

jurnal1

by dodi mulyadi

General metrics

12,247

characters

1,569

words

45

sentences

6 min 16 secreading
time**12 min 4 sec**speaking
time

Score



99

This text scores better than 99%
of all texts checked by Grammarly

Writing Issues

8

Issues left

4

Critical

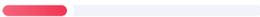
4Advanced

Plagiarism



This text seems 100% original. Grammarly found no matching text on the Internet or in ProQuest's databases.

Writing Issues

8	Correctness	
2	Unknown words	
1	Misspelled words	
4	Punctuation in compound/complex sentences	
1	Confused words	

Unique Words

Measures vocabulary diversity by calculating the percentage of words used only once in your document

4%unique words

Rare Words

Measures depth of vocabulary by identifying words that are not among the 5,000 most common English words.

38%rare words

Word Length

Measures average word length

0.3characters per word

Sentence Length

Measures average sentence length

34.9words per sentence

jurnal1

Statistika, Vol. 6, No. 2, November 2018

160

PEMODELAN PRODUKSI PADI MENGGUNAKAN REGRESI SEMIPARAMETRIK KERNEL

Tiani Wahyu Utami¹, Indah Manfaati Nur², Endah Suryaningsih³

^{1,2,3} Program Studi Statistika, FMIPA, Universitas Muhammadiyah Semarang

Alamat e-mail : tianitutami@unimus.ac.id

ABSTRAK

Negara Indonesia merupakan negara agraris dengan salah satu bahan makanan pokok di Indonesia adalah padi. Produksi padi di provinsi Jawa Timur merupakan tertinggi di Indonesia. Tujuan penelitian ini adalah melakukan estimasi parameter parametrik dan nonparametrik serta memodelkan produksi padi di 38 kabupaten/kota provinsi Jawa Timur menggunakan regresi semiparametrik kernel. Regresi semiparametrik merupakan gabungan dari regresi parametrik dan nonparametrik. Regresi parametrik memiliki kurva yang berpola, misal linier, dll. Regresi nonparametrik memiliki kurva yang smooth

yang tidak diketahui polanya, sehingga dalam hal ini diperlukan teknik smoothing yang digunakan untuk menghaluskan kurva yaitu salah satunya adalah kernel dengan estimator parameter Nadaraya-Watson dan metode estimasi parameter menggunakan Weighted Least Square (WLS), serta pemilihan bandwidth (h) optimal dengan metode Generalized Cross Validation (GCV). Variabel yang digunakan dalam penelitian yaitu produksi padi sebagai variabel respon, sedangkan variabel prediktornya yaitu luas panen, produktivitas, dan luas kerusakan. Data yang digunakan merupakan data sekunder dari website resmi Badan Pusat Statistika (BPS) provinsi Jawa Timur. Berdasarkan hasil estimasi menunjukkan bahwa regresi semiparametrik kernel mendapatkan nilai bandwidth (h) optimalnya 0,006 dengan $GCV=15058865548$. Koefisien determinasi (R^2) sebesar 92,46% dan MSE sebesar 5797077303.

Kata kunci : Produksi Padi, Regresi Semiparametrik, GCV, Nadaraya-Watson, WLS

PENDAHULUAN

Negara Indonesia merupakan negara agraris. sehingga memiliki hasil pangan yang melimpah. Hasil pangan di Indonesia berupa tanaman padi, jagung, dll. Tanaman pangan yang utama di Indonesia yaitu tanaman padi. Orang Indonesia lebih memilih mengonsumsi hasil olahan padi sebagai makanan pokok dibandingkan yang lainnya. Meskipun demikian, bahan makanan pokok lainnya tetap dikonsumsi sebagai produk sampingan.

Indonesia sebagai negara agraris seharusnya memiliki hasil produksi padi yang melimpah, sehingga tidak perlu mengadakan impor beras dari negara tetangga.

Jawa Timur merupakan provinsi yang paling banyak menghasilkan produksi padi di Negara Indonesia pada tahun 2015 [2]. Meskipun pada tahun 2015 provinsi tersebut berada diposisi pertama di Indonesia mengenai produksi padinya. Meskipun produksi padinya tertinggi produksinya, harus memiliki simpanan padi untuk kebutuhan pangan dimasa yang akan datang guna memenuhi konsumsi sehari-harinya.

Faktor-faktor yang mempengaruhi produksi padi antara lain yaitu luas panen, produktivitas, dan luas puso/kerusakan[7]. [3] mendefinisikan luas panen merupakan luas lahan sawah yang biasa diambil hasilnya. Produktivitas merupakan hasil yang diperoleh tiap satuan luas. Luas rusak adalah jika tanaman mengalami serangan organisme pengganggu tumbuhan, bencana alam, sedemikian rupa sehingga hasilnya kurang dari 11% keadaan normal. Analisis regresi merupakan salah satu metode untuk mengetahui hubungan antara variabel respon (Y) dengan variabel prediktor (X). Analisis regresi terdapat tiga macam pendekatan, yaitu parametrik, nonparametrik, dan semiparametrik. Pendekatan parametrik yaitu pendekatan dengan variabel data yang polanya diketahui. Pendekatan nonparametrik yaitu pendekatan dengan variabel data dengan kurva smooth yang polanya tidak diketahui, sehingga data tersebut akan mencari bentuk kurvanya sendiri [5]. Regresi semiparametrik merupakan gabungan dari regresi parametrik dan nonparametrik [8]. Estimasi kurva sangat bergantung pada perilaku data, sehingga diperlukan teknik smoothing. Pendekatan semiparametrik yaitu gabungan dari pendekatan parametrik dan nonparametrik, sehingga pada pendekatan semiparametrik lebih tepat apabila diterapkan pada kasus yang mengandung komponen parametrik dan nonparametrik [9]. Pada umumnya, teknik smoothing yang sering digunakan dalam penelitian adalah kernel. Pada

penelitian ini digunakan estimator Nadaraya-Watson dan fungsi kernel Gaussian.

Regresi kernel merupakan teknik statistika nonparametrik yang digunakan untuk mengestimasi fungsi yang terdapat dalam model regresi nonparametrik. Nadaraya dan Watson pada tahun 1994 mendefinisikan estimator regresi kernel, sehingga disebut estimator Nadaraya-Watson [6]. Berikut merupakan estimator Nadaraya-Watson:

Bandwidth (h) adalah parameter pemulus (smoothing) yang memiliki fungsi untuk mengontrol kurva yang sedang diestimasi. Generalized Cross Validation (GCV) merupakan salah satu metode dalam analisis regresi untuk menentukan bandwidth yang optimal [10].

METODE PENELITIAN

Sumber Data dan Variabel Penelitian

Data yang digunakan bersumber dari website resmi Badan Pusat Statistika (BPS) di 38 kabupaten/kota provinsi Jawa Timur tahun 2015. Data yang digunakan yaitu sebagai berikut :

Tabel 1. Variabel Penelitian

Data

Variabel

Produksi Padi

Respon (y)

Luas Panen

Parametrik (x)

Produktivitas

Nonparametrik (t_1)

Luas Puso

Nonparametrik (t_2)

Variabel-variabel tersebut berpacu dari penelitian-penelitian terdahulu yaitu oleh penelitian [7] dan penelitian [1].

Metode Analisis

Langkah-langkah yang digunakan dalam melakukan penelitian regresi semiparametrik kernel Nadaraya-Watson yaitu sebagai berikut :

Mengkaji estimasi model regresi semiparametrik kernel sebagai berikut :

a. Diberikan data observasi (y_i, x_i, t_i) yang memenuhi model regresi

semiparametrik :

$, j=1,2,3, \dots, n$

b. Dari langkah a tersebut dapat dinyatakan sebagai berikut :

Atau dalam bentuk matriks dapat ditulis sebagai berikut

dengan

c. Menyatakan $m_j(t_j)$ dapat didekati dengan estimator Nadaraya-Watson :

d. Mengestimasi $m_j(t_j)$ dengan meminimalkan kriteria Weighted Least Square (WLS) :

Dengan $K_h = \frac{1}{h} K \cdot h$ fungsi pembobot kernel

e. Untuk mendapatkan matriks $A(h)$ yang berukuran $N \times N$ dengan

menyelesaikan persamaan berikut :

atau

f. Mensubstitusikan untuk memperoleh estimasi β yaitu

2. Memodelkan produksi padi menggunakan regresi semiparametrik kernel :

a. Mengumpulkan data-data yang dibutuhkan dalam penelitian

b. Melakukan deskripsi data yang akan digunakan dalam penelitian dengan

cara membuat scatterplot pada setiap variabel

- c. Menentukan variabel komponen parametrik dan variabel komponen nonparametrik
- d. Melakukan pemilihan bandwidth (h) yang optimal dengan nilai Generalized Cross Validation (GCV) yang minimum
- e. Memodelkan nilai bandwidth yang optimal
- f. Menghitung nilai Mean Square Error (MSE)

HASIL PENELITIAN

Estimasi Regresi Semiparametrik Kernel

Diberikan data observasi (y_i, x_i, t_i) sebanyak n , dengan y_i menyatakan variabel respon pengamatan untuk subjek ke- i , sedangkan x_i dan t_i menyatakan variabel prediktor. Berikut bentuk model dengan regresi semiparametrik :
dengan

Persamaan model tersebut mengandung dua komponen yaitu komponen parametrik dan komponen nonparametrik yang dapat diasumsikan bahwa nilai memiliki bentuk fungsi yang diketahui, sehingga dapat dibentuk sebagai berikut :

dengan

Bentuk dari fungsi $m(t_{ij})$ belum diketahui, sehingga didekati dengan estimator Nadaraya-Watson yang dilakukan dengan cara meminimumkan kriteria Weighted Least Square (WLS). Metode Weighted Least Square (WLS) memiliki pembobot sebagai K_h , dengan $K_h = \text{diag}(K_h(t_{i1}-t), K_h(t_{i2}-t), \dots, K_h(t_{i38}-t))$.

Matriks K_h merupakan matriks yang berisi fungsi pembobot dengan $K(\cdot)$ merupakan fungsi kernel dan h yaitu bandwidth. $K(\cdot)$ merupakan fungsi kernel Gaussian, yaitu

Metode estimasi Weighted Least Square (WLS) dalam bentuk matriks sebagai berikut :

Dengan meminimumkan WLS tersebut, didapatkan matriks sebagai berikut :

Sehingga, didapatkan matriks $A(h)$ yaitu

Dengan

Atau dalam bentuk estimator Nadaraya-Watson sebagai berikut :

Pola Hubungan Data Respon dan Prediktor

Gambar 1. Luas panen¹ VS produksi²

Gambar 2. Produktivitas VS produksi³

Gambar 3. Luas Puso VS Produksi

Berdasarkan scatterplot diatas, dapat dikatakan bahwa gambar 1. hubungan produksi padi dengan luas panen termasuk komponen parametric, sedangkan variabel produksi padi dengan produktivitas serta hubungan produksi padi dengan luas puso termasuk komponen nonparametrik.

Tabel 2. Pemilihan Bandwidth (h) Optimal

No

Bandwidth

GCV

1

0,0001

15187967591

2

0,002

15187964224

3

0,006

15058865548

4

0,01

15779117764

Pemilihan bandwidth (h) optimal dilihat dari nilai GCV yang minimum, yaitu dengan nilai GCV 15058865548 terpilih bandwidth (h) optimalnya 0,006.

Penentuan estimasi parameter β pada regresi parametrik yang diperoleh yaitu sebesar 6,20.

Pemodelan Produksi Padi Menggunakan Regresi Semiparametrik Kernel

Estimasi masing-masing parameter yang sudah diperoleh, baik komponen parametrik maupun nonparametrik, maka dapat dibentuk pemodelan regresi semiparametrik kernel sebagai berikut :

Berikut persamaan model semiparametrik kernel adalah sebagai berikut :

Perbandingan Hasil Prediksi Produksi Padi dengan Data Aktual

Gambar 4. Perbandingan Prediksi Padi dengan Data Aktual

Berdasarkan grafik tersebut, dapat dikatakan bahwa data respon aktual dengan data respon prediksi memiliki pola data yang sama, artinya data prediksi mampu menutupi data aktual. Hal tersebut dapat dinyatakan bahwa pemodelan yang diperoleh telah sesuai untuk memprediksi produksi padi. Wilayah dengan nilai prediksi tertinggi yaitu berada di kabupaten Jember, sedangkan wilayah dengan nilai prediksi terendah berada di kota Mojokerto.

Evaluasi Ketepatan Model

Evaluasi ketepatan model dapat dilihat pada nilai koefisien determinasi (R^2) dan Mean Square Error (MSE). Semakin tinggi nilai koefisien determinasi maka

model semakin bagus, serta semakin kecil nilai MSE maka model semakin baik. Hasil olahan program R menunjukkan bahwa nilai koefisien determinasi dan MSE yaitu sebesar 0.9246 dan 5797077303. Artinya, bahwa 92,46% variabel prediktor dapat menjelaskan variansi terhadap variabel respon, sedangkan sisanya yaitu 7,54% dipengaruhi oleh faktor lain.

KESIMPULAN⁴

Hasil estimasi regresi Semiparametrik kernel

sehingga,

atau dalam estimator Nadaraya-Watson adalah sebagai berikut :

Pemodelan Produksi Padi menggunakan Regresi Semiparametrik Kernel :

Model regresi tersebut memiliki nilai koefisien determinasi (R^2) sebesar 92,46% dan nilai MSE sebesar 5797077303. Hal tersebut menunjukkan bahwa variabel prediktor luas panen (x) dan variabel prediktor produktivitas (t_1) dan luas puso (t_2) memiliki kemampuan menjelaskan variansi terhadap variabel respon (y) sebesar 92,46% dan 7,54% dijelaskan oleh variabel lainnya yang tidak termasuk dalam penelitian ini.

UCAPAN TERIMKASIH

Penulis mengucapkan terimakasih kepada DRPM Kemenristek Dikti atas Hibah Penelitian Dosen Pemula 2017-2018 dengan no SK 014/UNIMUS.J/PJ/PG/2018.

DAFTAR PUSTAKA

Arnanda, F, dan Karim, A. 2016. Pemodelan Produksi Padi di Provinsi Jawa Tengah Dengan Pendekatan Spatial Econometrics. Jurnal Statistika, 4(2): 20-

27.

Badan Pusat Statistik. 2015. Data BPS Provinsi Jawa Timur. BPS Jawa Timur. <http://jatim.bps.go.id>. 5 Desember 2017 (10:26).

Badan Pusat Statistik. 2015. Produksi Padi dan Palawija di Jawa Timur 2015-2016. BPS Jawa Timur. Surabaya.

Djiwandi, 1980. Penyuluhan Pertanian. Universitas Sebelas Maret. Surakarta

Eubank, R. L. 1999⁵. Nonparametric Regression and Spline Smoothing. New York: Marcel Dekker, Inc.

Hardle, W. 1994⁶. Applied Nonparametric Regression. Cambridge University Press. New York.

Ishaq, M, dkk. 2016. Analisis Faktor-faktor yang Mempengaruhi Produksi Padi di Provinsi Padi di Jawa Timur Menggunakan Regresi Semiparametrik Spline. Jurnal Sains dan Seni ITS 5(2): 420-425.

Ruppert, D., Wand, M. P., & Carrol, R. J. 2003⁷. Semiparametric Regression. New York: Cambridge University Press.

Utami, T. W. 2013. Estimasi Kurva Regresi Semiparametrik Pada Data Longitudinal Berdasarkan Estimator Polinomial Lokal. Jurnal Statistika 1(1): 30-36.

Wu, H., dan Zhang, J.T. 2006⁸. Nonparametric Regression Methods for Longitudinal Data Analysis. New York: John Wiley and Sons, Inc.

1.	panen → panel	Misspelled Words	Correctness
2.	<i>produksi</i>	Unknown Words	Correctness
3.	<i>produksi</i>	Unknown Words	Correctness
4.	KESIMPULAN → kesimpulan	Confused Words	Correctness
5.	, 1999	Punctuation in Compound/Complex Sentences	Correctness
6.	, 1994	Punctuation in Compound/Complex Sentences	Correctness
7.	, 2003	Punctuation in Compound/Complex Sentences	Correctness
8.	, 2006	Punctuation in Compound/Complex Sentences	Correctness