

Estimator Nadaraya-Watson dengan Fungsi Kernel Normal dan Fungsi Kernel Kuadratik

Saskia Amalia Putri, Ayudita Rahmi Aristya, Nur Azizah Janad, Yudy Novindri Tadale,
Lilies Handayani,
Prodi Statistika, Universitas Tadulako, Kota Palu, Sulawesi Tengah

*e-mail:

saskia.amalia.pitri0108@gmail.com, lilies.stath@gmail.com

Abstrak

Indeks pembangunan manusia (IPM) atau Human Development Index (HDI) adalah pengukuran perbandingan dari harapan hidup, melek huruf, pendidikan dan standar hidup untuk semua negara seluruh dunia. Indeks pembangunan manusia digunakan untuk mengklasifikasikan apakah sebuah negara adalah negara maju, negara berkembang atau negara terkebelakang dan juga untuk mengukur pengaruh dari kebijaksanaan ekonomi terhadap kualitas hidup. Tujuan penelitian ini yaitu untuk mengetahui bagaimana hubungan persentase penduduk miskin dan tingkat pengangguran terhadap indeks pembangunan manusia di Sulawesi Tengah. Dalam penelitian ini digunakan metode regresi kernel Nadaraya-Watson dengan perbandingan regresi antara fungsi kernel normal dan fungsi kernel kuadratik. Berdasarkan hasil penelitian Untuk model terbaik pada X_1 (persentase penduduk miskin) dengan nilai MSE terkecil yaitu pada metode CV.LS dengan nilai bandwidth sebesar 1.369349, dan untuk model terbaik pada X_2 (tingkat pengangguran terbuka) dengan nilai MSE terkecil yaitu pada metode CV.AIC dengan nilai bandwidth sebesar 1.331878.

Kata Kunci : *Estimator Kernel, Indeks Pembangunan Manusia, Persentase Penduduk miskin, Tingkat Pengangguran Terbuka*

1. Pendahuluan

Statistik dipakai untuk menyatakan kumpulan fakta, umumnya berbentuk angka yang disusun dalam tabel atau diagram yang melukiskan atau menggambarkan suatu persoalan. Sedangkan statistika adalah ilmu yang terdiri dari teori dan metode yang merupakan cabang dari matematika terapan dan membicarakan tentang bagaimana mengumpulkan data, bagaimana meringkas data, mengolah dan menyajikan data, bagaimana menarik kesimpulan dari analisis, bagaimana menentukan keputusan dalam batas-batas resiko tertentu berdasarkan strategi yang ada. Statistika sebagai pengetahuan berhubungan dengan cara-cara pengumpulan fakta, pengolahan serta pembuatan keputusan yang cukup beralasan berdasarkan fakta dan Analisa yang dilakukan (Sudjana,1996).

Statistika menurut tingkat atau tahapan kegiatannya dibagi menjadi dua kelompok, yaitu statistika deskriptif dan statistika inferensia (Khasanah, 2016). Statistika deskriptif adalah bagian dari statistika yang mempelajari cara pengumpulan data dan penyajian data sehingga mudah dipahami. Statistika deskriptif mencakup distribusi frekuensi beserta bagian- bagiannya seperti grafik distribusi (histogram, poligon frekuensi, dan ogif), ukuran nilai pusat (rata-rata, median, modus, kuartil dan sebagainya), ukuran dispersi (jangkauan, simpangan rata-rata, variasi, simpangan baku, dan sebagainya), kemencengan dan keruncingan kurva (Hasan, 2001). Sementara itu, statistika inferensia adalah statistika yang menyediakan aturan atau cara yang dapat

digunakan untuk menarik kesimpulan, membuat ramalan, dan penaksiran. Dengan kata lain, statistika inferensia adalah metode yang berhubungan dengan analisis data pada sampel dan hasilnya dipakai untuk generalisasi pada populasi yang bertujuan untuk melakukan estimasi, menguji hipotesis, dan mengambil keputusan (Nisfiannoor, 2009).

Pada statistika inferensia, dapat dibagi ke dalam dua golongan, yaitu statistika parametrik dan statistika nonparametrik. Statistika parametrik adalah suatu penggunaan teknik yang didasarkan pada asumsi bahwa data yang diambil mempunyai distribusi normal dan jenis data yang digunakan adalah data berjenis interval atau rasio. Sedangkan statistika nonparametrik adalah suatu penggunaan teknik yang tidak mengharuskan data yang diambil mempunyai distribusi normal atau jenis data yang digunakan adalah data berjenis nominal atau ordinal.

Pada statistika inferensia baik itu statistika parametrik maupun statistika nonparametrik dapat dilakukan suatu analisis data. Biasanya analisis yang dilakukan untuk mengetahui hubungan dua variabel atau lebih adalah analisis regresi. Analisis regresi sendiri secara umum merupakan alat analisis statistik yang memanfaatkan hubungan antara dua variabel atau lebih dengan tujuan untuk membuat perkiraan atau prediksi yang dapat dipercaya untuk nilai suatu variabel (disebut variabel terikat atau dependen), jika nilai variabel lain yang berhubungan dengannya diketahui disebut sebagai variabel bebas atau independen (Quadratullah, 2013). Untuk menyelesaikan suatu analisis regresi, dapat dilakukan dengan dua model pendekatan yaitu pendekatan model regresi parametrik dan pendekatan model regresi nonparametrik. Pendekatan model regresi parametrik sering dilakukan, dikarenakan kemudahan dalam estimasinya akan tetapi, pendekatan parametrik sangat terikat dengan asumsi-asumsi antara lain *error* berdistribusi normal, *homokedastisitas*, non-autokorelasi serta antara variabel dependen dan variabel independen tidak terjadi multikolinearitas. Sehingga jika suatu permasalahan dimodelkan menggunakan pendekatan parametrik, akan tetapi model yang didapatkan asumsi-asumsinya tidak terpenuhi, tentunya model tersebut akan menjadi bias untuk digunakan.

Pada model regresi parametrik mempunyai asumsi yang sangat kaku seperti yang sudah dijelaskan di atas, dan bentuk kurva regresi telah diketahui, misalnya linear, kuadratik, kubik, eksponensial dan lain-lain. Sedangkan dalam regresi nonparametrik bentuk kurva regresi tidak diketahui, data diharapkan mencari sendiri bentuk estimasinya sehingga memiliki fleksibilitas yang tinggi (Sukarsa & Srinadi, 2012). Estimasi fungsi regresi nonparametrik dilakukan berdasarkan data pengamatan dengan menggunakan teknik *smoothing*. Terdapat beberapa teknik *smoothing* dalam model regresi nonparametrik antara lain histogram, estimator kernel, deret orthogonal, estimator *spline*, *k-NearestNeighbor*, deret *fourier*, dan *wavelet* (Eubank, 1998).

Estimasi densitas kernel adalah suatu metode estimasi terhadap fungsi densitas yang belum diketahui dengan menggunakan fungsi kernel. Estimator Kernel merupakan pengembangan dari estimator histogram. Estimator kernel diperkenalkan oleh Rosenbalt (1956) dan Parzen (1962) sehingga disebut estimator densitas kernel Rosenbalt-Parzen (Hardle, 1994). Penghalusan dengan estimasi kernel yang selanjutnya dikenal sebagai penghalusan kernel (*kernel smoothing*) yang sangat bergantung pada fungsi kernel dan *bandwidth* (Hardle, 1994). Terdapat beberapa jenis fungsi kernel, antara lain fungsi kernel *Uniform*, *Triangle*, *Epanechnikov*, Kuartik, *Triweight*, *Cosinus*, dan *Gaussian*.

Pada penelitian ini akan digunakan estimator kernel Nadaraya-Watson yang akan mengestimasi fungsi kernel. Estimator kernel Nadaraya-Watson tersebut akan diterapkan pada beberapa fungsi kernel yaitu fungsi kernel *Epanechnikov* dan fungsi kernel Kuartik.

Pada penelitian ini akan menunjukkan perbedaan hasil yang diperoleh diantara kedua fungsi kernel tersebut.

Pada implementasinya, penulis mengambil masalah yang dapat diselesaikan dengan analisis regresi nonparametrik kernel dengan menggunakan estimator Nadaraya-Watson, mengenai permasalahan pengangguran yang ada di Provinsi Sulawesi Tengah tahun 2020. Indeks pembangunan Manusia (IPM) sendiri digunakan sebagai pengukuran kualitas hidup suatu negara atau suatu daerah. Indeks pembangunan manusia di Indonesia masih tergolong rendah dibandingkan negara-negara di dunia. Sedangkan di provinsi Sulawesi Tengah sendiri, walaupun indeks pembangunan manusia (IPM) tahun 2020 mengalami peningkatan tetapi masih tergolong rendah dibandingkan daerah-daerah lainnya di Indonesia.

Salah satu faktor yang mempengaruhi perbedaan IPM antar provinsi adalah Pendapatan masih dipercaya sebagai indikator mengukur kesejahteraan masyarakat, artinya seseorang harus bekerja untuk mendapatkan pendapatan. Pendapatan yang sampai saat ini belum tergolong sejahtera mengakibatkan banyaknya persentase penduduk miskin. Fakta yang terjadi bukannya masyarakat enggan untuk bekerja namun sulitnya mencari pekerjaan menyebabkan masyarakat tidak bisa memenuhi kebutuhannya sehingga tingkat pengangguran di Indonesia semakin tinggi. Pengangguran adalah penduduk yang tidak bekerja namun sedang mencari pekerjaan atau sedang mempersiapkan suatu usaha baru atau penduduk yang tidak mencari pekerjaan karena sudah diterima bekerja tetapi belum mulai bekerja.

Oleh karena itu pada penelitian ini, penulis ingin melihat bagaimana hubungan persentase penduduk miskin dan tingkat pengangguran terhadap indeks pembangunan manusia di Sulawesi Tengah menggunakan metode regresi kernel Nadaraya-Watson dengan perbandingan regresi antara fungsi kernel normal dan fungsi kernel kuadratik.

2. Tinjauan Pustaka

2.1 Regresi Nonparametrik

Menurut Eubank (1988) dalam Fathurahman (2011:54) regresi nonparametrik adalah suatu teknik analisis data yang menjelaskan hubungan antara variabel prediktor dengan variabel respon yang tidak diketahui bentuk fungsinya. Tetapi diasumsikan mulus (*smooth*) dalam suatu ruang fungsi tertentu sehingga regresi nonparametrik sangat mempertahankan fleksibilitasnya.

Ada beberapa teknik *smoothing* yang digunakan dalam metode nonparametrik antara lain: estimator histogram, kernel, deret orthogonal, spline, k-NN, deret fourier, dan wavelet (Komang dan Gusti, 2012).

Model regresi nonparametrik secara matematis dapat ditulis:

$$y = m(x) + \varepsilon$$

(1)

keterangan:

y = variabel respon

$m(x)$ = fungsi regresi nonparametrik yang memuat variabel prediktor

ε = *error* (galat) absolute, dirumuskan dengan $\varepsilon = y - m(x)$. Jika tanda *error* (positif atau negatif) tidak dipertimbangkan maka diperoleh *error* mutlak sebagai berikut: $|\varepsilon| = |y - m(x)|$

2.2 Estimator Kernel

Estimator kernel merupakan pengembangan dari estimator histogram. Suatu histogram disusun dengan meletakkan titik-titik data ke dalam suatu bin atau kelas. Setiap bin dinyatakan secara grafik oleh segiempat dengan lebar sama dan tinggi proporsional dengan banyaknya titik-titik data yang terletak dalam bin tersebut (Adisantoso, 2010 dalam Kurniasih, 2013). Estimator kernel ini umum digunakan dalam model pendekatan nonparametrik. Hal ini disebabkan estimator densitas mempunyai beberapa kelebihan, yaitu:

- Estimator kernel mempunyai bentuk yang fleksibel dan secara matematik mudah dikerjakan.
- Estimator kernel mempunyai rata-rata kekonvergenan yang relatif cepat (Hardle, 1990).

Estimator kernel ini disebut estimator densitas kernel Rosenblatt-Parzen karena dikenalkan pertama kali oleh Parzen (1962) dan Rosenblatt (1956) (Klemela, 1957). Menurut Eubank (1998) pada dasarnya estimator kernel sama dengan estimator linier lainnya hanya saja metode kernel lebih khusus dalam penggunaan metode *bandwidth*.

Menurut Halim dan Bisono (2006:75) estimator kernel dibagi menjadi tiga macam, yaitu:

- Estimator Nadaraya-Watson

$$\hat{m}(x) = \frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)Y_i}{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)}$$

(2)

- Estimator Priestley-Chao

$$\hat{m}(x) = \frac{1}{h} \sum_{i=1}^n (x - x_{i-1}) Y_i K\left(\frac{x-X_i}{h}\right)$$

(3)

- Estimator Gasser-Muller

$$\hat{m}(x) = \frac{1}{h} \sum_{i=1}^n Y_i \int_{s_{i-1}}^{s_i} K\left(\frac{x-X_i}{h}\right) dx$$

(4)

Dimana :

$$s_0 = 0$$

$$s_i = \left(\frac{x_i + x_{i+1}}{2}\right)$$

$$i = 1, \dots, n - 1 \text{ dan } s_n = 1$$

Keterangan:

$\hat{m}(x)$ = fungsi taksiran regresi

Y_i = variabel respon pada data ke-i

X_i = variabel prediktor pada data ke-i

K = fungsi kernel

n = ukuran sampel/banyaknya sampel

h = *bandwidth* atau *smoothing parameter*.

2.3 Fungsi Kernel

Suatu fungsi kernel harus merupakan fungsi kontinyu, berharga riil, simetris, dan terbatas. Menurut Hardle (1994), secara umum fungsi kernel didefinisikan sebagai berikut:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right) \quad (5)$$

untuk $-\infty < x < \infty, h > 0$

Dengan : K = fungsi kernel

h = *bandwidth* atau *smoothing parameter*

Fungsi kernel di atas harus memenuhi beberapa syarat, yaitu:

- (i) $K(x) \geq 0$, untuk semua x
- (ii) $\int_{-\infty}^{\infty} K(x) dx = 1$
- (iii) $\int_{-\infty}^{\infty} x^2 K(x) dx = \sigma^2 > 0$
- (iv) $\int_{-\infty}^{\infty} x K(x) dx = 0$

Pada estimator kernel terdapat beberapa fungsi kernel yang umum digunakan untuk estimasi data, seperti pada tabel berikut.

Tabel 1. Fungsi-fungsi Kernel

No	Kernel	$K(x)$
1	Uniform	$\frac{1}{2} I(x \leq 1)$
2	Segitiga (<i>Triangle</i>)	$(1 - x) I(x \leq 1)$
3	Epanechnikov	$\frac{3}{4} (1 - x^2)^2 I(x \leq 1)$
4	Kuadrat (Quartik)	$\frac{15}{16} (1 - x^2)^2 I(x \leq 1)$
5	Triweight	$\frac{35}{32} (1 - x^2)^3 I(x \leq 1)$
6	Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) - \infty < x < \infty$
7	Cosinus	$\frac{\pi}{4} \cos\left(\frac{\pi}{2} x\right) I(x \leq 1)$
8	Tricube	$\frac{70}{81} (1 - x ^3)^3 I(x \leq 1)$
9	Logistik	$\frac{1}{e^x + 2 + e^{-x}}$

(Sumber: Indrayanti, 2004)

Pada estimator kernel terdapat beberapa fungsi kernel yang umum digunakan untuk estimasi data, salah satunya fungsi kernel normal dan fungsi kernel kuadrat.

2.4 Regresi Kernel

Salah satu metode untuk mengestimasi model regresi nonparametrik adalah regresi kernel. Regresi kernel merupakan teknik statistik nonparametrik untuk menaksir nilai ekspektasi bersyarat dari suatu variabel acak. Nilai ekspektasi umumnya dinotasikan dengan $E(Y|X)$. Regresi kernel bertujuan untuk mendapatkan hubungan nonlinier antara variabel X dan Y .

Salah satu teknik regresi nonparametrik yang sering digunakan untuk menaksir fungsi regresi $m(x)$ adalah dengan menggunakan estimator Nadaraya- Watson. Estimator ini diperoleh dengan menggunakan metode penaksiran fungsi densitas kernel. Fungsi densitas peluang bersama $f(x, y)$ diduga dengan perkalian kernel sebagai berikut :

$$\hat{f}(x, y) = \hat{f}_{h_1 h_2}(x, y) = \frac{1}{n} \sum_{i=1}^n K_{h_1}(x - X_i) K_{h_2}(y - Y_i)$$

(6)

Taksiran penyebut pada persamaan Nadaraya-Watson dapat diperoleh dari integralfungsi kepadatan bersama $f(x, y)$:

$$\begin{aligned} \int y \hat{f}(x, y) dy &= \frac{1}{n} \sum_{i=1}^n K_{h_1}(x - X_i) \int y K_{h_2}(y - Y_i) dy \\ &= \frac{1}{n} \sum_{i=1}^n K_{h_1}(x - X_i) \int \frac{y}{h_2} K\left(\frac{y - Y_i}{h_2}\right) dy \\ &= \frac{1}{n} \sum_{i=1}^n K_{h_1}(x - X_i) \int (sh_2 + Y_i) K(s) ds \\ &= \frac{1}{n} \sum_{i=1}^n K_{h_1}(x - X_i) Y_i \end{aligned}$$

(7)

dimana $\int (sh_2 + Y_i) K(s) ds = Y_i$ (Kartika, 2000 dalam Indrayanti, 2014).

Sedangkan taksiran penyebutnya adalah taksiran kepadatan kernel. Sehingga dari kombinasi kedua taksiran probabilitas bersyarat pada persamaan akan diperoleh persamaan Nadaraya-Watson, yaitu:

$$\hat{m}(x) = \frac{\frac{1}{n} \sum_{i=1}^n K_h(x - X_i) Y_i}{\frac{1}{n} \sum_{i=1}^n K_h(x - X_i)}$$

(8)

Dimana

$$\sum_{i=1}^n K_h(x - X_i) = \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)$$

(9)

Sehingga

$$\hat{m}(x) = \frac{\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) Y_i}{\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)}$$

(10)

dengan K adalah fungsi kernel dan h adalah *bandwidth* atau *smoothing parameter* dan pengontrol kemulusan (Halim dan Bisono, 2006).

Kemudian disubstitusikan sehingga menjadi :

$$y = \frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)Y_i}{\sum_{i=1}^n K\left(\frac{x-X_i}{h}\right)} + \varepsilon$$

(11)

2.5 Pemilihan Bandwidth Optimum

Bandwidth dinotasikan dengan h adalah konstanta positif untuk menentukan kemulusan dari kurva taksiran regresi. Dengan memilih *bandwidth* yang mendekati 0 maka taksiran yang didapatkan tidak mulus bahkan akan menonjolkan plot datanya saja.

Pada regresi kernel pemilihan *bandwidth* jauh lebih penting daripada pemilihan fungsi kernel. Jika *bandwidth* yang dipilih terlalu kecil maka akan menghasilkan estimasi kurva kurang mulus (*under-smoothing*), sebaliknya jika *bandwidth* terlalu besar maka akan menghasilkan estimasi kurva sangat mulus (*over-smoothing*) yang tidak sesuai dengan pola sebaran data. Sehingga harus dipilih nilai *bandwidth* yang optimum agar dihasilkan estimasi terbaik.

Terdapat beberapa metode yang digunakan dalam pemilihan *bandwidth* optimum, salah satunya adalah menggunakan kriteria *Generalized Cross Validation* (GCV) (Galub, dkk, 1979), didefinisikan dengan:

$$GCV = \frac{MSE}{\left(\frac{1}{n}tr(I-H(h))\right)^2}$$

(12)

dengan

n = banyaknya data

I = matriks identitas

h = bandwidth

X = matriks data

$H(h) = X(X'X + nhl)^{-1}X'$

$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - m_h(x_i))^2$

Menurut Komang dan Gusti (2012), kebaikan suatu estimator dapat dilihat dari tingkat kesalahannya. Terdapat beberapa kriteria untuk menentukan estimator terbaik dalam model regresi nonparametrik, di antaranya:

a. *Mean Square Error* (MSE)

Untuk mengukur *error* biasanya digunakan *Mean Square Error*. Estimator terbaik dipilih berdasarkan nilai MSE terkecil. *Mean Square Error* (MSE) adalah rata-rata dari kuadrat kesalahan.

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

(13)

b. *Root Mean Square Error* (RMSE)

$$RMSE = \sqrt{MSE}$$

(14)

c. *Mean Absolute Deviation* (MAD)

$$MAD = \frac{1}{n} \sum_{i=1}^n |e_i| = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

(15)

Keterangan:

n = banyaknya data

y_i = data sebenarnya

\hat{y}_i = nilai prediksi dari variabel y_i

2.6 Indeks Pembangunan Manusia

Indeks pembangunan manusia merupakan indikator strategis yang banyak digunakan untuk melihat upaya dan kinerja program pembangunan secara menyeluruh di suatu wilayah. Dalam hal ini IPM dianggap sebagai gambaran dari hasil program pembangunan yang telah dilakukan beberapa tahun sebelumnya. Demikian juga kemajuan program pembangunan dalam suatu periode dapat diukur dan ditunjukkan oleh besaran IPM pada awal dan akhir periode tersebut (Saputra, 2011). Indeks pembangunan manusia (IPM) atau Human Development Index (HDI) adalah pengukuran perbandingan dari harapan hidup, melek huruf, pendidikan dan standar hidup untuk semua negara seluruh dunia. Indeks pembangunan manusia digunakan untuk mengklasifikasikan apakah sebuah negara adalah negara maju, negara berkembang atau negara terkebelakang dan juga untuk mengukur pengaruh dari kebijaksanaan ekonomi terhadap kualitas hidup.

2.7 Persentase Penduduk Miskin

Penduduk Miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan dibawah garis kemiskinan. Daerah tertinggal (atau Kabupaten tertinggal) adalah daerah kabupaten yang wilayah serta masyarakatnya kurang berkembang dibandingkan dengan daerah lain dalam skala nasional. Kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Daerah/kabupaten tertinggal ditetapkan berdasarkan kriteria sebagai berikut: 1. Perekonomian masyarakat 2. Sumber daya manusia 3. Sarana dan prasarana 4. Kemampuan keuangan daerah 5. Aksesibilitas 6. Karakteristik daerah.

2.8 Pengangguran

Pengangguran adalah seseorang yang sudah digolongkan dalam angkatan kerja, yang secara aktif sedang mencari pekerjaan pada suatu tingkat upah tertentu, tetapi tidak dapat memperoleh pekerjaan yang diinginkannya (Sukirno :2010). Sedangkan pengangguran terbuka dapat berarti angkatan kerja yang belum bekereja dan sedang aktif mencari pekerjaan (Masriah dan Muhahid, 2011: 54).

3. Metode Penelitian

3.1 Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) Sulawesi Tengah. Data tersebut merupakan

data Indeks Pembangunan Manusia, Tingkat Pengangguran Terbuka dan Persentase Penduduk Miskin di Sulawesi Tengah pada Tahun 2020. Adapun datanya sebagai berikut:

Tabel 2. Data Indeks Pembangunan Manusia, Persentase Penduduk Miskin dan Tingkat Pengangguran terbuka di Sulawesi Tengah Tahun 2020

Kabupaten/Kota	Y	X ₁	X ₂
Banggai Kepulauan	65,42	14,04	72,97
Banggai	70,52	7,39	69,79
Morowali	72,21	13,43	64,70
Poso	71,28	15,45	74,78
Donggala	65,56	17,39	63,93
Toli-toli	65,69	12,85	65,95
Buol	67,82	13,93	69,75
Parigi Moutong	65,44	15,85	72,50
Tojo Una-una	64,59	16,39	75,33
Sigi	68,12	12,45	69,24
Banggai Laut	65,43	14,60	67,20
Morowali Utara	68,36	14,10	69,85
Kota Palu	81,47	6,80	66,46

Keterangan :

Y = Indeks Pembangunan Manusia

X₁ = Persentase Penduduk Miskin

X₂ = Tingkat Pengangguran terbuka

3.2 Variabel Penelitian

Pada penelitian ini menggunakan dua variabel yakni variabel respon dan variabel prediktor sebagai berikut:

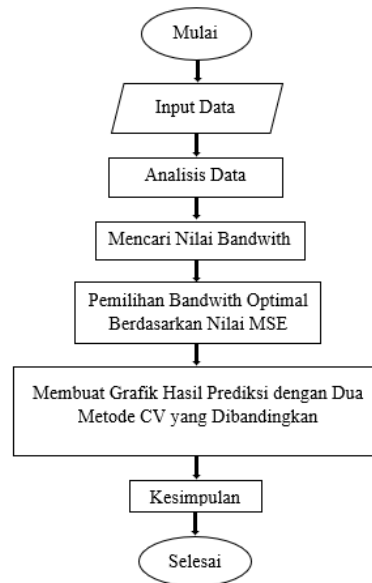
Varibel	Definisi Operasional	Skala
Variabel Respon (Y)	Indeks Pembangunan Manusia (IPM) Menurut Kabupaten/Kota Provinsi Sulawesi Tengah 2020	Rasio
Varibel Prediktor (X ₁)	Penduduk Miskin(%)Menurut Kabupaten/Kota Provinsi Sulawesi Tengah 2020	Rasio
Varibel Prediktor (X ₂)	Tingkat Pengangguran Terbuka (TPT) Menurut Kabupaten/Kota Provinsi Sulawesi Tengah 2020	Rasio

3.3 Analisis Data

Pada penelitian ini menggunakan Analisis Kuantitatif karena data yang digunakan adalah Data Sekunder yang diperoleh dari Publikasi BPS Sulawesi Tengah Tahun 2020. Metode Analisis yang digunakan yaitu Statistika Deskriptif, selain itu metode selanjutnya menggunakan Analisis Regresi Statistika Nonparametrik Regresi Kernel.

3.4 Kerangka Pikir

Pada penelitian ini menggunakan Analisis Kuantitatif karena data yang digunakan adalah Data Sekunder yang diperoleh dari Publikasi BPS Sulawesi Tengah Tahun 2020. Metode Analisis yang digunakan yaitu Statistika Deskriptif, selain itu metode selanjutnya menggunakan Analisis Regresi Statistika Nonparametrik Regresi Kernel.



Gambar 1. Kerangka Pikir

4. Hasil dan Pembahasan

4.1 Statistik Deskriptif

Tabel 3. Statistik Deskriptif

	Minimum	Median	Mean	Maximum
Y	64.59	67.82	68.61	81.47
X_1	6.80	14.04	13.44	17.39
X_2	63.93	69.75	69.42	75.33

Berdasarkan output di atas, untuk Persentase Indeks Pembangunan Manusia memiliki rata-rata sebesar 68.61%. Untuk Persentase Indeks Pembangunan Manusia terendah adalah Kabupaten Tojo Una-Una yaitu 64.59%, sedangkan Persentase Indeks Pembangunan Manusia tertinggi adalah Kota Palu yaitu 81.47%. Pada data Presentase Penduduk Miskin memiliki rata-rata sebesar 13.44%. Data Presentase Penduduk Miskin terendah berada pada Kota Palu yaitu sebesar 6.80%, sedangkan Presentase Penduduk Miskin tertinggi adalah Kabupaten Donggala yaitu 17.38%. Untuk data Persentase Tingkat Pengangguran Terbuka rata-rata sebesar 69.42%, dengan Persentase Tingkat Pengangguran Terbuka terendah adalah Kabupaten Donggala yaitu 63.93%, sedangkan Persentase Tingkat Pengangguran Terbuka tertinggi adalah Kabupaten Tojo Una-Una yaitu sebesar 75.33%.

4.2 Bandwidth Optimum

Tabel 4. Nilai Bandwidth Optimum

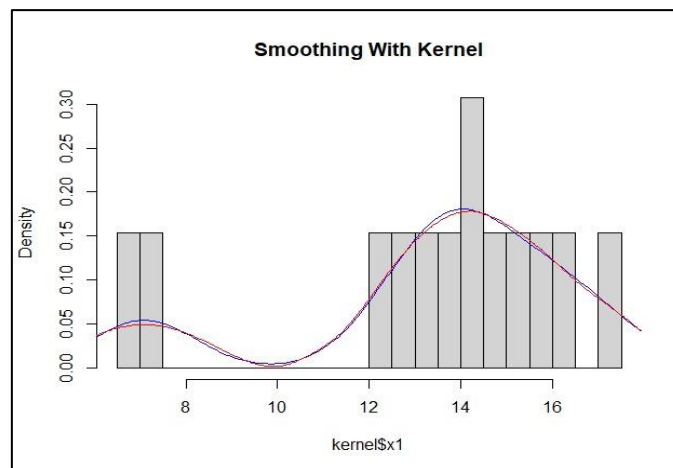
	Jenis Metode CV	Bandwidth	R-Square	MSE
X_1	CV.AIC	2.825092	0.5643457	1.523095
	CV.LS	3.294538	0.5644725	1.369349
X_2	CV.AIC	35230241	0.09443642	1.331878
	CV.LS	31240304	0.001502402	1.96234

Berdasarkan tabel di atas, dapat dilihat bahwa nilai MSE pada X_1 dan X_2 dengan metode CV.AIC adalah sebesar 1.523095 dan 1.331878, sedangkan nilai MSE pada X_1 dan X_2 dengan metode CV.LS sebesar 1.369349 dan 1.96234. Untuk melihat bandwidth optimal untuk X_1 dan X_2 dapat dilakukan dengan melihat nilai MSE terkecil. Untuk model terbaik pada X_1 dengan nilai MSE terkecil yaitu pada metode CV.LS dengan nilai bandwidth sebesar 1.369349, dan untuk model terbaik pada X_2 dengan nilai MSE terkecil yaitu pada metode CV.AIC dengan nilai bandwidth sebesar 1.331878.

4.3 Perbandingan Fungsi Kernel Normal dan Fungsi Kernel Kuadratik

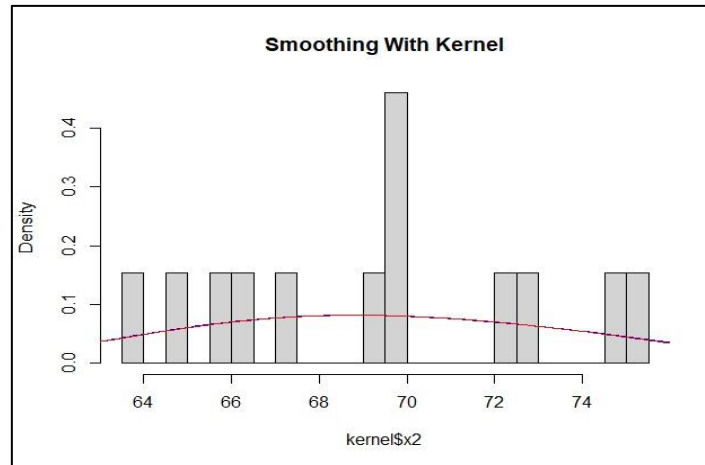
Tabel 5. Nilai Bandwidth Optimum

	Normal	Kuadratik
X_1	1.085826	2.847704
X_2	2.736757	7,177462



Gambar 2. Smoothing With Kernel Persentase Penduduk Miskin

Berdasarkan plot hasil estimasi variabel Presentase Penduduk Miskin untuk fungsi Kernel Normal dan fungsi Kernel Kuadratik dengan menggunakan bandwidth optimal, garisnya menunjukkan sedikit perbedaan dan berjauhan dititik puncak sehingga dapat dikatakan bahwa penggunaan Fungsi kernel yang berbeda dengan bandwidth yang optimal untuk masingmasing fungsi kernel tersebut akan menghasilkan estimasi kurva yang berbeda.



Gambar 2. Smoothing With Kernel Tingkat Pengangguran terbuka

Berdasarkan plot hasil estimasi variabel Pengangguran untuk fungsi Kernel Normal dan fungsi Kernel Kuadratik dengan menggunakan bandwidth optimal, garisnya sangat berhimpit sehingga dapat dikatakan bahwa penggunaan Fungsi kernel yang berbeda dengan bandwidth yang optimal untuk masing-masing fungsi kernel tersebut akan menghasilkan estimasi kurva yang sama.

5. Kesimpulan

- Estimasi densitas kernel adalah suatu metode estimasi terhadap fungsi densitas yang belum diketahui dengan menggunakan fungsi kernel. Estimator Kernel merupakan pengembangan dari estimator histogram. Estimator kernel diperkenalkan oleh Rosenbalt (1956) dan Parzen (1962) sehingga disebut estimator densitas kernel Rosenbalt-Parzen (Hardle, 1994). Penghalusan dengan estimasi kernel yang selanjutnya dikenal sebagai penghalusan kernel (*kernel smoothing*) yang sangat bergantung pada fungsi kernel dan *bandwidth* (Hardle, 1994). Terdapat beberapa jenis fungsi kernel, antara lain fungsi kernel *Uniform*, *Triangle*, *Epanechnikov*, Kuartik, *Triweight*, *Cosinus*, dan *Gaussian*.
- Dalam regresi kernel yang terpenting adalah pemilihan nilai bandwidth optimal, bukan pemilihan fungsi kernel. Karena penggunaan fungsi kernel yang berbeda dengan nilai bandwidth optimal menghasilkan estimasi kurva yang hampir sama. Pada kasus ini, Untuk model terbaik pada X_1 dengan nilai MSE terkecil yaitu pada metode CV.LS dengan nilai bandwidth sebesar 1.369349, dan untuk model terbaik pada X_2 dengan nilai MSE terkecil yaitu pada metode CV.AIC dengan nilai bandwidth sebesar 1.331878.

Daftar Pustaka

- Abdullah. 2003. Tafsir Ibnu Katsir, Jilid 4. Terjemahan M. Abdul Ghoffar E.M. Bogor: Pustaka Imam Asy-Syafi'i.
- Abdullah. 2004a. Tafsir Ibnu Katsir, Jilid 5. Terjemahan M. Abdul Ghoffar E.M, dkk. Bogor: Pustaka Imam Asy-Syafi'i.
- Abdullah. 2004b. Tafsir Ibnu Katsir, Jilid 6. Terjemahan M. Abdul Ghoffar E.M dan Abu Ihsan Al-Atsari. Bogor: Pustaka Imam Asy-Syafi'i.

- Abdullah. 2004c. Tafsir Ibnu Katsir, Jilid 8. Terjemahan M. Abdul Ghoffar E.Mdan Abu Ihsan Al-Atsari. Bogor: Pustaka Imam Asy-Syafi'i.
- Ad-Dimasyqi, A.. 2000. Tafsir Ibnu Katsir (Juz 6). Jakarta: Sinar Baru Algensindo.
- Al-Jazairi, S.. 2007. Tafsir Al-Qur'an Al-Aisar Surat: Al-A'raaf-Yusuf (Jilid 3). Jakarta: Darus Sunnah Press.
- Al-Jazairi, S.. 2008. Tafsir Al-Qur'an Al-Aisar Surat: Al-Mu'minuun-Al-Ahzab(Jilid 5). Jakarta: Darus Sunnah Press.
- Al-Jazairi, S.. 2009. Tafsir Al-Qur'an Al-Aisar Surat: Saba'-Al-Hujuraat (Jilid 6). Jakarta: Darus Sunnah Press.
- Dasuki, H. dan Dahlan, Z.. 1995. Al-Qur'an dan Tafsirnya. Yogyakarta: PT. Dana Bhakti Wakaf.
- Eubank, R.. 1998. Spline Smoothing and Nonparametric Regression. New York:Marcel Dekker.
- Fathurahman, M.. 2011. Estimasi Parameter Model Regresi Spline. *Jurnaleksponensial* 2 (1): 2085–7829.
- Fox, C.G.. 2009. Solar-Geophysical Data (January 2009 Number 773-Part I).(Online), (<http://www.ngdc.noaa.gov>), diakses 18 Juli 2014.
- Galub, G.H., Heath, M., dan Wahba, G.. 1979. Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics*, 21 (2): 215–223 .
- Halim, S. dan Bisono, I.. 2006. Fungsi-Fungsi Kernel pada Metode Regresi Nonparametrik dan Aplikasinya pada Priest River Experimental Forest's Data. *Jurnal Teknik Industri*, 8 (1): 73–81.
- Hardle, W.. 1990. Smoothing Techniques with Implementation in S. New York: Springer Verlag.
- Hardle, W.. 1994. Applied Nonparametric Regression. Cambridge: Cambridge University Press.
- <http://diomishinegi.blogspot.co.id/2014/02/pengertian-dan-penjelasan-ipm-indeks>
- Indrayanti, A.I.. 2014. Estimator Kernel Cosinus dan Kernel Gaussian dalam Model Regresi Nonparametrik pada Data Butterfly Diagram Siklus Aktivitas ke-23 (Studi Kasus di BPD LAPAN Watukosek). Skripsi tidak dipublikasikan. Malang: UIN Maulana Malik Ibrahim Malang.
- Kaufmann, W.J.. 1978. Exploration of The Solar System. New York: Macmillan Publishing Co.
- Klemela, J.. 1965. Multivariate Nonparametric Regression and Visualization with R and Applications to Finance. Finland: Wiley.
- Komang, G. dan Gusti, A.. 2012. Estimator Kernel dalam Model Regresi Nonparametrik. *Jurnal Matematika*, 2 (1): 19–30.
- Kurniasih, D.. 2013. Efisiensi Relatif Estimator Fungsi Kernel Gaussian terhadap Estimator Polinomial dalam Peramalan USD terhadap JPY. Skripsi tidak dipublikasikan. Semarang: Universitas Negeri Semarang.
- Nugraha, N.. 2009. Regresi Polinomial Lokal. Skripsi tidak dipublikasikan. Depok: Universitas Indonesia.
- Puspitasari, I., Suparti, dan Wilandari, Y.. 2012. Analisis Indeks Harga Saham Gabungan (IHSG) dengan Menggunakan Model Regresi Kernel. *Jurnal Gaussian*, 1 (1): 93–102.
- Saputra, 2011. Analisis Pengaruh Jumlah Penduduk, PDRB, IPM, Pengangguran Terhadap Tingkat Kemiskinan Di Jawa Tengah. Skripsi, Universitas Diponegoro.

- Sudarno. 2011. Pemulusan Sebaran Data Menggunakan Penaksiran Kernel Nadaraya-Watson dan Linier Lokal untuk Kernel Normal. Prosiding Seminar Nasional Statistik.
- Volvacea, V.. 2012. Klasifikasi dan Perubahan Jumlah Sunspot Diamati dari Laboratorium Astronomi Jurusan Fisika FMIPA UM pada Bulan Agustus–Oktober 2012. Jurnal Fisika, 1 (1): 1–6.