

**PEMODELAN INDEKS LQ45 DENGAN FUNGSI  
KERNEL *TRIWEIGHT* DAN *EPANECHNIKOV*  
MENGUNAKAN *BANDWIDTH GENERALIZED*  
*CROSS VALIDATION***

**SKRIPSI**



**ZULFA PUTRI ASMAWI**

**H051171305**

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS HASANUDDIN  
MAKASSAR  
DESEMBER 2023**

**PEMODELAN INDEKS LQ45 DENGAN FUNGSI  
KERNEL *TRIWEIGHT* DAN *EPANECHNIKOV*  
MENGUNAKAN *BANDWIDTH GENERALIZED*  
*CROSS VALIDATION***

**SKRIPSI**

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada  
Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu  
Pengetahuan Alam Universitas Hasanuddin

**ZULFA PUTRI ASMAWI**

**H051171305**

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS HASANUDDIN**

**MAKASSAR**

**DESEMBER 2023**

## LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

**Pemodelan Indeks LQ45 dengan Fungsi Kernel *Triweight* dan *Epanechnikov* menggunakan *Bandwidth Generalized Cross Validation***

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 11 Desember 2023



Zulfa Putri Asmawi

NIM H051171305

PEMODELAN INDEKS LQ45 DENGAN FUNGSI KERNEL  
*TRIWEIGHT* DAN *EPANECHNIKOV* MENGGUNAKAN  
*BANDWIDTH GENERALIZED CROSS VALIDATION*

Disetujui Oleh:

Pembimbing Utama

Sitti Sahriman, S.Si., M.Si.  
NIP. 19881018 201504 2 002

Pembimbing Pendamping

Dr. Nirwan, M.Si.  
NIP. 19630306 198702 1 002

Ketua Program Studi



Fitria Istiqomiyati, S.Si., M.Si.  
NIP. 19770808 200501 2 002

Pada 11 Desember 2023

## HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Zulfa Putri Asmawi  
NIM : H051171305  
Program Studi : Statistika  
Judul Skripsi : *Pemodelan Indeks LQ45 dengan Fungsi Kernel Triweight dan Epanechnikov menggunakan Bandwidth Generalized Cross Validation*

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

### DEWAN PENGUJI

1. Ketua : Sitti Sahriman, S.Si., M.Si. (.....)
2. Sekretaris : Dr. Nirwan, M.Si. (.....)
3. Anggota : Drs. Raupong, M.Si. (.....)
4. Anggota : Siswanto, S.Si., M.Si. (.....)

Ditetapkan di : Makassar

Tanggal : 11 Desember 2023

## KATA PENGANTAR

AlhamdulillahRobbil'alamin, Segala Puji senantiasa disenandungkan untuk menyampaikan rasa syukur penulis kepada **Allah Subhanahu Wata'ala** atas segala limpahan rahmat, nikmat, dan hidayah yang diberikan kepada penulis sehingga dapat menyelesaikan penulisan skripsi dengan judul "**Pemodelan Indeks LQ45 dengan Fungsi Kernel *Triweight* dan *Epanechnikov* menggunakan *Bandwidth Generalized Cross Validation***" sebagai salah satu syarat untuk mendapat gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin Makassar.

Penulis menyadari dalam penulisan skripsi ini tidak lepas dari hambatan dan masalah namun dapat terselesaikan berkat dukungan dan bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan rasa terima kasih yang setulus-tulusnya serta penghargaan yang setinggi-tingginya untuk orang tua penulis, Ayahanda **Asmawi, S.E** dan Ibunda tercinta **Musdalipa**, yang telah membesarkan dan mendidik dengan penuh kesabaran, memberikan cinta dan limpahan kasih sayang, dukungan dan doa yang tulus tanpa henti kepada penulis. Rasa terima kasih juga kepada adik tersayang **Zulfahmi Asmawi, Nurul Isma Asmawi, Mutia Asmawi, Salwa Saswana, Ahmad Fachri Rahimahullah**, dan **Ahmad Mufli**. Tak lupa pula terima kasih untuk **Keluarga Besar** atas doa, dukungan, semangat, dan bantuannya kepada penulis.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin, M.Si.** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Anna Islamiyati, S.Si., M.Si.** selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf Departemen Statistika yang telah membekali ilmu dan kemudahan-kemudahan kepada Penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.

4. **Bapak Siswanto, S.Si., M.Si.** selaku Penasehat Akademik atas saran, nasehat, motivasi, dan dukungan yang telah diberikan kepada penulis selama menjadi mahasiswa dan meluangkan waktu untuk membimbing dan memberikan masukan dalam penulisan skripsi ini.
5. **Ibu Sitti Sahrinan, S.Si., M.Si.** selaku Pembimbing Utama dan **Bapak Dr. Nirwan, M.Si.** selaku Pembimbing Pertama yang telah ikhlas meluangkan waktu dan pemikirannya untuk memberikan arahan, pengetahuan, motivasi dan bimbingan ditengah kesibukan beliau serta menjadi tempat berkeluh kesah untuk penulis.
6. **Bapak Drs. Raupong, M.Si.** dan **Bapak Siswanto, S.Si., M.Si.** selaku Tim Penguji atas saran dan kritikan yang membangun dalam penyempurnaan penyusunan skripsi ini serta waktu yang telah diberikan kepada Penulis.
7. Sahabat seperjuangan dalam penyusunan skripsi, **Nurhidayatullah, Zakiah Fitri** dan **Sriulan S.S. Abdullah** yang selalu setia membersamai penulis, memberikan dukungan satu sama lain dan menjadi tempat bertukar cerita. Terima kasih juga kepada **Kak Muhammad Fadil** untuk segala bantuan dan motivasi yang tidak ada putusnya.
8. Teman-teman **Statistika 2017**, terima kasih atas kebersamaan, suka, dan duka selama menjalani pendidikan di Departemen Statistika. Penulis senang mengenal kalian semua. Terkhusus untuk **Aqilah, Haura, Ata, Zulfa, Ukhty, Ica, Mirna, Putri, Wiwi, Iklil, Aii** dan **Rahman** terima kasih atas segala bantuan dan motivasinya untuk penulis dalam menyelesaikan skripsi ini.
9. Keluarga besar **DISKRIT 2017**, terima kasih telah memberikan pelajaran yang berharga dan arti kebersamaan selama ini kepada penulis. Pengalaman yang berharga telah penulis dapatkan dari teman-teman selama berproses.
10. **Keluarga Mahasiswa FMIPA Unhas** terkhusus anggota keluarga **Himatika FMIPA Unhas** dan **Himastat FMIPA Unhas**, terima kasih atas ilmu yang mungkin tidak bisa didapatkan di proses perkuliahan dan telah menjadi keluarga selama penulis kuliah di Universitas Hasanuddin.
11. Semua pihak yang telah banyak berpartisipasi, baik secara langsung maupun tidak langsung yang tak sempat penulis sebutkan satu per satu. Terima kasih untuk segala bantuan dan dukungannya.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini. Untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga skripsi ini dapat memberikan manfaat bagi pembaca.

*Wassalamu'alaikum Warahmatullahi Wabarakatuh*

Makassar, 11 Desember 2023



Zulfa Putri Asmawi

## PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

---

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama : Zulfa Putri Asmawi  
NIM : H051171305  
Program Studi : Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam  
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-eksklusif** (*Non-exclusive Royalty- Free Right*) atas tugas akhir saya yang berjudul:

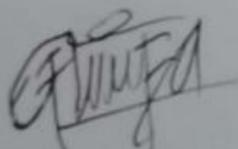
**“Pemodelan Indeks LQ45 dengan Fungsi Kernel *Triweight* dan *Epanechnikov* menggunakan *Bandwidth Generalized Cross Validation*”**

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal 11 Desember 2023.

Yang menyatakan



(Zulfa Putri Asmawi)

## ABSTRAK

Analisis regresi nonparametrik merupakan metode pendugaan yang digunakan jika tidak ada informasi sebelumnya tentang bentuk kurva regresi. Salah satu teknik untuk mengestimasi kurva regresi nonparametrik yaitu kernel. Kelebihan dari regresi kernel memiliki bentuk yang fleksibel dan relatif mudah perhitungan matematikanya. Regresi kernel dengan estimator *Priestley-Chao* adalah salah satu metode untuk mengestimasi data yang memiliki fluktuasi serta sulit diestimasi bentuknya. Estimasi dengan pendekatan kernel dipengaruhi oleh parameter fungsi kernel *triweight* dan *epanechnikov* serta pemilihan *bandwidth*. Tujuan penelitian ini adalah untuk mendapatkan model terbaik data indeks saham LQ45 menggunakan fungsi kernel *triweight* dan *epanechnikov* dengan *bandwidth* optimum berdasarkan nilai *generalized cross validation*. Variabel pada penelitian ini terdiri dari variabel repon yaitu data indeks saham LQ45 dan variabel prediktor berupa waktu pengamatan pergerakan indeks saham periode 1 Januari 2022 sampai dengan 31 Desember 2022. Hasil penelitian menunjukkan bahwa pemilihan *bandwidth* dari 0.01 sampai 10 diperoleh *bandwidth* optimum fungsi kernel *triweight* dan *epanechnikov* adalah  $h = 1.11$  dan  $h = 0.75$  dengan nilai *generalized cross validation* sebesar 2.637 dan  $4.59 \times 10^{-27}$ . *Mean square error* fungsi kernel *triweight* adalah 2.614 dan *epanechnikov* adalah 1.758, sehingga fungsi kernel *epanechnikov* merupakan model terbaik.

**Kata kunci:** Indeks Saham LQ45, Fungsi Kernel, *Triweight*, *Epanechnikov*, *Bandwidth*, *Generalized Cross Validation*.

## ABSTRACT

Nonparametric regression analysis is an estimation method used when there is no prior information about the shape of the regression curve. One of the techniques to estimate the nonparametric regression curve is kernel. The advantage of kernel regression is that it has a flexible form and is relatively easy to calculate mathematically. Kernel regression with the Priestley-Chao estimator is one method to estimate data that has fluctuations and is difficult to estimate its shape. Estimation with the kernel approach is influenced by the triweight and epanechnikov kernel function parameters and bandwidth selection. The purpose of this research is to get the best model of LQ45 stock index data using triweight and epanechnikov kernel functions with optimum bandwidth based on generalized cross validation values. The variables in this study consist of response variables, namely LQ45 stock index data and predictor variables in the form of observation time of stock index movements from January 1, 2022 to December 31, 2022. The results showed that the selection of bandwidth from 0.01 to 10 obtained the optimum bandwidth of the triweight kernel function and epanechnikov is  $h = 1.11$  and  $h = 0.75$  with generalized cross validation values of 2.637 and  $4.59 \times 10^{-27}$ . Mean square error of triweight kernel function is 2.614 and epanechnikov is 1.758, so epanechnikov kernel function is the best model.

**Keywords:** LQ45 Stock Index, Kernel Fungsi, *Triweight*, *Epanechnikov*, *Bandwidth*, *Generalized Cross Validation*.

## DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PERNYATAAN KEOTENTIKAN.....	ii
LEMBAR PERSETUJUAN PEMBIMBING.....	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR .....	v
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS .....	viii
ABSTRAK .....	ix
ABSTRACT.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR .....	xiii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN.....	xv
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>4</b>
2.1 Regresi Nonparametrik.....	4
2.2 Regresi Kernel.....	4
2.3 Estimator Kernel.....	4
2.4 Estimator Priestley Chao .....	7
2.5 Pemilihan <i>Bandwidth</i> Optimum .....	8
2.6 <i>Generalized Cross Validation</i> .....	8
2.7 <i>Mean Square Error</i> .....	9
2.8 Saham .....	9
2.9 Pasar Modal .....	9
2.10 Indeks Harga Saham.....	9

<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>11</b>
3.1 Sumber Data .....	11
3.2 Deskripsi Penelitian.....	11
3.3 Metode Analisis.....	11
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>13</b>
4.1 Deskripsi Data .....	13
4.2 Fungsi Kernel <i>Triweight</i> dan <i>Epanechnikov</i> .....	14
4.2.1 <i>Bandwidth</i> 0.01.....	15
4.2.2 <i>Bandwidth</i> 10.....	15
4.3 Pemilihan <i>Bandwidth</i> Optimum pada Fungsi Kernel.....	16
4.4 Penduga Model Regresi Kernel <i>Triweight</i> .....	18
4.5 Penduga Model Regresi Kernel <i>Epanechnikov</i> .....	20
4.6 Perbandingan Nilai <i>Mean Square Error</i> .....	22
<b>BAB V KESIMPULAN DAN PEMBAHASAN .....</b>	<b>23</b>
5.1 Kesimpulan.....	23
5.2 Saran .....	23
<b>DAFTAR PUSTAKA .....</b>	<b>24</b>
<b>LAMPIRAN.....</b>	<b>26</b>

## DAFTAR GAMBAR

<b>Gambar 2.1</b> Fungsi Kernel.....	6
<b>Gambar 2.2</b> Fungsi Kernel .....	6
<b>Gambar 4.1</b> Kurva Harga Saham LQ45 .....	13
<b>Gambar 4.2</b> Kurva Regresi Kernel <i>Triweight</i> dan <i>Epanechnikov</i> dengan <i>Bandwidth</i> 0.01.....	15
<b>Gambar 4.3</b> Kurva Regresi Kernel <i>Triweight</i> dan <i>Epanechnikov</i> dengan <i>Bandwidth</i> 10.....	16
<b>Gambar 4.4</b> Kurva Regresi Kernel <i>Triweight</i> dengan <i>Bandwidth</i> 1.11 .....	18
<b>Gambar 4.5</b> Kurva Error Regresi Kernel <i>Triweight</i> dengan <i>Bandwidth</i> 1.11 .....	18
<b>Gambar 4.6</b> Kurva Regresi Kernel <i>Epanechnikov</i> dengan <i>Bandwidth</i> 0.75 .....	20
<b>Gambar 4.7</b> Kurva Error Regresi Kernel <i>Epanechnikov</i> dengan <i>Bandwidth</i> 0.75 .....	20

## DAFTAR TABEL

<b>Tabel 2.1</b> Fungsi Kernel .....	5
<b>Tabel 4.1</b> Nilai GCV minimum Kernel <i>Triweight</i> dan <i>Epanechnikov</i> .....	17
<b>Tabel 4.2</b> Nilai MSE.....	22

## DAFTAR LAMPIRAN

<b>Lampiran 1</b> Data Harga Saham LQ45 Periode 1 Januari 2022 sampai dengan 31 Desember 2022.....	27
<b>Lampiran 2</b> Pemilihan <i>Bandwidth</i> Optimum Berdasarkan GCV .....	30
<b>Lampiran 3</b> Nilai Hasil Estimasi dan Error Fungsi Kernel <i>Triweight</i> Dengan <i>Bandwidth</i> $h = 0.01$ .....	45
<b>Lampiran 4</b> Nilai Hasil Estimasi dan Error Fungsi Kernel <i>Epanechnikov</i> Dengan <i>Bandwidth</i> $h = 0.01$ .....	49
<b>Lampiran 5</b> Nilai Hasil Estimasi dan Error Fungsi Kernel <i>Triweight</i> Dengan <i>Bandwidth</i> $h = 10$ .....	53
<b>Lampiran 6</b> Nilai Hasil Estimasi dan Error Fungsi Kernel <i>Epanechnikov</i> Dengan <i>Bandwidth</i> $h = 10$ .....	57
<b>Lampiran 7</b> Nilai Hasil Estimasi dan Error Fungsi Kernel <i>Triweight</i> Dengan <i>Bandwidth</i> $h = 1.11$ .....	61
<b>Lampiran 8</b> Nilai Hasil Estimasi dan Error Fungsi Kernel <i>Epanechnikov</i> Dengan <i>Bandwidth</i> $h = 0.75$ .....	65

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Indeks LQ45 merupakan indeks yang terdapat dalam Bursa Efek Indonesia, yang terdiri atas 45 saham likuid atau saham dengan transaksi terbanyak. Indeks LQ45 hadir dalam Bursa Efek Indonesia untuk melengkapi Indeks Harga Saham Gabungan (IHSG) yakni untuk menyediakan sarana yang obyektif dan terpercaya bagi para analisis keuangan, manajer investasi, investor dan pemerhati pasar modal lainnya dalam memonitor pergerakan harga saham yang aktif diperdagangkan (Lexy dkk, 2015). Oleh karena itu, sangat penting untuk mengetahui kenaikan atau penurunan harga saham dengan menggunakan suatu metode analisis.

Data indeks saham LQ45 umumnya tidak memiliki bentuk pola sebaran data tertentu, sehingga bisa menggunakan pendekatan regresi nonparametrik untuk mendekati pola sebaran data aslinya. Salah satu teknik untuk mengestimasi kurva regresi nonparametrik yaitu kernel. Regresi kernel merupakan salah satu analisis nonparametrik dengan metode *smoothing* yang bertujuan untuk membuang variabilitas dari data yang tidak memiliki efek sehingga ciri-ciri dari data akan tampak jelas. Regresi kernel memiliki bentuk yang fleksibel dan perhitungan matematisnya mudah sehingga dapat mencapai tingkat kekonvergenan yang relatif cepat . (Budiantara & Mulianah, 2007).

Regresi kernel dengan estimator *Priestley-Chao* bergantung pada dua parameter yakni fungsi kernel dan pemilihan *bandwidth* (Suparti dkk, 2018). Fungsi kernel yang umum digunakan yaitu kernel *triweight* dan kernel *epanechnikov* (Hardle dkk, 2004). Fungsi kernel *triweight* relatif lebih mudah perhitungannya dan fungsi kernel *epanechnikov* mempunyai laju konvergensi lebih cepat menuju nilai yang diestimasi (Hardle,1990). *Bandwidth* adalah parameter *smoothing* yang berfungsi untuk mengatur kemulusan kurva yang diestimasi (Ningsi, 2019). *Bandwidth* yang relatif kecil menghasilkan estimasi kurva kurang mulus (*under-smoothing*) dan sebaliknya *bandwidth* relatif besar menghasilkan estimasi kurva yang sangat mulus (*over-smoothing*) yang tidak sesuai dengan pola data aslinya. Oleh karena itu, *bandwidth* yang dipilih harus optimal. Pemilihan *bandwidth* yang

optimal dilakukan dengan cara memperkecil tingkat kesalahan. *Generalized Cross Validation* (GCV) adalah salah satu metode untuk mendapatkan *bandwidth* optimal (Silverman, 1986). Untuk mengetahui tingkat kesalahan suatu estimator dapat dilihat dari *Mean Square Error* (MSE), semakin kecil tingkat kesalahan semakin baik estimasinya.

Nur Azizah Komara Rifa (2019) telah melakukan penelitian menggunakan data saham *Jakarta Composite Index* (JCI) dengan metode regresi kernel berdasarkan estimator *Nadaraya-Watson*. Hasil penelitiannya menunjukkan bahwa estimator *Nadaraya Watson* dengan fungsi kernel *triweight* dengan *Mean Square Error* 9030,63 dan *bandwidth* 108,2 adalah model terbaik untuk prediksi. Febriolah Lamusu dkk (2020) telah melakukan penelitian menggunakan estimator *Nadaraya-Watson* dengan pendekatan *Cross Validation* dan *Generalized Cross Validation* yang bertujuan untuk mengestimasi produksi jagung dan memperoleh hasil *Generalized Cross Validation* (GCV) lebih baik daripada *Cross Validation* (CV) untuk mengestimasi data produksi jagung.

Subian Saidi dkk (2021) telah melakukan penelitian menggunakan data Covid-19 Indonesia 2020 menggunakan regresi kernel dengan estimator *Nadaraya-Watson* dan beberapa fungsi kernel dari hasil penelitian menunjukkan bahwa regresi kernel *triweight* dengan *Mean Square Error* 9030,63 dan *bandwidth* 108,2 adalah model terbaik untuk prediksi. Meilani Nurjaina dkk (2022) telah melakukan penelitian mengenai estimator kurva *Priestley-Chao* menggunakan fungsi Kernel *Triangle* untuk data rata-rata bulanan bilangan sunspot matahari dapat disimpulkan bahwa estimator kurva *Priestley-Chao* dengan menggunakan fungsi kernel *Triangle* menghasilkan kurva regresi yang hampir mengikuti pola kurva data asli yang sebenarnya. Berdasarkan uraian di atas maka peneliti mengambil judul **“Pemodelan Indeks Saham LQ45 Dengan Fungsi Kernel *Triweight* Dan *Epanechnikov* Menggunakan *Bandwidth Generalized Cross Validation*”**.

## 1.2 Rumusan Masalah

Berdasarkan pada latar belakang yang telah dikemukakan diatas, maka rumusan masalah yang akan dikaji dalam penelitian ini adalah:yaitu:

1. Bagaimana pemilihan *bandwidth* optimum menggunakan *Generalized Cross Validation*?
2. Bagaimana model terbaik data indeks LQ45 menggunakan fungsi kernel *Triweight* dan kernel *Epanechnikov*?

## 1.3 Batasan Masalah

Berdasarkan permasalahan yang telah diuraikan untuk menghindari terjadinya penyimpangan dalam pembahasan dalam penelitian ini sangat diperlukan batasan masalah yaitu:

1. Penelitian ini menggunakan data harga saham LQ45 periode 1 Januari 2022 sampai dengan 31 Desember 2022.
2. Estimator kernel yang akan digunakan adalah estimator *Priestley Chao*.
3. Pemilihan nilai *bandwidth* dibatasi antara interval 0.01 sampai 10.

## 1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, adapun tujuan yang ingin dicapai pada penelitian ini adalah:

1. Mendapatkan nilai *bandwidth optimum* dengan *Generalized Cross Validation*.
2. Mendapatkan model terbaik data indeks LQ45 menggunakan fungsi kernel *Triweight* dan kernel *Epanechnikov*.

## 1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Dapat menjadi referensi dan pengembangan kepada pembaca mengenai model regresi nonparametrik pada penelitian selanjutnya.
2. Mengembangkan wawasan keilmuan dan pegetahuan tentang *Generalized Cross Validation* (GCV)) untuk memodelkan indeks LQ45 menggunakan fungsi kernel *Triweight* dan kernel *Epanechnikov*.

## BAB II

### TINJAUAN PUSTAKA

#### 3.1 Regresi Nonparametrik

Analisis regresi merupakan analisis yang mempelajari hubungan fungsional antara satu atau beberapa variabel prediktor dengan satu variabel respon (Puspitasari dkk, 2012). Pendekatan yang digunakan untuk menduga fungsi tersebut bisa dilakukan dengan parametrik dan nonparametrik. Pendekatan parametrik digunakan untuk mengestimasi model jika kurva regresinya diketahui, sedangkan pendekatan regresi nonparametrik merupakan suatu teknik analisis data dalam statistika yang digunakan untuk melihat hubungan antara variabel prediktor dan variabel respon dimana kurva regresinya tidak diketahui dan tidak ada informasi apapun terkait bentuk kurva regresinya (Eubank, 1988).

#### 3.2 Regresi Kernel

Regresi kernel adalah teknik statistika nonparametrik untuk mengestimasi fungsi regresi  $(x_i)$ . Tujuan analisis regresi adalah menemukan hubungan antara variabel respon dan variabel prediktor. Secara umum model regresi kernel adalah sebagai berikut:

$$y_i = m(x_i) + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.1)$$

dengan  $y_i$  merupakan variabel respon ke- $i$ ,  $m(x_i)$  adalah fungsi regresi yang tidak diketahui bentuk kurvanya dan  $\varepsilon_i$  adalah sisaan yang diasumsikan independen dan berdistribusi normal dengan mean nol dan varian  $\sigma^2$  (Eubank, 1999).

#### 3.3 Estimator Kernel

Estimator kernel diperkenalkan oleh Rosenblatt (1956) dan Parzen (1962) sehingga disebut estimator densitas kernel Rosenblatt-Parzen. Menurut Eubank (1999) estimator kernel merupakan estimator linier yang sama dengan estimator lainnya, perbedaannya hanya karena estimator kernel lebih khusus dalam penggunaan metode *bandwidth*. Penghalusan dengan pendekatan kernel dikenal sebagai pemulus kernel (*kernel smoothing*) yang sangat bergantung pada fungsi kernel dan *bandwidth* (Saputra, 2016).

Estimator kernel merupakan salah satu pendekatan untuk mengestimasi kurva regresi nonparametrik. Metode ini sering digunakan karena memiliki bentuk yang lebih fleksibel dan perhitungan matematisnya mudah dikerjakan (Famalika dan Sihombing, 2022). Beberapa kelebihan estimator kernel adalah fleksibel, bentuk matematisnya mudah, dan dapat mencapai tingkat kekonvergenan yang relatif cepat (Budiantara dan Mulianah, 2007).

Kernel  $K$  dengan *bandwidth*  $h$  secara umum didefinisikan sebagai:

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right), \text{ untuk } -\infty < x < \infty, h > 0 \tag{2.2}$$

serta memenuhi:

- 1)  $K(x) \geq 0$ , untuk semua  $x$
- 2)  $\int_{-\infty}^{\infty} K(x)dx = 1$
- 3) Memiliki nilai variansi berupa konstanta

$$\int_{-\infty}^{\infty} x^2 K(x)dx = \sigma^2 > 0$$

- 4) Memiliki nilai rata-rata sama dengan nol

$$\int_{-\infty}^{\infty} xK(x)dx = 0$$

- 5)  $K(-x) = K(x)$ , untuk semua  $x$  (*simetris*)

Menurut Hardle (1994) secara umum estimator densitas kernel untuk fungsi densitas  $\hat{f}_h(x)$  didefinisikan sebagai berikut:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \tag{2.3}$$

Menurut Hardle (1994) menyebutkan bahwa terdapat beberapa jenis fungsi kernel, sebagai berikut.

**Tabel 2.1** Fungsi Kernel

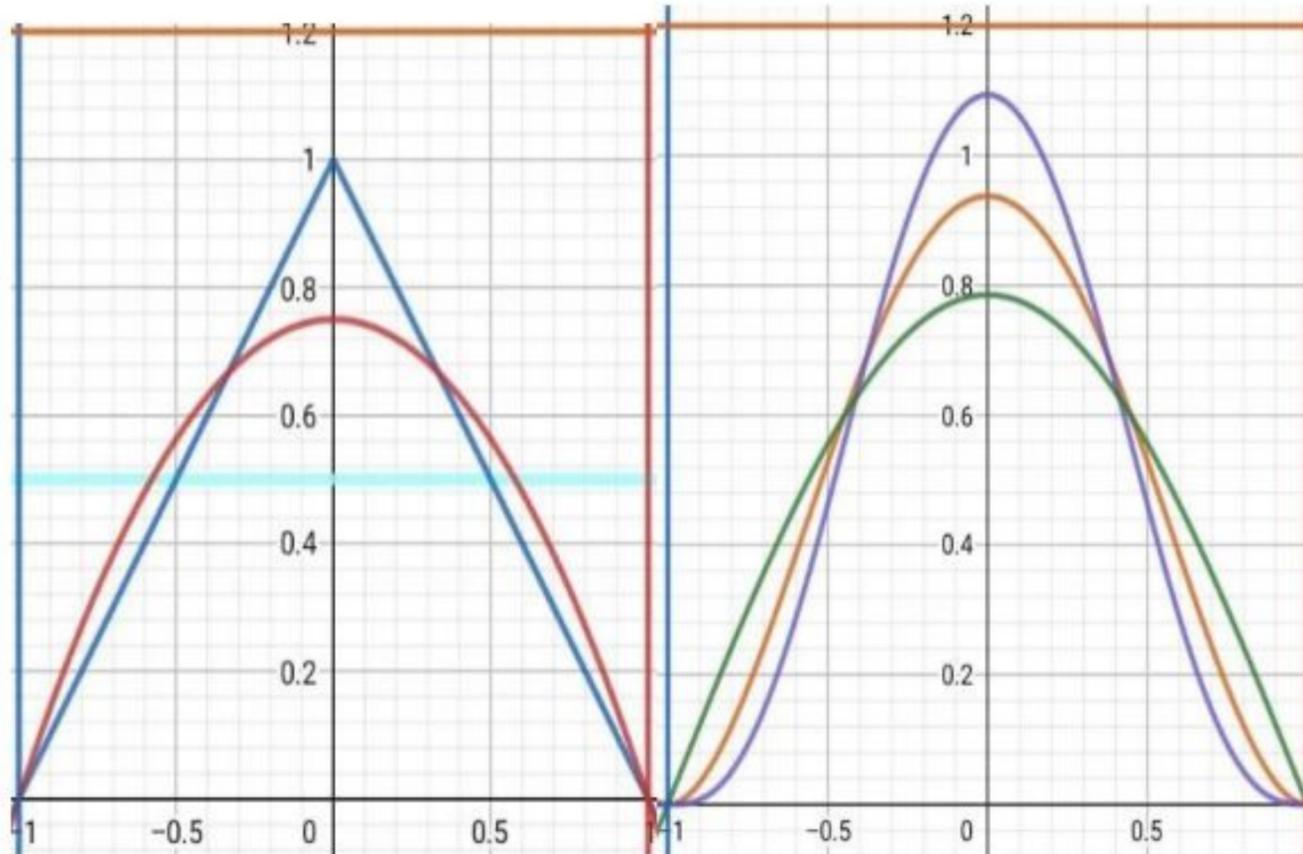
Kernel	$K(x)$	Kondisi
Gaussian	$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) I_A$	$(-\infty < x < \infty)$
Epanechnikov	$K(x) = \frac{3}{4} (1 - x^2) I_A$	$( x  \leq 1)$
Uniform	$K(x) = \frac{1}{2} I_A$	$( x  \leq 1)$
Triangle	$K(x) = (1 -  x ) I_A$	$( x  \leq 1)$

Kernel	$K(x)$	Kondisi
Triweight	$K(x) = \frac{35}{32}(1 - x^2)^3 I_A$	$( x  \leq 1)$
Biweight	$K(x) = \frac{15}{16}(1 - x^2)^2 I_A$	$( x  \leq 1)$
Cosine	$K(x) = \frac{\pi}{4} \cos\left(\frac{\pi}{2}x\right) I_A$	$( x  \leq 1)$

Dengan  $I_A$  adalah fungsi indikator, dengan

$$I_A(x) = \begin{cases} 1 & \text{jika } |x| \leq 1 \\ 0 & \text{jika } |x| > 1 \end{cases}, \text{ untuk } -\infty < x < \infty$$

Fungsi kernel *Gaussian* dengan fungsi  $I_A(-\infty < x < \infty)$  sedangkan fungsi kernel lainnya dengan fungsi indikator  $I_A(|x| \leq 1)$ .



Gambar 2.1 Fungsi Kernel

Gambar 2.2 Fungsi Kernel

- Keterangan:
- a. Biru muda : Kernel *Uniform*
  - b. Biru tua : Kernel *Triangle*
  - c. Merah : Kernel *Epanechnikov*
  - d. Oranye : Kernel *Quartik*
  - e. Ungu : Kernel *Triweight*
  - f. Hijau : Kernel *Cosinus*

Pada penelitian ini yang digunakan ada dua jenis fungsi kernel, yaitu kernel *triweight* dan kernel *epanechnikov*. Bentuk umum fungsi kernel *triweight* adalah:

$$K(x) = \frac{35}{32}(1 - x^2)^3 I_A \quad (2.4)$$

Sedangkan bentuk umum dari fungsi kernel *epanechnikov* adalah sebagai berikut:

$$K(x) = \frac{3}{4}(1 - x^2) I_A \quad (2.5)$$

### 3.4 Estimasi Priestley Chao

Berdasarkan Persamaan (2.1) diketahui bahwa  $m(x_i)$  merupakan fungsi yang tidak diketahui persamaannya dan ditaksir dengan  $\hat{m}(x_i)$ . Adapun  $\hat{m}(x_i)$  didefinisikan di interval  $[a, b]$  serta dapat diperkirakan secara nonparametrik menggunakan data yang ada. Salah satu estimator yang dipertimbangkan untuk digunakan adalah estimator *Priestley-Chao*. estimator ini diusulkan oleh Priestley dan Chao (1972) dan dinyatakan:

$$\hat{m}(x_i) = \frac{\delta}{h} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) y_i \quad (2.6)$$

dengan  $\delta = \frac{(b-a)}{n}$  untuk  $x \in (a, b)$ , fungsi *kernel* diasumsikan simetris nol, maka menjadi  $\int K(u)^2 du < \infty$  dan memiliki momen kedua yang terbatas  $\int u^2 K(u) du = \sigma_K^2 < \infty$ . Estimator *Priestley-Chao* merupakan rata-rata terboboti dari variabel respon  $y_1, y_2, \dots, y_n$  dan bobot  $K\left(\frac{x-x_i}{h}\right)$  dengan hasil sebenarnya ditentukan oleh kedekatan  $x$  ke  $x_i$  relatif terhadap nilai  $h$ . Jika data tidak memiliki jarak yang sama maka estimatornya adalah (He, 2019):

$$\hat{m}(x_i) = \frac{1}{h} \sum_{i=1}^n (x_i - x_{i-1}) K\left(\frac{x-x_i}{h}\right) y_i, \text{ dengan } x_0 = 0, \quad (2.7)$$

Keterangan :

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

$y_i$  = variabel respon pada pengamatan ke-i

$x_i$  = variabel prediktor pada pengamatan ke-i

$x$  = sampel random

$K$  = kernel

$n$  = banyak pengamatan

$h$  = *bandwidth*

$i$  = 1, 2, ...,  $n$  maka  $x_0 = 0$

### 3.5 Pemilihan *Bandwidth* Optimum

Ketika *bandwidth* terlalu besar maka bias pemodelannya akan besar dan variansi akan kecil sebaliknya jika *bandwidth* terlalu kecil maka bias pemodelannya akan kecil dan variansi akan besar (Suparti dan Prahutama, 2016).

Pemilihan *bandwidth* yang optimal melibatkan penyeimbangan antara bias dan variansi. Pemilihan *bandwidth* yang optimal dilakukan dengan cara memperkecil tingkat kesalahan. Semakin kecil tingkat kesalahan maka semakin baik estimasinya. Salah satu metode yang digunakan untuk melihat hubungan antara bias dan variansi adalah *Mean Square Error*/MSE (Karimuse dkk, 2023).

Apabila *bandwidth* yang digunakan terlalu kecil maka estimasi kurva regresi yang dihasilkan kurang halus (*under-smoothing*). Sebaliknya, apabila *bandwidth* yang digunakan terlalu besar maka estimasi kurva regresi yang dihasilkan sangat mulus (*over-smoothing*) sehingga tidak sesuai dengan pola distribusi data. Oleh karena itu, nilai *bandwidth* yang optimum harus dipilih supaya estimasi terbaik dapat dihasilkan (Hardle, 1994). Dalam penelitian ini, pemilihan *bandwidth* dilakukan menggunakan metode *Generalized Cross Validation* (GCV).

### 2.6 *Generalized Cross Validation*

Pemilihan *bandwidth* dilakukan menggunakan metode *Generalized Cross Validation* (GCV). Metode GCV digunakan untuk menentukan parameter pemulus dalam regresi kernel. Bentuk umum metode GCV adalah sebagai berikut :

$$GCV = \frac{MSE}{\left(\frac{1}{n}tr(I-H(h))\right)^2} \quad (2.8)$$

dengan :

$n$  = banyak data

$I$  = matriks identitas

$h$  = *bandwidth*

$X$  = matriks data

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

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

### 2.7 Mean Square Error

Menurut Komang dan Gusti (2012), kebaikan suatu estimator dapat dilihat dari tingkat kesalahannya. Untuk mengukur *error* biasanya digunakan *Mean Square Error* (MSE) adalah rata-rata dari kuadrat kesalahan.

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.9)$$

dengan :

- $n$  = banyak data
- $y_i$  = nilai variabel respon
- $\hat{y}_i$  = nilai prediksi dari variabel respon

### 2.8 Saham

Saham adalah tanda bukti memiliki perusahaan dimana pemiliknya disebut sebagai pemegang saham. Dengan kata lain, saham dapat diartikan sebagai surat bukti pemilikan modal seseorang atau badan usaha sebagai pihak tertentu dalam perusahaan atau perseroan terbatas yang memperoleh penghasilan dan kekayaan perusahaan tersebut (Defrizal & Mulyawan, 2015). Masing-masing lembar saham biasa mewakili suatu suara tentang segala hal dalam pengurusan perusahaan dan menggunakan suara tersebut dalam rapat tahunan perusahaan dan pembagian keuntungan (Hadi, 2013).

### 2.9 Pasar Modal

Pasar Modal adalah pasar keuangan untuk dana jangka panjang dan dalam arti sempit merupakan pasar yang konkrit. Pasar Modal dalam arti sempit adalah suatu tempat dalam pengertian fisik yang terorganisasi tempat efek diperdagangkan yang disebut Bursa Efek (Rustiana dan Ramadhani, 2022).

### 2.10 Indeks Harga Saham

Indeks harga saham berfungsi sebagai indikator yang menunjukkan pergerakan harga saham dan tren pasar untuk memperlihatkan gambaran kondisi pasar, apakah saham sedang aktif atau lemah. Terdapat beberapa indeks harga saham di Bursa Efek Indonesia (BEI), diantaranya yaitu indeks individual, indeks harga saham sektoral, indeks harga saham gabungan (IHSG), indeks LQ45, indeks syariah atau *Jakarta Index*

*Islamic* (JII), indeks papan utama dan papan pengembangan, indeks KOMPAS 100, indeks bisnis-27, indeks PERFINDO-25, dan indeks Sri-kehati.

### **2.10.1 Indeks Saham LQ45**

Saham LQ45 merupakan saham yang diminati oleh banyak investor. Hal ini bisa diperhatikan dari data frekuensi perdagangan. Frekuensi perdagangan menunjukkan berapa kali suatu saham diperjualbelikan pada periode tertentu. Pergerakan harga saham salah satunya diakibatkan adanya interaksi pasar. Dikatakan saham tersebut likuid jika frekuensi perdagangan sahamnya besar. Sedangkan, dikatakan saham tidak likuid atau tidak membuat investor tertarik jika frekuensi perdagangan sahamnya kecil. Oleh karena itu, frekuensi perdagangan menjadi perhatian investor untuk menentukan strategi investasi yang tepat (Cahyani dkk, 2021).