yang mencapai 40.665.696 jiwa. Jumlah penduduk miskin di Jawa Timur pada tahun 2020 sebanyak 4.419.100 jiwa, persentase penduduk miskin sebesar 11,09, dan indeks kedalaman kemiskinan sebesar 1,82 (BPS, 2021). Adanya masalah kemiskinan ini apabila tidak cepat diatasi, maka akan berdampak di wilayah Jawa Timur. Masalah kemiskinan dapat dipengaruhi oleh beberapa faktor yaitu tingkat pengangguran terbuka, tingkat partisipasi angkatan kerja, rata-rata lama sekolah, dan laju pertumbuhan penduduk (Tediwibawa, Yuniarti, dan Hayati, 2019).

Pada penelitian ini pendekatan regresi nonparametrik birespon *spline* digunakan untuk memodelkan persentase penduduk miskin dan indeks kedalaman kemiskinan di Provinsi Jawa Timur pada tahun 2020. Faktor-faktor yang dapat mempengaruhi persentase penduduk miskin dan indeks kedalaman kemiskinan di Jawa Timur yaitu tingkat pengangguran terbuka dan tingkat partisipasi angkatan kerja.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka diperoleh rumusan masalah yaitu bagaimana model regresi nonparametrik birespon *spline* terbaik pada persentase penduduk miskin dan indeks kedalaman kemiskinan.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan dari penelitian ini adalah untuk menentukan model regresi nonparametrik birespon *spline* terbaik pada persentase penduduk miskin dan indeks kedalaman kemiskinan.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah data persentase penduduk miskin, indeks kedalaman kemiskinan, tingkat pengangguran terbuka, dan tingkat partisipasi angkatan kerja di 38 kabupaten/kota Provinsi Jawa Timur pada tahun 2020.
2. Pemilihan titik knot optimal menggunakan nilai *Generalized Cross Validation* (GCV) minimum.
3. Titik knot yang digunakan adalah satu, dua, dan tiga titik knot.
4. Analisis menggunakan pendekatan regresi nonparametrik birespon *spline* dengan pembobot W sebagai invers matriks variansi-kovariansi respon.

1.5 Tinjauan Pustaka

Padatuan, Sifriyani, dan Prangga (2021) melakukan penelitian dengan metode regresi nonparametrik birespon *spline* dalam menentukan model terbaik pada angka harapan hidup dan angka kematian bayi di 41 Kabupaten/kota Pulau Kalimantan. Hasil dari penelitian menunjukkan bahwa model terbaik yaitu model regresi nonparametrik *spline* linier birespon dengan nilai R^2 sebesar 80,51% dan model *spline* tiga titik knot dengan nilai GCV minimum sebesar 7,1454.

Dewanti, Suparti, dan Prahutama (2020) melakukan pemodelan pada Indeks Harga Saham Gabungan (IHSG) dan Jakarta *Islamic Index* (JII) menggunakan regresi nonparametrik birespon *spline truncated* berbasis GUI R. Data pada

penelitiannya dibagi menjadi dua yaitu data *in sample* digunakan untuk menentukan model terbaik dari regresi nonparametrik birespon *spline truncated*, sedangkan data *out sample* digunakan untuk mengevaluasi kinerja dari model terbaik yang diperoleh. Hasil analisis diperoleh bahwa model terbaik pada saat Indeks Harga Saham Gabungan (IHSG) berada pada orde dua dan Jakarta *Islamic Index* (JII) berada pada orde tiga dengan dua titik knot dengan nilai GCV sebesar 10524,50. Nilai R^2 sebesar 85,54437% dan MAPE data *out sample* sebesar 2,65595% artinya memiliki kemampuan yang sangat baik dalam peramalan.

Rahmawati, Suparti, dan Prahutama (2019) menentukan model terbaik pada inflasi dan Indeks Harga Saham Gabungan (IHSG) menggunakan metode regresi nonparametrik birespon *spline truncated* dengan pembobot matriks variansi-kovariansi *error* respon. Pada penelitiannya terdapat pembagian data yaitu data *in sample* digunakan untuk menentukan model terbaik dari regresi nonparametrik birespon *spline truncated*, sedangkan data *out sample* digunakan untuk mengevaluasi kinerja dari model terbaik yang diperoleh. Hasil dari analisis didapat model terbaik pada saat inflasi dan Indeks Harga Saham Gabungan (IHSG) berada pada orde dua dengan empat titik knot dan nilai MSE sebesar 4781,697. Nilai R^2 sebesar 96,2679% dan MAPE data *out sample* sebesar 43,446% menunjukkan bahwa kemampuan yang cukup baik dalam peramalan.

Penelitian dengan menggunakan metode regresi nonparametrik birespon *spline truncated* juga dilakukan oleh Pangestikasari, Rahmawati, dan Ispriyanti (2018) untuk memodelkan pengeluaran per kapita dan persentase penduduk miskin di 35 kabupaten/kota Provinsi Jawa Tengah. Berdasarkan hasil analisis yang telah dilakukan, bahwa model terbaik didapat pada saat pengeluaran per kapita dan persentase penduduk miskin berorde dua dengan lima titik knot dan nilai MSE sebesar 662634,2. Nilai R^2 yang dihasilkan sebesar 43,21% dan nilai MAPE sebesar 14,25% berarti memiliki kemampuan yang baik dalam peramalan.

1.6 Metodologi Penelitian

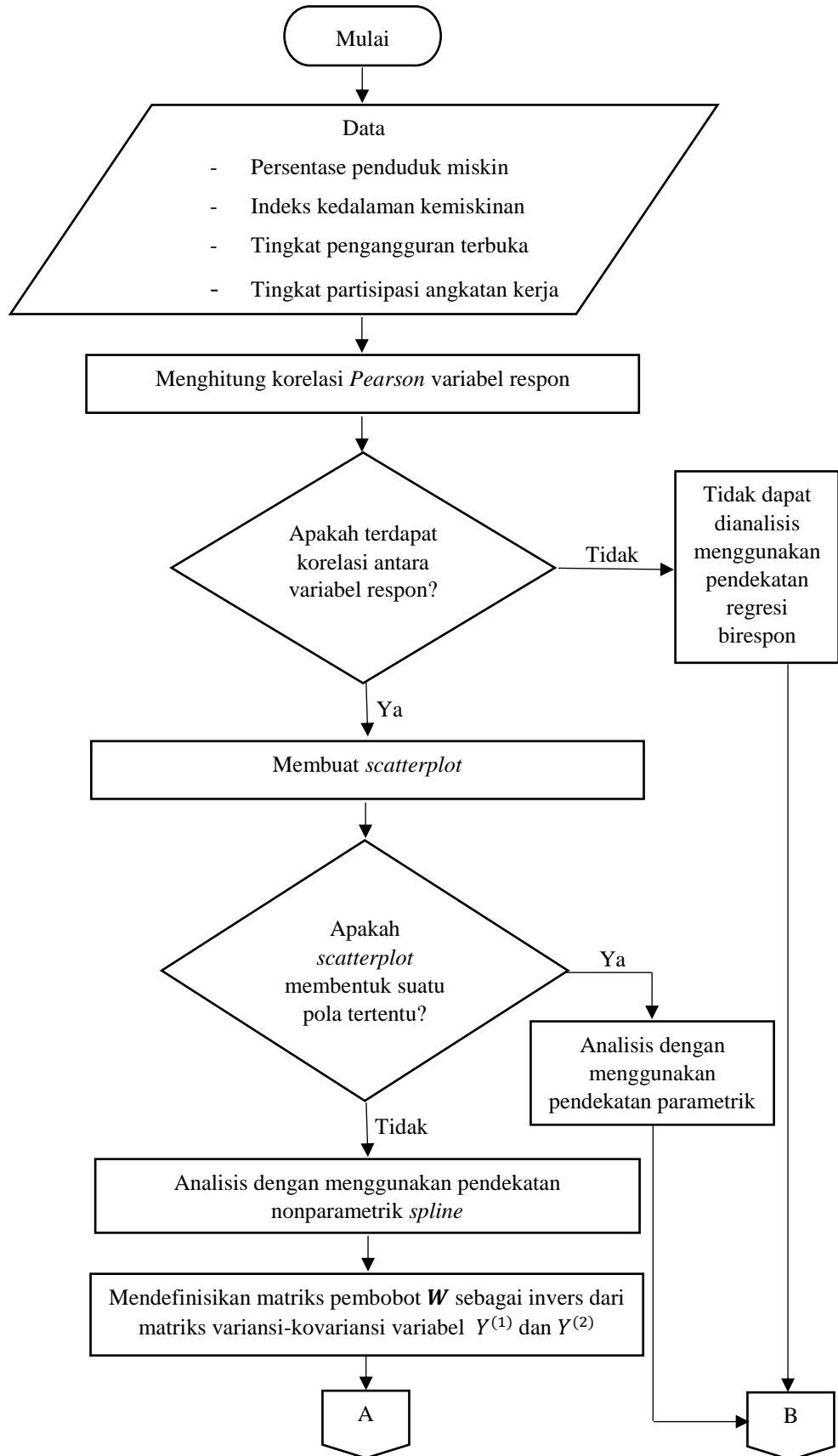
Penelitian ini menggunakan data kabupaten/kota di Provinsi Jawa Timur pada tahun 2020 yang diperoleh dari *website* Badan Pusat Statistik (BPS) Provinsi Jawa

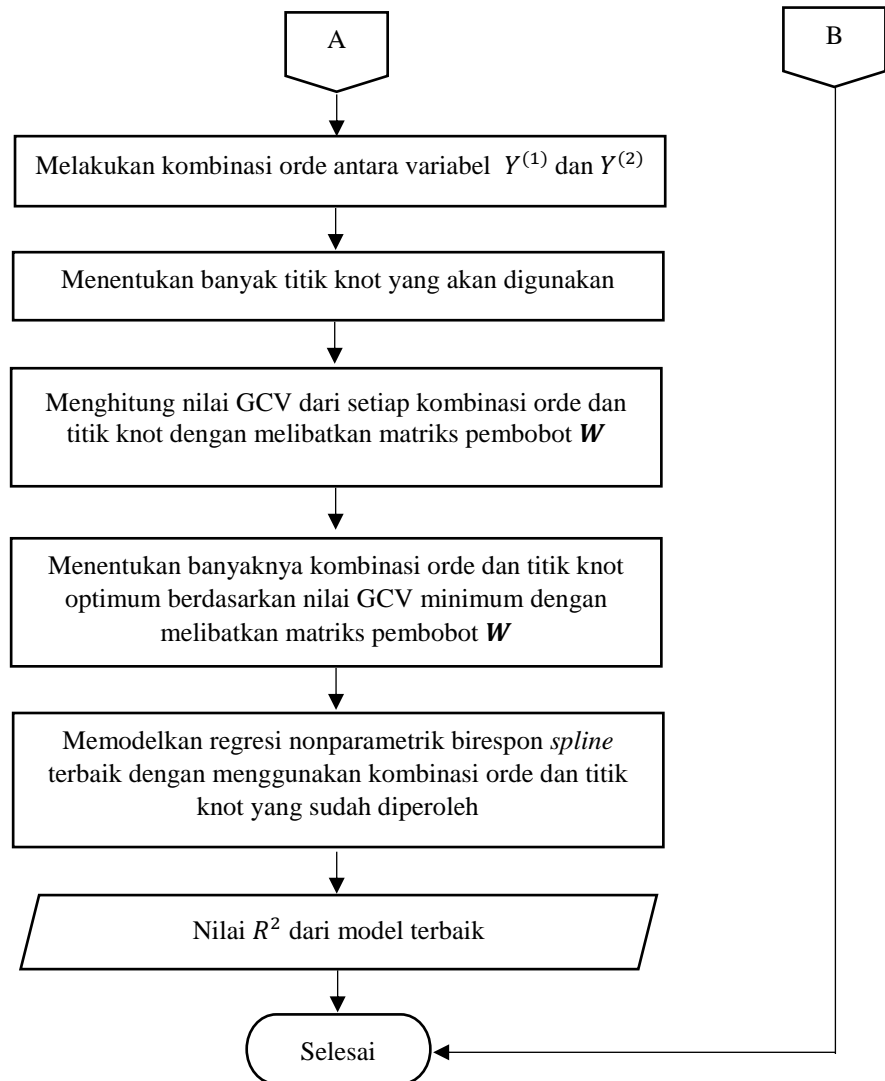
Timur dengan 38 observasi. Variabel yang digunakan dalam penelitian ini adalah persentase penduduk miskin ($Y^{(1)}$), indeks kedalaman kemiskinan ($Y^{(2)}$), tingkat pengangguran terbuka (X_1), dan tingkat partisipasi angkatan kerja (X_2). Analisis yang digunakan pada penelitian ini adalah pendekatan regresi nonparametrik birespon *spline* dengan pembobot invers matriks variansi-kovariansi respon. Penelitian ini dianalisis menggunakan *software* R.4.1.0.

Tahapan pertama yang dilakukan dalam menganalisis data adalah melakukan uji korelasi *Pearson* antara variabel $Y^{(1)}$ dan variabel $Y^{(2)}$ untuk melihat berapa besar korelasi antara kedua variabel respon tersebut. Dari nilai korelasi *Pearson* yang diperoleh dapat digunakan uji t untuk melihat apakah terdapat korelasi antara variabel $Y^{(1)}$ dan variabel $Y^{(2)}$. Apabila terdapat korelasi antara variabel respon maka analisis dapat menggunakan pendekatan regresi birespon dan apabila tidak terdapat korelasi antara variabel respon maka analisis tidak dapat menggunakan pendekatan regresi birespon.

Tahap kedua mengidentifikasi data dengan membuat *scatterplot* untuk masing – masing variabel respon terhadap variabel prediktor. Dari bentuk *scatterplot* dapat terlihat apakah terdapat pola data tertentu atau tidak untuk variabel respon terhadap variabel prediktor. Tahap ketiga mendefinisikan matriks pembobot W sebagai invers dari matriks variansi-kovariansi variabel $Y^{(1)}$ dan variabel $Y^{(2)}$. Tahap kelima dilakukan kombinasi orde antara variabel $Y^{(1)}$ dan variabel $Y^{(2)}$ dengan menentukan banyak titik knot yang akan digunakan. Tahapan keenam menghitung nilai *Generalized Cross Validation* (GCV) dari setiap kombinasi orde dan titik knot yang melibatkan matriks pembobot W .

Tahapan selanjutnya menentukan banyaknya kombinasi orde dan titik knot optimum berdasarkan nilai GCV minimum yang melibatkan matriks pembobot W . Kemudian diperoleh model terbaik pada regresi nonparametrik birespon *spline*. Berdasarkan model terbaik diperoleh nilai R^2 untuk melihat ukuran kebaikan dari model yang sudah didapat. Langkah-langkah dalam pemodelan persentase penduduk miskin ($Y^{(1)}$) dan indeks kedalaman kemiskinan ($Y^{(2)}$) untuk lebih jelasnya disajikan pada Gambar 1.1.





Gambar 1.1 Flowchart Pemodelan Regresi Nonparametrik Birespon *Spline* pada Persentase Penduduk Miskin dan Indeks Kedalaman Kemiskinan