

**ESTIMATOR KERNEL GAUSSIAN, KERNEL
EPANECHNIKOV DAN KERNEL KUARTIK DALAM
MODEL REGRESI SEMIPARAMETRIK**

SKRIPSI

Untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-1

Program Studi Matematika



diajukan oleh

SRIYANTI MUSTIKA NINGRUM

16610009

Kepada

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UIN SUNAN KALIJAGA
YOGYAKARTA**

2020



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu 'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Sriyanti Mustika Ningrum

NIM : 16610009

Judul Skripsi : Estimator Kernel *Gaussian*, Kernel *Epanechnikov* dan Kernel Kuartik dalam Model Regresi Semiparametrik

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu 'alaikum wr. wb.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Yogyakarta, 11 November 2020
Pembimbing

Moh. Farhan Oudratullah, M.Si

NIP: 19790922 200801 1 011



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-2683/Un.02/DST/PP.00.9/12/2020

Tugas Akhir dengan judul : Estimator Kernel Gaussian, Kernel Epanechnikov, dan Kernel Kuartik dalam Model Regresi Semiparametrik

yang dipersiapkan dan disusun oleh:

Nama : SRIYANTI MUSTIKA NINGRUM
Nomor Induk Mahasiswa : 16610009
Telah diujikan pada : Senin, 30 November 2020
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Mohammad Farhan Qudratullah, S.Si., M.Si
SIGNED

Valid ID: 5fe1760e35924



Penguji I

Dr. Muhammad Wakhid Musthofa, S.Si.,
M.Si.
SIGNED

Valid ID: 5fc9a7917b758



Penguji II

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.
SIGNED

Valid ID: 5fe15f4bc1b5c



Yogyakarta, 30 November 2020

UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Dr. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 5fe180bf48e96

UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Sriyanti Mustika Ningrum

NIM : 16610009

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 11 November 2020



Menyatakan

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Sriyanti Mustika Ningrum

PERSEMBAHAN

Dengan mengucapkan syukur atas berkat dan rahmat Allah SWT.

Skripsi ini saya persembahkan untuk:

Kedua orang tuaku yang telah membesarkan, mendidik, membimbing, dan memberi banyak nasehat, serta iringan doanya yang selalu menyertai setiap langkah saya.

Kakak dan adikku serta semua keluarga besarku yang senantiasa menyayangiku, memberikan motivasi dan dukungannya.

Almamater tercinta Universitas Islam Negeri Sunan Kalijaga Yogyakarta

Bapak Ibu dosen serta Teman-teman yang selalu memberi motivasi dan semangat dalam berkarya.



MOTTO



Sesuatu akan terlihat tidak mungkin hingga semuanya selesai

~Nelson Mandela~



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

Puji syukur kehadiran Allah SWT. atas limpahan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Estimator Kernel *Gaussian*, Kernel *Epanechnikov* dan Kernel Kuartik dalam Model Regresi Semiparametrik” ini dengan baik. Sholawat serta salam semoga tercurahkan kepada Nabi Muhammad Saw. yang telah mengantarkan manusia ke jalan kebenaran.

Keberhasilan penulisan skripsi ini tidak lepas dari bimbingan, arahan dan dukungan dari berbagai pihak, baik berupa pikiran, motivasi tenaga maupun doa. Oleh karena itu, dengan kerendahan hati penulis mengucapkan rasa terima kasih kepada:

1. Prof. Dr. Phil. Al Makin S.Ag.,MA., selaku Rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
2. Dr.Hj. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Muchammad Abrori, S.Si, M.Kom., selaku Ketua Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
4. Mohammad Farhan Qudratullah, M.Si., selaku pembimbing skripsi yang telah memberikan ilmu pengetahuan, pengalaman yang berharga kepada penulis, sehingga ilmu yang telah didapat mempermudah dalam penyusunan skripsi ini.

5. Bapak/Ibu Dosen dan Staff Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta atas ilmu, bimbingan dan pelayanan selama perkuliahan hingga penyusunan skripsi ini selesai.
6. Bapak Mulyono dan Ibu Sri Yati Eko Purwandari, selaku orangtua penulis yang telah memberikan kasih sayang, mendoakan dan memberikan dukungan tiada henti kepada penulis, yang selalu setia menjadi tempat curahan dan merestui setiap langkah penulis.
7. Sri Wahyono dan Sri Ayu Setiawati, selaku kakak dan adik penulis serta semua saudara yang telah memberikan kasih sayang, perhatian, semangat mendoakan dan dukungan tiada henti kepada penulis.
8. Sahabat-sahabat penulis, Laela Rif'atuzzulfa, Emsa Citra Saputri dan Zainul Khozin yang telah banyak membantu, memberikan semangat, memberikan motivasi, dan selalu setia mendengarkan curahan hati penulis.
9. Teman-teman satu bimbingan yang selalu menjadi teman mengeluh bersama, bercanda, berbagi kerumitan pikiran dan teman saling menyemangati serta memberikan motivasi sehingga penulis lebih bersemangat dalam menyelesaikan skripsi ini.
10. Teman-teman prodi Matematika angkatan 2016 yang selalu menemani dan memberikan dukungan dan pelajaran berharga selama ini.
11. Semua pihak yang memberikan dukungan dan doa kepada penulis dalam menyelesaikan skripsi ini yang tidak bisa penulis sebutkan satu per satu.

Semoga Allah SWT menerima amal kebaikan beliau sekalian dan memberikan balasan dan pahala yang berlipat. Penulis menyadari bahwa skripsi

ini masih banyak kekurangan dan jauh dari kata sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun agar penulis dapat membuat karya dengan lebih baik. Penulis berharap, semoga skripsi ini dapat memberikan manfaat kepada para pembaca.

Yogyakarta, 11 November 2020



Penulis



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN.....	iv
PERSEMBAHAN	v
MOTTO	vi
KATA PENGANTAR	vii
DAFTAR ISI	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN.....	xvi
DAFTAR SIMBOL.....	xvii
INTISARI	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	4
1.6 Tinjauan Pustaka	5
1.7 Sistematika Penulisan	8
BAB II LANDASAN TEORI.....	10
2.1 Data	10
2.2 Skala Pengukuran.....	11
2.3 Matriks	13
2.3.1 Jenis Matriks.....	14
2.3.2 Operasi Matriks	16
2.3.3 Transpose Matriks	19

2.3.4 Matriks Simetris	19
2.3.5 Determinan Matriks.....	20
2.3.6 Matriks Singular	20
2.3.7 Invers Matriks.....	20
2.3.8 Aturan Cramer	21
2.3.9 <i>Trace</i> Matriks	22
2.4 Analisis Korelasi	23
2.4.1 Korelasi <i>Product Moment</i> Pearson.....	24
2.4.2 Korelasi Spearman	24
2.4.3 Korelasi Kendall Tau	26
2.5 Regresi Parametrik.....	27
2.6 Analisis Regresi Linear Sederhana	27
2.6.1 Estimasi Regresi Linear Sederhana dengan Metode Kuadrat Terkecil (<i>Ordinary Least Square (OLS)</i>).....	29
2.6.2 Inferensi Analisis Regresi Sederhana.....	33
2.6.3 Asumsi Model Regresi Linear	37
2.6.4 Sifat-sifat Estimator Regresi Linear Sederhana	38
2.7 Analisis Regresi Linear Berganda.....	44
2.8 Uji Asumsi Klasik	45
2.8.1 Normalitas.....	45
2.8.2 Multikolinearitas	47
2.8.3 Heterokedastisitas.....	47
2.8.4 Autokorelasi.....	49
2.9 Regresi Nonparametrik	52
2.10 Regresi Semiparametrik.....	52
2.11 Regresi Kernel	53
2.12 Estimator Densitas Kernel	54
2.13 <i>Mean Square Error (MSE)</i>	57
2.14 Pemilihan <i>Bandwidth</i> Optimal	59
2.15 Deret Taylor.....	60
BAB III METODOLOGI PENELITIAN	63

3.1 Metode Penelitian.....	63
3.2 Metode Pengumpulan Data	63
3.3 Variabel Penelitian	64
3.4 Pengolahan Data.....	64
3.5 <i>Flowchart</i>	66
BAB IV PEMBAHASAN	67
4.1 Regresi Semiparametrik Kernel	67
4.1.1 Regresi Semiparametrik Kernel <i>Gaussian</i>	68
4.1.2. Regresi Semiparametrik Kernel <i>Epanechnikov</i>	69
4.1.3 Regresi Semiparametrik Kernel Kuartik.....	70
4.2 Estimator Priestley-Chao	71
4.2.1 Estimasi Kernel <i>Gaussian</i>	72
4.2.2 Estimasi Kernel <i>Epanechnikov</i>	73
4.2.3 Estimasi Kernel Kuartik.....	74
4.3 Estimasi Bias dan Variansi	74
4.3.1 Estimasi Bias dan Variansi pada Fungsi Kernel <i>Gaussian</i>	78
4.3.2 Estimasi Bias dan Variansi pada Fungsi Kernel <i>Epanechnikov</i>	83
4.4.3 Estimasi Bias dan Variansi pada Fungsi Kernel Kuartik.....	88
4.4 <i>Mean Square Error</i> (MSE).....	93
4.4.1 MSE pada Fungsi Kernel <i>Gaussian</i>	94
4.4.2 MSE pada Fungsi Kernel <i>Epanechnikov</i>	95
4.4.3 MSE pada Fungsi Kernel Kuartik	95
4.5 <i>Mean Integrated Square Error</i> (MISE)	96
4.6 Pemilihan <i>Bandwidth</i> Optimum.....	97
4.6.1 Pemilihan <i>Bandwidth</i> Optimum pada Fungsi Kernel <i>Gaussian</i>	99
4.6.2 Pemilihan <i>Bandwidth</i> Optimum pada Fungsi Kernel <i>Epanechnikov</i> ..	100
4.6.3 Pemilihan <i>Bandwidth</i> Optimum pada Fungsi Kernel Kuartik.....	100
BAB V SIMULASI.....	101
5.1 Deskripsi Data	101
5.2 Hubungan Antar Variabel.....	104
5.3 Pemilihan <i>Bandwidth</i>	106

5.4 Estimator Priestley-Chao	108
5.4.1 Estimator Priestley-Chao dengan Fungsi Kernel <i>Gaussian</i>	109
5.4.2 Estimator Priestley-Chao dengan Fungsi Kernel <i>Epanechnikov</i>	112
5.4.3 Estimator Priestley-Chao dengan Fungsi Kernel Kuartik.....	115
5.5 Perbandingan Nilai <i>Mean Square Error</i> (MSE)	119
BAB VI PENUTUP	121
6.1 Kesimpulan	121
6.2 Saran.....	123
DAFTAR PUSTAKA	125
LAMPIRAN	127
CURRICULUM VITAE.....	143



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
 YOGYAKARTA

DAFTAR TABEL

Tabel 1. 1 Tinjauan Pustaka.....	7
Tabel 2. 1 Kategori Keeratan Hubungan	23
Tabel 2. 2 Jenis-jenis Fungsi Kernel	57
Tabel 4. 1 Nilai $\mu_2(K)$ dan $P(K)$ Fungsi Kernel	99
Tabel 5. 1 Nilai <i>Bandwidth</i> Fungsi Kernel	108
Tabel 5. 4 Nilai <i>Mean Square Error</i> (MSE).....	119



DAFTAR GAMBAR

Gambar 3. 1 <i>Flowchart</i>	66
Gambar 5. 1 Diagram Batang Variabel Dependen.....	102
Gambar 5. 2 Diagram Batang Variabel Independen X_1	103
Gambar 5. 3 Diagram Batang Variabel Independen X_2	104
Gambar 5. 4 <i>Scatterplot</i> Hubungan Variabel Dependen dengan Independen X_1	105
Gambar 5. 5 <i>Scatterplot</i> Hubungan Variabel Dependen dengan Independen X_2	106
Gambar 5. 6 Kurva Estimasi Regresi Kernel <i>Gaussian Bandwidth</i> Optimal	110
Gambar 5. 7 Kurva Estimasi Regresi Kernel <i>Gaussian Bandwidth</i> = 0,1	111
Gambar 5. 8 Kurva Estimasi Regresi Kernel <i>Gaussian Bandwidth</i> = 1	112
Gambar 5. 9 Kurva Estimasi Regresi Kernel <i>Epanechnikov Bandwidth</i> Optimal	113
Gambar 5. 10 Kurva Estimasi Regresi Kernel <i>Epanechnikov Bandwidth</i> = 0,1.	114
Gambar 5. 11 Kurva Estimasi Regresi Kernel <i>Epanechnikov Bandwidth</i> = 1	115
Gambar 5. 12 Kurva Estimasi Regresi Kernel <i>Kuartik Bandwidth</i> Optimal	116
Gambar 5. 13 Kurva Estimasi Regresi Kernel <i>Kuartik Bandwidth</i> = 0,1	117
Gambar 5. 14 Kurva Estimasi Regresi Kernel <i>Kuartik Bandwidth</i> = 1	118

DAFTAR LAMPIRAN

Lampiran 1 <i>R-Script</i> Diagram Batang dan <i>Scatterplot</i>	128
Lampiran 2 <i>R-Script</i> Pemilihan <i>Bandwidth</i> Optimal	129
Lampiran 3 Output R-Studio Pemilihan <i>Bandwidth</i> Optimal	130
Lampiran 4 <i>Source Code</i> Matlab Estimasi Priestley-Chao	131
Lampiran 5 <i>Source Code</i> Matlab Kurva Data Hasil Estimasi	140



DAFTAR SIMBOL

$m(t)$ = fungsi regresi nonparametrik

Y = variabel dependen

X = variabel independen

t = variabel bebas nonparametrik

β = koefisien variabel independen

ε = *error*

n = banyaknya pengamatan

$E(Y)$ = ekspektasi variabel Y

$\text{var}(Y)$ = variansi variabel Y

$f_h(x)$ = fungsi densitas peluang

$K(u)$ = fungsi kernel

h = *bandwidth*

I = fungsi indikator

v = kernel order ke- v

MSE = *Mean Square Error*

$MISE$ = *Mean Integrated Square Error*

**Estimator Kernel Gaussian, Kernel Epanechnikov dan Kernel Kuartik
dalam Model Regresi Semiparametrik**

Oleh : Sriyanti Mustika Ningrum

INTISARI

Analisis regresi merupakan salah satu metode dalam statistika untuk mengetahui hubungan antara variabel independen dengan variabel dependen. Pada analisis regresi terdapat tiga macam pendekatan, yaitu parametrik, nonparametrik, dan semiparametrik. Pendekatan parametrik merupakan pendekatan dengan variabel data yang diketahui bentuk polanya. Pendekatan nonparametrik merupakan pendekatan dengan variabel data dengan kurva *smooth* yang tidak diketahui polanya, sehingga data tersebut akan mencari bentuk polanya sendiri. Sedangkan pendekatan semiparametrik merupakan gabungan dari pendekatan parametrik dan pendekatan nonparametrik. Estimasi kurva sangat bergantung pada perilaku data, sehingga diperlukan teknik *smoothing*. Salah satu teknik *smoothing* yang dapat digunakan dalam penelitian adalah kernel.

Pada penelitian ini, penulis membahas salah satu estimator kernel yaitu estimator Priestley-Chao yang akan mengestimasi tiga fungsi kernel yaitu fungsi kernel *gaussian*, fungsi kernel *epanechnikov* dan fungsi kernel kuartik. Selanjutnya ketiga fungsi kernel yang diestimasi tersebut digunakan untuk menganalisis data simulasi yang dibangkitkan dari *software* R dengan $x_1 \sim N(0,4)$, $x_2 \sim U(-1,1)$ dengan $\varepsilon \sim N(0,1)$.

Hasil analisis yang diperoleh menunjukkan bahwa nilai MSE untuk fungsi kernel *gaussian* sebesar 179,1222 untuk fungsi kernel *epanechnikov* sebesar 74,5913 dan fungsi kernel kuartik sebesar 76,6575. Pada simulasi ini dapat disimpulkan bahwa pada data simulasi yang dianalisis penulis menunjukkan bahwa estimasi regresi kernel *epanechnikov* lebih baik daripada estimasi regresi kernel *gaussian* dan estimasi regresi kernel kuartik.

Kata kunci: Kernel *Epanechnikov*, Kernel *Gaussian*, Kernel Kuartik, Priestley-Chao, Regresi Semiparametrik

BAB I

PENDAHULUAN

1.1 Latar Belakang

Statistika merupakan suatu disiplin ilmu yang mempelajari sekumpulan konsep dan metode pengumpulan, penyajian, analisis, dan interpretasi data sampai pada pengambilan keputusan (Qudratullah, 2013). Statistika menurut tingkat atau tahapan kegiatannya dibagi menjadi dua kelompok, yaitu statistika deskriptif dan statistika inferensia (Khasanah, 2016). Statistika deskriptif adalah statistika yang mengumpulkan, menyajikan, dan mengolah data, misal membuat tabel, grafik, dan perhitungan yang menentukan nilai statistika. Sedangkan statistika inferensia adalah statistika yang berhubungan dengan pengambilan keputusan.

Statistika inferensia berhubungan dengan penaksiran tentang karakteristik populasi, penentuan ada tidaknya hubungan karakteristik dalam populasi, hingga pembuatan prediksi. Salah satu bagian dari statistika inferensia adalah analisis regresi. Analisis regresi merupakan sebuah kajian terhadap hubungan suatu variabel yang disebut sebagai variabel yang diterangkan (variabel dependen) dengan satu atau lebih variabel menerangkan (variabel independen). Analisis regresi umumnya digunakan untuk memodelkan pola hubungan antara satu variabel atau lebih (Qudratullah, 2013).

Tujuan utama dalam analisis regresi adalah untuk mengestimasi kurva regresi. Untuk mengestimasi kurva regresi terdapat tiga bentuk pendekatan, yaitu pendekatan regresi parametrik, pendekatan regresi nonparametrik, dan pendekatan

regresi semiparametrik. Pendekatan regresi parametrik digunakan jika bentuk kurva regresinya diasumsikan diketahui seperti linear, kuadratik, kubik, eksponensial dan lain sebagainya (Gujarati, 2004). Namun dalam kenyataannya tidak semua data memiliki pola-pola tertentu. Jika hubungan variabel respon dan variabel prediktor tidak diketahui bentuk polanya, maka untuk memodelkan hubungan variabel tersebut dapat menggunakan pendekatan regresi nonparametrik. Sedangkan, pendekatan regresi semiparametrik merupakan gabungan dari komponen parametrik dan nonparametrik. Regresi semiparametrik digunakan karena adanya kasus-kasus pemodelan dimana hubungan antar variabelnya sebagian memiliki pola tertentu dan sebagian lainnya tidak diketahui bentuk polanya.

Pada metode analisis regresi semiparametrik terdapat beberapa teknik *smoothing* antara lain histogram, penduga kernel, deret orthogonal, penduga spline, k-NN, deret fourier, dan wavelet. Dalam penduga kernel terdapat beberapa fungsi antara lain kernel *Uniform*, *Triangle*, *Epanechnikov*, *Gaussian*, *Kuartik*, *Cosinus*. Penduga kernel memiliki sifat yang fleksibel dan secara matematik mudah diselesaikan. Pada penduga kernel yang terpenting adalah pemilihan parameter pemulus (*bandwidth*) yang optimal untuk mendapatkan kurva regresi yang optimal. Salah satu metode untuk mencari nilai *bandwidth* yang optimal adalah dengan menggunakan kriteria *Rule of Thumb*.

Menurut Halim dan Bisono (2006) estimator kernel dibagi menjadi tiga macam, yaitu: estimator Nadaraya-Watson, estimator Priestley-Chao dan estimator Gasser-Muller.

Pada penelitian ini akan digunakan estimator kernel Priestley-Chao yang akan mengestimasi fungsi kernel. Estimator kernel Priestley-Chao tersebut akan diterapkan pada beberapa fungsi kernel yaitu fungsi kernel *Gaussian*, fungsi kernel *Epanechnikov* dan fungsi kernel Kuartik. Penelitian ini pada akhirnya akan menunjukkan perbedaan hasil yang diperoleh diantara ketiga fungsi kernel yang dipilih penulis.

Pada aplikasinya, penulis mengambil masalah yang dapat diselesaikan dengan analisis regresi semiparametrik kernel dengan menggunakan estimator Priestley-Chao, yaitu dengan membangkitkan data simulasi dari *software R*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah di atas, berikut beberapa rumusan masalah yang akan dikaji dalam penelitian ini:

1. Bagaimana langkah-langkah analisis regresi semiparametrik kernel dengan menggunakan estimator Priestley-Chao?
2. Bagaimana perbandingan hasil analisis antara regresi kernel *Gaussian* Priestley-Chao, *Epanechnikov* Priestley-Chao dan regresi kernel Kuartik Priestley-Chao yang di aplikasikan terhadap data simulasi *software R*?

1.3 Batasan Masalah

Untuk menghindari terdapatnya ruang lingkup masalah yang terlalu luas dalam penelitian ini, maka penulis memiliki batasan masalah sebagai berikut:

1. Analisis yang digunakan pada penelitian ini adalah analisis regresi semiparametrik kernel *Gaussian*, *Epanechnikov* dan Kuartik dengan estimator Priestley-Chao.

2. Kriteria pemilihan model terbaik adalah dengan menggunakan kriteria nilai MSE (*Mean Square Error*).
3. Untuk menunjang kegiatan analisis data, penulis menggunakan bantuan *software* R Studio.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka diketahui tujuan dari penelitian ini adalah:

1. Mengetahui langkah-langkah analisis regresi semiparametrik kernel menggunakan estimator Priestley-Chao.
2. Mengetahui perbandingan hasil analisis antara regresi kernel *Gaussian* Priestley-Chao, *Epanechnikov* Priestley-Chao dan regresi kernel Kuartik Priestley-Chao yang di aplikasikan terhadap data simulasi *software* R.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Bagi Penulis
 - a. Sebagai salah satu syarat untuk kelulusan dan mendapatkan gelar Strata 1 (S1).
 - b. Menambah pengetahuan tentang pengaplikasian matematika khususnya dalam bidang statistika.
 - c. Mengetahui dan memahami analisis regresi semiparametrik kernel dengan estimator Priestley-Chao.

2. Bagi Program Studi Matematika

Menjadi salah satu acuan atau tambahan referensi dengan harapan dapat dimanfaatkan, dipelajari dan dipahami dengan baik sebagai salah satu alat bantu untuk penelitian selanjutnya.

3. Bagi Pembaca

- a. Memberikan pengetahuan mengenai langkah-langkah serta hasil estimasi fungsi kernel menggunakan estimator Priestley-Chao.
- b. Sebagai sumbangan referensi yang kemudian dapat digunakan sebagai salah satu alat bantu untuk penelitian selanjutnya.

1.6 Tinjauan Pustaka

Pada penelitian ini, penulis menggunakan beberapa tinjauan pustaka seperti buku, jurnal matematika dan skripsi atau penelitian-penelitian terdahulu yang berkaitan dengan penelitian ini. Diantara penelitian yang berkaitan dalam penelitian ini adalah penelitian Adella Fitria Marlin (2016) membahas tentang *Penduga Kurva Regresi Nonparametrik Linear dan Nonlinear dengan Metode Priestley-Chao, Nadaraya-Watson dan Metode Fourier*, penelitian ini bertujuan untuk menduga kurva regresi menggunakan metode Priestley-Chao, Nadaraya-Watson dan metode Fourier pada fungsi regresi linear dan nonlinear berdasarkan nilai kuadrat tengah galat dan *bandwidth* optimal. Sehingga hasil penelitian ini menunjukkan bahwa metode Priestley-Chao, metode Nadaraya-Watson dan metode Fourier dapat digunakan dalam mengestimasi fungsi regresi linear.

Sedangkan pada fungsi regresi nonlinear, metode Fourier lebih baik daripada metode Priestley-Chao dan metode Nadaraya-Watson (Marlin A. F., 2016).

Penelitian Deden Aditya Nanda (2016) membahas tentang *Analisis Pengaruh Jumlah Uang Beredar dan Nilai Tukar Rupiah terhadap Indeks Harga Saham Gabungan Menggunakan Pemodelan Regresi Semiparametrik Kernel*, dalam penelitian ini dijelaskan mengenai regresi semiparametrik dengan menggunakan pendekatan fungsi kernel *Gaussian* dan *Triangle* yang di estimasi menggunakan estimator Nadaraya-Watson. Berdasarkan hasil analisis yang telah dilakukan, model regresi semiparametrik kernel yang terbaik menggunakan fungsi kernel *Gaussian* dengan *bandwidth* sebesar 47,94 dan nilai $GCV = 34675,27047$. Hasil evaluasi ketepatan model untuk nilai *Mean Absolute Percentage Error* (MAPE) data *out sample* sebesar 4,036% yang menunjukkan bahwa model yang dihasilkan adalah sangat akurat (Nanda, 2016).

Penelitian Riswanti Oktavian (2019) membahas tentang *Estimasi Model Regresi Semiparametrik dengan Penduga Nadaraya Watson Kernel Uniform*, dalam penelitian ini dijelaskan mengenai regresi semiparametrik fungsi kernel *Uniform* yang di estimasi menggunakan estimator Nadaraya-Watson, yang diaplikasikan pada bidang kesehatan, yaitu mengenai keterhubungan apakah suhu tubuh, kadar hemoglobin dan kadar hematokrit dapat memengaruhi lama kesembuhan pasien DBD. Berdasarkan hasil analisis yang telah dilakukan, hasil dari penelitian ini menunjukkan bahwa model regresi semiparametrik dengan kernel *uniform* dapat mengestimasi data asli dan data simulasi dengan sangat baik (Oktavian, 2019).

Adapun detail dari ketiga penelitian di atas dapat di lihat pada tabel.1.1 berikut ini:

Tabel 1. 1 Tinjauan Pustaka

No	Nama Peneliti	Model Regresi	Metode Estimasi	Studi Kasus
1	Adella Fitria Marlin (2016)	Linear dan Nonlinear	Priestley-Chao, Nadaraya-Watson, dan Fourier	Data bangkitan software Matlab R2013b yang berdistribusi Uniform
2	Deden Aditya Nanda (2016)	Kernel <i>Gaussian</i> dan Kernel <i>Triangle</i>	Nadaraya-Watson	Pengaruh Jumlah Uang Beredar dan Nilai Tukar Rupiah terhadap Indeks Harga Saham Gabungan
3	Riswanti Oktavian (2019)	Kernel <i>Uniform</i>	Nadaraya-Watson	Rekam Medis Pasien DBD bulan Januari-Agustus 2018
4	Sriyanti Mustika Ningrum (2020)	Kernel <i>Gaussian</i> , Kernel <i>Epanechnikov</i> dan Kernel Kuartik	Priestley-Chao	Data Simulasi Bangkitan Software R

Perbedaan penelitian penulis dengan tinjauan pustaka yang pertama adalah jika pada penelitian Adella Fitri Marlin menggunakan model regresi linear dan nonlinear dengan metode estimasi Priestley-Chao, Nadaraya-Watson, dan Fourier. Sementara penulis mengangkat penelitian ini menggunakan model regresi *gaussian*, *epanechnikov* dan kuartik dengan metode estimasi Priestley-Chao.

Perbedaan penelitian penulis dengan tinjauan pustaka yang kedua adalah jika pada penelitian Deden Aditya Nanda menggunakan model regresi kernel

Gaussian dan kernel *Triangle* dengan metode estimasi Nadaraya-Watson. Sementara penulis mengangkat penelitian ini menggunakan model regresi *gaussian*, *epanechnikov* dan kuartik dengan metode estimasi Priestley-Chao.

Perbedaan penelitian penulis dengan tinjauan pustaka yang ketiga adalah jika pada penelitian Riswanti Oktavian menggunakan model regresi kernel *uniform* dengan metode estimasi Nadaraya-Watson. Sementara penulis mengangkat penelitian ini menggunakan model regresi *gaussian*, *epanechnikov* dan kuartik dengan metode estimasi Priestley-Chao.

1.7 Sistematika Penulisan

Sistematika penulisan ini bertujuan agar memudahkan pembaca untuk memahami penulisan dalam penelitian ini secara sederhana, runtut, jelas dan dengan harapan mudah untuk dipahami. Adapun sistematika penulisan dalam penelitian ini adalah sebagai berikut:

1. BAB I : PENDAHULUAN

Pada bab ini, akan membahas tentang hal-hal yang melatar belakangi penulis dalam menyusun penelitian ini, pembatasan masalah yang diangkat, rumusan masalah, tujuan penelitian, manfaat penelitian, dan tinjauan pustaka yang merupakan acuan dalam penyusunan penelitian ini serta sistematika penulisan sebagai gambaran sederhana dari penelitian ini secara garis besar.

2. BAB II : LANDASAN TEORI

Pada bab ini, berisi tentang teori-teori menunjang atau membantu sebagai penguat pembahasan analisis dalam penelitian yang dilakukan

terkait analisis data dengan menggunakan regresi semiparametrik kernel dengan estimator Priestley-Chao.

3. BAB III : METODOLOGI PENELITIAN

Pada bab ini akan dibahas mengenai cara, metode atau langkah-langkah analisis yang digunakan dalam penelitian. Selain itu juga memuat sumber data dan identifikasi variabel yang digunakan.

4. BAB IV : PEMBAHASAN

Pada bab ini akan dibahas mengenai metode analisis yang diteliti, dalam hal ini adalah yang terkait dengan analisis regresi semiparametrik kernel dengan estimator Priestley-Chao.

5. BAB V : SIMULASI

Pada bab ini akan dilakukan analisis terhadap simulasi yang dilakukan dalam penelitian dengan menggunakan metode yang telah dibahas pada bab sebelumnya.

6. BAB VI : PENUTUP

Pada bab ini, membahas tentang kesimpulan-kesimpulan yang diperoleh berdasarkan hasil dan pembahasan pada bab sebelumnya. Selain kesimpulan, pada bab ini juga terdapat saran-saran untuk penelitian selanjutnya yang sekiranya masih relevan dengan penelitian ini.

BAB VI

PENUTUP

Regresi Semiparametrik kernel yang di estimasi menggunakan metode estimasi Priestley-Chao telah dijelaskan dalam pembahasan dan dilakukan simulasi dengan menggunakan data yang dibangkitkan dari *software* R, sehingga diperoleh kesimpulan dan saran sebagai berikut.

6.1 Kesimpulan

Berdasarkan pembahasan pada bab-bab sebelumnya, maka dapat di ambil kesimpulan sebagai berikut.

1. Berikut ini adalah langkah-langkah menganalisis regresi kernel dengan estimator Priestley-Chao:
 - a. Menentukan variabel-variabel yang akan dianalisis menggunakan regresi kernel.
 - b. Menentukan fungsi kernel yang akan diestimasi dengan estimator Priestley-Chao. Terdapat tiga fungsi kernel yang digunakan dalam penelitian ini, yaitu fungsi kernel *gaussian*, fungsi kernel *epanechnikov* dan fungsi kernel kuartik dengan formula sebagai berikut.

- 1) Fungsi Kernel *Gaussian*

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{1}{2}(-u^2)\right)$$

- 2) Fungsi Kernel *Epanechnikov*

$$K(u) = \frac{3}{4}(1-u^2)I(|u| \leq 1)$$

3) Fungsi Kernel Kuartik

$$K(u) = \frac{15}{16}(1-u^2)^2 I(|u| \leq 1)$$

- c. Menentukan nilai *bandwidth* optimum dari masing-masing variabel independen dengan menggunakan kriteria pemilihan nilai *bandwidth* optimum “*Rule of Thumb*”.
- d. Mengestimasi fungsi kernel dengan estimator Priestley-Chao dengan menggunakan rumus $\hat{m}(x) = \frac{1}{h} \sum_{i=1}^n (x - x_{i-1}) y_i K\left(\frac{x - x_i}{h}\right)$, dimana $u = \left(\frac{x - x_i}{h}\right)$.
- e. Menentukan nilai *Mean Square Error* (MSE) pada masing-masing fungsi kernel yang diestimasi dengan menggunakan estimator Priestley-Chao.
- f. Memilih model terbaik dengan menggunakan kriteria nilai MSE terkecil.
2. Berdasarkan analisis fungsi kernel *gaussian*, fungsi kernel *epanechnikov*, dan fungsi kernel kuartik yang diestimasi dengan menggunakan estimator Priestley-Chao yang diaplikasikan terhadap simulasi dengan data yang dibangkitkan dari *software* R, diperoleh nilai *bandwidth* optimal kernel *gaussian* adalah 0,3283762, selanjutnya nilai *bandwidth* optimal kernel *epanechnikov* adalah 0,7268765, sedangkan nilai *bandwidth* optimal kernel kuartik adalah 0,8620809. Kemudian nilai MSE untuk fungsi kernel *gaussian*,

kernel *epanechnikov* dan kernel kuartik dengan *bandwidth* optimum masing-masing berturut-turut adalah 179,12222 , 74,5913 dan 76,6575. Dengan hasil MSE tersebut, diperoleh model terbaik dalam menganalisis simulasi yang diangkat peneliti adalah model regresi kernel *epanechnikov* lebih baik daripada model regresi kernel *gaussian* dan kernel kuartik. Berdasarkan hasil simulasi dengan membangkitkan data simulasi dari *software* R diperoleh hasil bahwa analisis regresi kernel *epanechnikov* merupakan model terbaik daripada model regresi kernel *gaussian* dan kernel kuartik. Model terbaik untuk regresi semiparametrik kernel *epanechnikov* adalah

$$\hat{y}_i = 0,7502 + 2,5014 * x_1 + \left(\frac{1}{0,7268765} \sum_{i=1}^{1000} (x_2 - x_{2i-1}) y_i \left(\frac{3}{4} \left(1 - \left(\frac{x_2 - x_{i2}}{0,7268765} \right)^2 \right) I \right) \right)$$

6.2 Saran

Demi memperluas ilmu pengetahuan tentang regresi semiparametrik, ada beberapa saran yang penulis sampaikan sebagai berikut.

1. Masih banyak pilihan metode dalam regresi semiparametrik yang dapat dilakukan dalam mengestimasi kurva regresi seperti spline, deret fourier, dan lain-lain.
2. Masih banyak pilihan fungsi kernel yang lain seperti fungsi kernel *uniform*, fungsi kernel *cosinus*, fungsi kernel *triangle*, dan lain-lain.
3. Menggunakan estimator fungsi kernel selain estimator Priestley-Chao seperti estimator Nadaraya-Watson, estimator Gasser-Muller, dan lain-lain.

4. Menggunakan kriteria pemilihan bandwidth optimum selain *Rule of Thumb* seperti kriteria *Complete Cross Validation (CCV)*, atau *Generalized Cross Validation (GCV)*.



DAFTAR PUSTAKA

- Anton, H., & Rorres, C. (2004). *Aljabar Linear Elementer*. Jakarta: Erlangga.
- Awat, N. J. (1995). *Metode Statistik dan Ekonometri*. Yogyakarta: Liberty.
- Baisuni, M. H. (1986). *Kalkulus*. Jakarta: Universitas Indonesia Press.
- Budiantara, I. N. (2009). *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: Institut Teknologi Sepuluh November.
- Djarwanto. (2001). *Statistik Nonparametrik Edisi Ketiga*. Yogyakarta: BPFE.
- Efendi, T. (2013). *Pemodelan Persamaan Regresi Spline Kuadratik dengan Menentukan Titik-titik Knot Optimal*. Skripsi.
- Guidoum, A. C. (2013). *Kernel Estimator and Bandwidth Selection for Density and its Derivatives*. The Kedd Packages.
- Gujarati, D. N. (2004). *Basic Econometrics*. New York: McGraw-Hill.
- Halim, S., & Bisono, I. (2006). Fungsi-fungsi Kernel pada Metode Regresi Nonparametrik dan Aplikasinya pada Priest River Experimental Forest's Data. *Jurnal Teknik Industri*, 73-81.
- Hardle, W. (1990). *Applied Nonparametric Regression*. New York: Cambridge University Press.
- Harini, S., & Turmudi. (2008). *Metode Statistika*. Malang: UIN Malang Press.
- Hasan, M. I. (2002). *Pokok-Pokok Materi Metodologi Penelitian dan Aplikasinya*. Jakarta : Ghalia Indonesia.
- Johannes, & Budiono. (1994). *Pengantar Matematika untuk Ekonomi*. Jakarta: LP3ES.
- Khasanah, U. (2016). *Statistika Non Parametrik*. Yogyakarta: UAD Press.
- Komang, G., & Gusti, A. (2012). Estimator Kernel dalam Regresi Nonparametrik. *Jurnal Matematika Volume 2 Nomor 1. Universitas Udayana, Bali*.

- Marlin, A. F. (2016). *Penduga Kurva Regresi Nonparametrik Linear dan Nonlinear dengan Metode Prestley-Chao, Nadaraya Watson dan Metode Fourier*. Lampung: Skripsi Jurusan Matematika FMIPA Universitas Lampung.
- Nanda, D. A. (2016). *Analisis Pengaruh Jumlah Uang Beredar dan Nilai Tukar Rupiah terhadap Indeks Harga Saham Gabungan Menggunakan Pemodelan Regresi Semiparametrik Kernel*. *Jurnal Gaussian*, 5(3).373-382.
- Oktavian, R. (2019). *Estimasi Model Regresi Semiparametrik dengan Penduga Nadaraya-Watson Kernel Uniform*. Skripsi Jurusan Matematika FMIPA Universitas Lampung.
- Quadratullah, M. F. (2013). *Analisis Regresi Terapan*. Yogyakarta: C.V Andi Offset.
- Quadratullah, M. F. (2017). *Statistika Nonparametrik Terapan: Teori, Contoh Kasus & Aplikasi dengan IBM SPSS*. Yogyakarta: Andi.
- Ruppert, D., Wand, M. P., & Carroll, J. R. (2003). *Semiparametric Regression*. New York: Cambridge University Press.
- Supangat, A. (2007). *Statistika dalam Kajian Deskriptif, Inferensi dan Nonparametrik*. Jakarta: Predana Media.
- Ulinuha, I. (2019). *Estimator Nadaraya-Watson dengan Fungsi Kernel Epanechnikov dan Fungsi Kernel Kuartik*. Skripsi Jurusan Matematika Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
- Wand, M. P., & Jones, M. C. (1995). *Kernel Smoothing*. New York: Chapman and Hall.
- Widarjono, A. (2010). *Analisis Statistika Multivariat Terapan*. Yogyakarta: Unit Penerbit dan Percetakan.
- Yamin, S., Rachmach, L. A., & Kurniawan, H. (2011). *Regresi dan Korelasi dalam Genggaman Ada: Aplikasi dengan Software SPSS, EViews, MINITAB, dan STATGRAPHICS*. Jakarta: Salemba Empat.
- Yatchew, A. (2003). *Semiparametric Regression for the Applied Econometrician*. New York: Cambridge University Press.