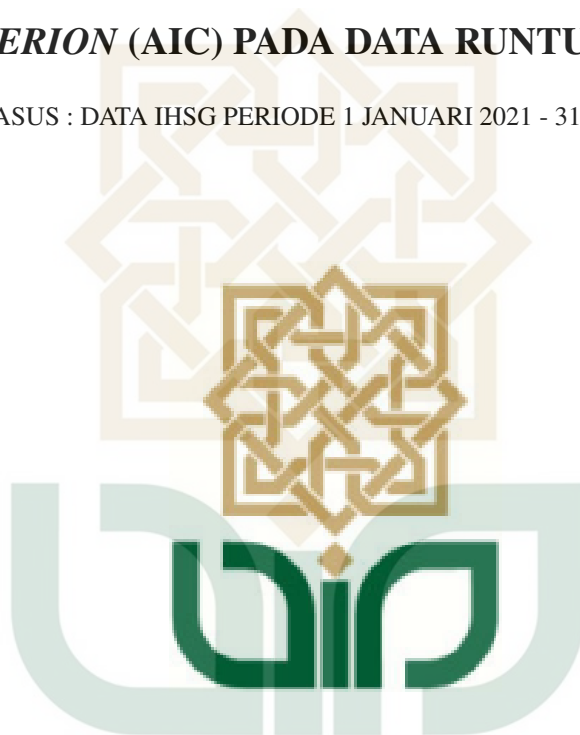


**SKRIPSI**

**PERBANDINGAN OPTIMASI MODEL REGRESI  
*PENALIZED SPLINE* MENGGUNAKAN *GENERALIZED  
CROSS VALIDATION (GCV)* DAN *AKAIKE INFORMATION  
CRITERION (AIC)* PADA DATA RUNTUN WAKTU**

(STUDI KASUS : DATA IHSG PERIODE 1 JANUARI 2021 - 31 DESEMBER 2022)



STATE ISLAMIC UNIVERSITY  
**DEVITA RISKA NURAINI**  
NIM. 19106010046  
SUNAN KALIJAGA  
YOGYAKARTA

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA  
YOGYAKARTA**

**2023**

**PERBANDINGAN OPTIMASI MODEL REGRESI  
*PENALIZED SPLINE* MENGGUNAKAN *GENERALIZED  
CROSS VALIDATION (GCV)* DAN *AKAIKE INFORMATION  
CRITERION (AIC)* PADA DATA RUNTUN WAKTU**

(STUDI KASUS : DATA IHSG PERIODE 1 JANUARI 2021 - 31 DESEMBER 2022)

Skripsi

Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh Derajat  
Sarjana Matematika



**DEVITA RISKA NURAINI**  
**NIM. 19106010046**  
STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA  
YOGYAKARTA**

**2023**



## **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 : Devita Riska Nuraini

NIM : 19106010046

Judul Skripsi : Perbandingan Optimasi Model Regresi *Penalized Spline* Menggunakan *Generalized Cross Validation* (GCV) dan *Akaike Information Criterion* (AIC) pada Data Runtun Waktu (Studi Kasus : Data Indeks Harga Saham Gabungan (IHSG) Periode 1 Januari 2021 – 31 Desember 2022)

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 dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

*Wassalamu'alaikum wr. wb.*

Yogyakarta, 9 Agustus 2023

Pembimbing I

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.  
NIP. 19741003 200003 2 002

Pembimbing II

Muhamad Rashif Hilmi, S.Si., M.Sc  
NIP. 19920309 202012 1 001



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-2215/Un.02/DST/PP.00.9/08/2023

Tugas Akhir dengan judul : PERBANDINGAN OPTIMASI MODEL REGRESI PENALIZED SPLINE MENGGUNAKAN GENERALIZED CROSS VALIDATION (GCV) DAN AKAIKE INFORMATION CRITERION (AIC) PADA DATA RUNTUN WAKTU (STUDI KASUS : DATA IHSG PERIODE 1 JANUARI 2021 - 31 DESEMBER 2022)

yang dipersiapkan dan disusun oleh:

Nama : DEVITA RISKA NURAINI  
Nomor Induk Mahasiswa : 19106010046  
Telah diujikan pada : Selasa, 15 Agustus 2023  
Nilai ujian Tugas Akhir : A-

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

TIM UJIAN TUGAS AKHIR



Ketua Sidang

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

Valid ID: 64e32a39edd54



Penguji I

Muhamad Rashif Hilmi, S.Si., M.Sc.  
SIGNED

Valid ID: 64e2cf8e8ac31



Penguji II

Sri Istiyarti Uswatun Chasanah, M.Si.  
SIGNED

Valid ID: 64e2e00e99c82



Yogyakarta, 15 Agustus 2023  
UIN Sunan Kalijaga  
Dekan Fakultas Sains dan Teknologi

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

Valid ID: 64e416515f320

## SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Devita Riska Nuraini  
NIM : 19106010046  
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, 10 Agustus 2023



Devita Riska Nuraini  
NIM. 19106010046

STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## **PERSEMBAHAN**

KARYA TULIS SEDERHANA INI PENULIS PERSEMBAHKAN

UNTUK :

KEDUA ORANG TUA, BAPAK HARTIYONO DAN IBU RUBIYATUN;

NENEK PENULIS, SIMBAH SUTIAH;

KEDUA ADIK PENULIS, AHMAD SOFYAN DWI MURYANTO DAN AHMAD FAHRI

BUDIANTO;

YANG TELAH MEMBERIKAN DO'A, DUKUNGAN, CINTA DAN KASIH SAYANG YANG TULUS TAK TERHINGGA SERTA MENJADI SUPPORT SYSTEM TERBAIK PENULIS.

KELUARGA BESAR SIMBAH SUHADI DAN SIMBAH MARDI, SAHABAT-SAHABAT, TEMAN-TEMAN, DAN ORANG-ORANG BAIK YANG PERNAH SAYA KENAL BAIK DALAM

WAKTU PERKULIAHAN MAUPUN DI WAKTU SEBELUMNYA.

SERTA ALMAMATER UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA  
MATEMATIKA ANGGKATAN 2019 UIN SUNAN KALIJAGA YOGYAKARTA.

STATE ISLAMIC UNIVERSITY  
**SUNAN KALIJAGA**  
YOGYAKARTA

## MOTTO

***”Sesungguhnya beserta kesulitan itu ada kemudahan”***

(QS.Al-Insyirah:6)

***”Ketika kamu ikhlas menerima semua kekecewaan hidup, maka Allah akan membayar tuntas kekecewaanmu dengan beribu-ribu kebaikan”***

(Ali bin Abi Thalib)

***”Jangan takut, Allah sebaik-baiknya perencana. Yakinlah apapun itu, akan selalu ada hikmah di balik rencana-Nya”***



STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## PRAKATA

*Assalamu'alaikum Warahmatullahi Wabarakatuh*

*Alhamdulillah*, puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Perbandingan Optimasi Model Regresi *Penalized Spline* Menggunakan *Generalized Cross Validation* (GCV) dan *Akaike Information Criterion* (AIC) pada Data Runtun Waktu. Tak lupa, shalawat serta salam senantiasa kita haturkan kepada Nabi Besar Muhammad SAW yang kita tunggu syafa'atnya di *yaumul qiyamah* nanti. Skripsi ini ditulis dalam rangka memenuhi salah satu syarat untuk memperoleh gelar Sarjana Matematika, Fakultas Sains dan Teknologi, UIN Sunan Kalijaga Yogyakarta. Penulis menyadari bahwa selama penyusunan skripsi ini telah banyak mendapatkan masukan, saran, bimbingan, bantuan, dan motivasi dari semua pihak. Oleh karena itu pada kesempatan kali ini penulis ingin menyampaikan ucapan terima kasih kepada :

1. Bapak Prof. Dr. Phil. Al Makin, S.Ag., M.A., selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Ibu Prof. Dr. Dra. Hj. Khurul Wardati, M.Si. selaku Dekan Fakultas Sains dan Teknologi, UIN Sunan Kalijaga Yogyakarta.
3. Bapak Muchammad Abrori, S.Si., M.Kom. selaku Ketua Program Studi Matematika beserta jajarannya.
4. Ibu Sri Utami Zuliana, S.Si., M.Sc., Ph.D. selaku Dosen Pembimbing skripsi yang telah berjasa dalam penyelesaian skripsi ini dan selalu membimbing dengan baik berupa arahan, motivasi, waktu, pikiran, dan lain sebagainya selama penulisan skripsi ini.
5. Bapak Muhamad Rashif Hilmi, S.Si., M.Sc. selaku Dosen Pembimbing skripsi yang telah berjasa dalam penyelesaian skripsi ini dan selalu membimbing dengan baik dan sabar selama penulisan skripsi.



6. Ibu Malahayati, S.Si., M.Sc. selaku dosen pembimbing akademik yang telah membimbing dari awal perkuliahan.
7. Seluruh dosen dan staf Fakultas Sains dan Teknologi yang selalu memberikan ilmu dan pelayanan yang terbaik selama perkuliahan sampai pada selesainya penyusunan skripsi ini.
8. Kedua orang tua tercinta, Bapak Hartiyono dan Ibu Rubiyatun yang selalu memberikan semangat, cinta, kasih sayang tulus tanpa batas, perhatian, doa, nasihat sehingga menjadi *support system* terbaik penulis.
9. Kedua adik tercinta, Ahmad Sofyan Dwi Muryanto dan Ahmad Fahri Budi-anto, yang menjadi sumber semangat penulis untuk selalu berjuang.
10. Sahabat Hateda, yaitu Marizka Hanani, Ranastya Ayyu, Diah Pratiwi, Elfira Diah, dan Aisyah Abdillah, yang setia menemani penulis sejak bangku SMA dalam keadaan suka maupun duka.
11. Sahabat Pulu-pulu, yaitu Ara Syaifha, Gayatri Putri, Tsalisa Zuraida, dan Indriani Wahyu, yang setia menemani di sepek terjang dunia perkuliahan sampai terselesaikannya skripsi ini.
12. Teman-teman satu bimbingan tugas akhir, Ara Syaifha, Gayatri Putri, dan Sheliea Candara, yang sudah berbagi ilmu dan menemani perjuangan penulis untuk menyelesaikan skripsi ini.
13. Teman-teman program studi Matematika angkatan 2019, yang telah berjuang bersama, selalu berbagi ilmu, dan pengalaman yang mengesankan selama masa perkuliahan.
14. Semua pihak yang telah memberikan dukungan dan doa yang tidak dapat saya sebutkan satu-persatu, terima kasih.

Semoga segala kebaikan dan pertolongan dari semua pihak mendapatkan berkah dan selalu dalam perlindungan Allah SWT. Penulis menyadari jika dalam penyusunan skripsi ini masih banyak terdapat kekurangan, baik dari segi tata bahasa maupun susunan kalimat. Oleh karena itu, penulis mengharapkan kritik dan

saran yang membangun untuk perbaikan selanjutnya. Semoga skripsi ini dapat bermanfaat, baik bagi penulis maupun semua pihak yang membutuhkan. Akhir kata, semoga kita selalu diberi kesehatan, kebahagiaan, dan perlindungan-Nya. *Aamiin ya robbal alamiin.*

*Wassalamu'alaikum Warahmatullahi Wabarakatuh*

Yogyakarta, 4 Agustus 2023

Penulis



## DAFTAR ISI

<b>HALAMAN JUDUL</b> . . . . .	<b>i</b>
<b>HALAMAN PERSETUJUAN</b> . . . . .	<b>ii</b>
<b>HALAMAN PENGESAHAN</b> . . . . .	<b>iii</b>
<b>HALAMAN PERNYATAAN</b> . . . . .	<b>iv</b>
<b>HALAMAN PERSEMBAHAN</b> . . . . .	<b>v</b>
<b>HALAMAN MOTTO</b> . . . . .	<b>vi</b>
<b>PRAKATA</b> . . . . .	<b>vii</b>
<b>DAFTAR ISI</b> . . . . .	<b>x</b>
<b>DAFTAR TABEL</b> . . . . .	<b>xiii</b>
<b>DAFTAR GAMBAR</b> . . . . .	<b>xiv</b>
<b>DAFTAR LAMBANG</b> . . . . .	<b>xv</b>
<b>INTISARI</b> . . . . .	<b>xvi</b>
<b>ABSTRACT</b> . . . . .	<b>xvii</b>
<b>I PENDAHULUAN</b> . . . . .	<b>1</b>
1.1. Latar Belakang Masalah . . . . .	1
1.2. Batasan Masalah . . . . .	3
1.3. Rumusan Masalah . . . . .	3
1.4. Tujuan Penelitian . . . . .	4
1.5. Manfaat Penelitian . . . . .	4
1.6. Tinjauan Pustaka . . . . .	4
1.7. Sistematika Penulisan . . . . .	6
<b>II DASAR TEORI</b> . . . . .	<b>8</b>
2.1. Data Runtun Waktu . . . . .	8
2.2. Model <i>Autoregressive Integrated Moving Average</i> (ARIMA) . . . . .	9

2.3. Asumsi Runtun Waktu . . . . .	11
2.4. Analisis Regresi . . . . .	14
2.5. Regresi Nonparametrik . . . . .	16
2.6. <i>Spline</i> . . . . .	16
2.7. Regresi <i>B-Spline</i> . . . . .	18
2.8. Basis-Basis <i>B-Spline</i> . . . . .	21
2.9. <i>Penalized Spline (P-Spline)</i> . . . . .	21
2.10. Pemilihan Orde, Jumlah Knot, dan $\lambda$ Optimal . . . . .	24
2.10.1. <i>Generalized Cross Validation (GCV)</i> . . . . .	24
2.10.2. <i>Akaike Information Criterion (AIC)</i> . . . . .	25
2.11. Pengujian Signifikansi Parameter . . . . .	27
2.12. Ketepatan Metode Peramalan . . . . .	28
2.13. Indeks Harga Saham Gabungan (IHSG) . . . . .	29
2.13.1. Fungsi IHSG . . . . .	29
<b>III METODE PENELITIAN . . . . .</b>	<b>31</b>
3.1. Jenis Penelitian . . . . .	31
3.2. Metode Pengumpulan Data . . . . .	31
3.3. Variabel Penelitian dan Sumber Data . . . . .	31
3.4. Metode Analisis Data . . . . .	32
3.5. <i>Software</i> yang Digunakan . . . . .	33
<b>IV HASIL DAN PEMBAHASAN . . . . .</b>	<b>34</b>
4.1. Deskripsi Data . . . . .	34
4.2. Pemodelan <i>Autoregressive Integrated Moving Average (ARIMA)</i> . . . . .	35
4.2.1. Uji Stasioneritas . . . . .	35
4.2.2. Model ARIMA . . . . .	37
4.2.3. Model ARIMA Terbaik . . . . .	38

4.2.4. Estimasi dan Uji Signifikansi Parameter . . . . .	39
4.3. Uji Asumsi Model ARIMA . . . . .	39
4.4. Pemilihan Orde, Jumlah Knot dan $\lambda$ Optimal Regresi <i>Penalized Spline</i>	41
4.5. Model Regresi <i>Penalized Spline</i> Terbaik . . . . .	42
4.6. Pengujian Signifikansi Parameter . . . . .	44
4.7. Interpretasi Model . . . . .	45
4.8. Ketepatan Metode Peramalan . . . . .	47
4.9. Prediksi Data IHSG Menggunakan Model Regresi <i>Penalized Spline</i> Terbaik . . . . .	49
4.10. Pembahasan . . . . .	50
<b>V PENUTUP . . . . .</b>	<b>54</b>
5.1. Kesimpulan . . . . .	54
5.2. Saran . . . . .	55
<b>DAFTAR PUSTAKA . . . . .</b>	<b>57</b>
<b>LAMPIRAN A . . . . .</b>	<b>59</b>
<b>LAMPIRAN B . . . . .</b>	<b>62</b>
<b>LAMPIRAN C . . . . .</b>	<b>94</b>
<b>LAMPIRAN D . . . . .</b>	<b>95</b>
<b>Curriculum Vitae . . . . .</b>	<b>97</b>

## DAFTAR TABEL

2.1	Analisis Variansi (ANOVA) Signifikansi Parameter . . . . .	28
4.1	Deskripsi Data Variabel $Y_{i-1}$ . . . . .	34
4.2	Deskripsi Data Variabel $Y_i$ . . . . .	34
4.3	Nilai AIC Model ARIMA . . . . .	39
4.4	Estimasi Parameter Model ARIMA . . . . .	39
4.5	Orde, Jumlah Knot, dan $\lambda$ Optimal . . . . .	42
4.6	Estimasi Parameter Menggunakan Metode AIC . . . . .	44
4.7	Analisis Uji Serentak Parameter Model GCV) . . . . .	45
4.8	Analisis Uji Serentak Parameter Model AIC . . . . .	45
4.9	Uji Kinerja Model Menggunakan Nilai <i>Mean Percentage Error</i> (MA- PE) . . . . .	49
4.10	Perbandingan Nilai Prediksi dan Nilai Aktual IHSG dengan Model AIC . . . . .	50

STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## DAFTAR GAMBAR

2.1	Contoh <i>B-Spline</i> derajat 1 (orde 2) (Eilers and Marx, 1996) . . . . .	18
2.2	Contoh <i>B-Spline</i> derajat 2 (orde 3) (Eilers and Marx, 1996) . . . . .	19
4.1	Grafik Data Indeks Harga Saham Gabungan (IHSG) Periode 1 Januari 2021 sampai 31 Desember 2022 . . . . .	35
4.2	Grafik Differensi Data Indeks Harga Saham Gabungan (IHSG) Periode 1 Januari 2022 sampai 31 Desember 2022 . . . . .	36
4.3	Uji <i>Autocorrelation Function</i> (ACF) . . . . .	37
4.4	Uji <i>Partial Autocorrelation Function</i> (PACF) . . . . .	38
4.5	Grafik Prediksi IHSG Model <i>ARIMA(0,1,1)</i> . . . . .	41
4.6	Grafik Hubungan Antara $\lambda$ dan GCV . . . . .	43
4.7	Grafik Hubungan Antara $\lambda$ dan AIC . . . . .	43
4.8	Grafik Perbandingan Data Aktual dengan Data Prediksi Menggunakan Metode <i>Generalized Cross Validation</i> (GCV) . . . . .	48
4.9	Grafik Perbandingan Data Aktual dengan Data Prediksi Menggunakan Metode <i>Akaike Information Criterion</i> (AIC) . . . . .	48

## DAFTAR LAMBANG

$Y_{i-1}$	: Variabel Prediktor
$Y_i$	: Variabel Respon
$f(x)$	: Fungsi Kurva Regresi
$\epsilon$	: Error
$\beta$	: Parameter
$K$	: Titik Knot
$m$	: Orde
$p$	: Derajat
$n$	: Banyak Data
$k$	: Jumlah Knot
$\lambda$	: Parameter Penghalus
$j$	: Letak Titik Knot



STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA



## INTISARI

# PERBANDINGAN OPTIMASI MODEL REGRESI *PENALIZED SPLINE* MENGUNAKAN *GENERALIZED CROSS VALIDATION (GCV)* DAN *AKAIKE INFORMATION CRITERION (AIC)* PADA DATA RUNTUN WAKTU

(Studi Kasus : Data IHSG Periode 1 Januari 2021 - 31 Desember 2022)

Oleh

DEVITA RISKA NURAINI

NIM. 19106010046

Data Indeks Harga Saham Gabungan (IHSG) merupakan data runtun waktu yang dapat dimodelkan dengan model parametrik, salah satunya yaitu dengan model *Autoregressive Integrated Moving Average (ARIMA)*. Pada ARIMA terdapat beberapa asumsi yang harus dipenuhi, yaitu asumsi stasioneritas, asumsi normalitas, dan asumsi *white noise*. Pada umumnya, data IHSG yang berfluktuasi melanggar asumsi tersebut. Alternatif dari permasalahan tersebut yaitu menggunakan pendekatan nonparametrik. Regresi *penalized spline* merupakan salah satu regresi nonparametrik yang dapat digunakan. Model regresi *penalized spline* optimal bergantung pada pemilihan kombinasi orde, jumlah knot, dan  $\lambda$  berdasarkan nilai *Generalized Cross Validation (GCV)* minimum dan *Akaike Information Criterion (AIC)* minimum. Diperoleh kombinasi orde, jumlah knot, dan  $\lambda$  optimal pada penelitian ini adalah pada orde 2 (linear) dan 1 titik knot yaitu 6829.922 dengan masing-masing  $\lambda$  optimal pada GCV sebesar  $2.828283 \times 10^{-4}$  dengan nilai GCV sebesar 15600.36, dan  $\lambda$  optimal AIC sebesar  $1 \times 10^{-16}$  dengan nilai AIC sebesar 1299.176. Selanjutnya akan dibandingkan nilai MAPE dari kedua model tersebut. Nilai MAPE pada model GCV sebesar 1.316928 dan nilai MAPE pada model AIC sebesar 1.316881. Dikarenakan nilai MAPE pada model AIC lebih kecil daripada model GCV, maka model AIC dinilai lebih baik dalam peramalan. Model terbaik yang diperoleh adalah  $\hat{y} = 5813.654B_{-1,2}(x) + 6858.172B_{0,2}(x) + 7220.146B_{1,2}(x)$

**Kata Kunci:** AIC, GCV, MAPE, Regresi *Penalized Spline*, Runtun Waktu

## ABSTRACT

### COMPARISON OF REGRESSION MODEL OPTIMIZATION PENALIZED SPLINE USING GENERALIZED CROSS VALIDATION (GCV) AND AKAIKE INFORMATION CRITERION (AIC) FOR TIME SERIES DATA

By

DEVITA RISKANURAINI

NIM. 19106010046

IHSG data is time series data that can be modeled using parametric models, one of which is the Autoregressive Integrated Moving Average (ARIMA) model. In ARIMA, there are several assumptions that must be met, namely the assumption of stationarity, the assumption of normality, and the assumption of white noise. In general, IHSG data fluctuates doesn't occupy these assumptions. An alternative to these problems is to use a nonparametric approach. Penalized spline regression is one nonparametric regression that can be used. The optimal penalized spline regression model depends on choosing a combination of order, number of knots, and  $\lambda$  based on the minimum Generalized Cross Validation (GCV) value and the Akaike Information Criteria (AIC) minimum. The optimal the combination of order, number of knots, and  $\lambda$  obtained in this study is at order 2 (linear) and 1 knot point is 6829.922, with each optimal  $\lambda$  at GCV of  $2.828283 \times 10^{-4}$  with a GCV value of 15600.36 and the optimal  $\lambda$  at AIC of  $1 \times 10^{-16}$  with an AIC value of 1299,176. Furthermore, the MAPE values of the two models will be compared. The MAPE value in the GCV model is 1.316928, and the MAPE value in the AIC model is 1.316881. Because the MAPE value in the AIC model is smaller than the GCV model, the AIC model is considered better for forecasting. The best model obtained is  $\hat{y} = 5813.654B_{-1,2}(x) + 6858.172B_{0,2}(x) + 7220.146B_{1,2}(x)$

**Keywords:** AIC, GCV, MAPE, Penalized Spline Regression, Time Series

# BAB I

## PENDAHULUAN

Bab pendahuluan merupakan bagian pertama dalam skripsi ini. Pada bab ini dibagi menjadi enam subbab, antara lain latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

### 1.1. Latar Belakang Masalah

Analisis regresi merupakan suatu teknik statistika yang bertujuan untuk pemeriksaan dan pemodelan hubungan antarvariabel. Analisis regresi secara umum dibagi menjadi tiga, yaitu analisis regresi parametrik, analisis regresi semiparametrik, dan analisis regresi nonparametrik. Apabila kurva regresi diasumsikan telah membentuk pola hubungan tertentu, misalnya berbentuk linear, kuadratik, kubik, dan lainnya, maka dapat digunakan analisis regresi dengan pendekatan parametrik. Sedangkan untuk memodelkan pola data yang tidak mengikuti pola tertentu menggunakan pendekatan nonparametrik. Regresi nonparametrik menjadi solusi masalah regresi parametrik karena pada regresi nonparametrik tidak memerlukan asumsi-asumsi yang harus dipenuhi.

Investasi di Indonesia terus mengalami perkembangan. Salah satu bentuk investasi yang menjanjikan di masyarakat adalah investasi di bidang pasar modal. Saat ini pasar modal sedang berada dalam tahap pertumbuhan. Terakhir semenjak pandemi Covid-19, di setiap tahunnya terdapat kenaikan jumlah investor di pasar modal, baik di bursa saham, reksadana, *C-Best*, maupun Surat Berharga Negara (SBN). Kenaikan jumlah investor tersebut pastinya berdampak positif bagi perkembangan pasar modal di Indonesia. Pasar modal (*capital market*) ini menyediakan atau sebagai tempat jual beli instrumen investasi jangka panjang. Salah satu instrumen investasi yang diperjualbelikan tersebut adalah saham.

Pada Otoritas Jasa Keuangan (OJK), saham didefinisikan sebagai tanda pe-

nyertaan modal seseorang atau pihak (badan usaha) pada suatu perusahaan atau Perseroan Terbatas. Dengan menyertakan modal tersebut, maka pihak tersebut memiliki klaim (hak) atas pendapatan perusahaan, aset perusahaan, dan berhak hadir dalam Rapat Umum Pemegang Saham (RUPS). Seperti halnya prinsip jual beli, dalam berinvestasi tentunya terdapat dua kemungkinan, yaitu mendapatkan keuntungan, atau harus menerima kerugian. Oleh karena itu, sebelum mengambil keputusan untuk memulai berinvestasi, diperlukan informasi penting, salah satunya yaitu Indeks Harga Saham Gabungan (IHSG).

Indeks Harga Saham Gabungan (IHSG) merupakan indeks gabungan dari seluruh jenis saham-saham perusahaan di Bursa Efek Indonesia (BEI). IHSG berubah setiap hari karena perubahan harga pasar yang terjadi setiap hari dan adanya saham tambahan. Pergerakan nilai IHSG biasanya menunjukkan perubahan situasi pasar yang terjadi. Seiring dengan perkembangan dan dinamika pasar, IHSG mengalami periode naik dan turun. Nilai IHSG yang mengalami kenaikan menggambarkan kondisi pasar yang sedang aktif. Nilai IHSG yang tetap menunjukkan keadaan pasar yang stabil, sedangkan nilai IHSG yang menurun menggambarkan kondisi pasar yang sedang lesu. Hal ini dikarenakan IHSG menggunakan seluruh saham yang tercatat di Bursa Efek Indonesia sebagai komponen penghitungan indeks (Anoraga & Pakarti, 2006).

Data Indeks Harga Saham Gabungan merupakan data yang berbentuk data runtun waktu. Terdapat beberapa metode yang dapat digunakan dalam memodelkan suatu data runtun waktu, salah satunya yaitu *Autoregressive Integrated Moving Average* (ARIMA). Model ini termasuk dalam pemodelan parametrik yang terikat oleh beberapa asumsi, diantaranya asumsi stasioneritas, asumsi normalitas, dan asumsi *white noise*. Data IHSG yang fluktuatif menyebabkan terjadinya pelanggaran asumsi parametrik, sehingga untuk mengatasinya digunakan regresi nonparametrik karena regresi nonparametrik bersifat fleksibel, tidak memerlukan asumsi-asumsi yang harus dipenuhi. Salah satu pendekatan nonparametrik yang dapat digunakan untuk membentuk model pada data runtun waktu yaitu regresi *spline*.

*Spline* merupakan analisis regresi yang berupa potongan polinomial tersegmen yang kontinu. Kelebihan dari regresi ini yaitu mampu menyesuaikan data menjadi lebih efektif. *Spline* dihubungkan oleh titik-titik knot yang dapat menjelaskan karakteristik dari data. Knot merupakan titik fokus dalam fungsi *spline* sehingga kurva yang dibentuk tersegmen pada titik tersebut. Dikarenakan dalam *spline* bentuk kurva yang dihasilkan masih kasar, sehingga diperlukan sebuah parameter kehalusan (*smoothing*) dengan menggunakan regresi *penalized spline*. Oleh karena itu, dengan menggunakan regresi *penalized spline* dengan metode *Generalized Cross Validation* (GCV) dan *Akaike Information Criterion* (AIC) dapat dilakukan pemodelan nilai IHSG yang akan berguna bagi para investor dan pelaku usaha sebagai bahan pertimbangan dalam mengambil keputusan investasi saham. Selanjutnya dari kedua metode tersebut, akan dibandingkan model mana yang menghasilkan model yang terbaik.

## 1.2. Batasan Masalah

1. Pemilihan model optimasi regresi *penalized spline* menggunakan metode *Generalized Cross Validation* (GCV), *Akaike Information Criterion* (AIC), dan *Mean Absolut Percentage Error* (MAPE).
2. Jumlah knot yang digunakan yaitu 1, 2, dan 3.
3. Orde yang digunakan yaitu orde 2, 3, dan 4.
4. Besar  $\lambda$  yang digunakan yaitu  $1 \times 10^{-16}$  sampai  $1 \times 10^{-3}$ .
5. Pengolahan data menggunakan bahasa pemrograman R versi 4.2.2.

## 1.3. Rumusan Masalah

Berdasarkan latar belakang dan batasan masalah di atas, terdapat beberapa rumusan masalah yang dapat dibuat sebagai berikut :

1. Bagaimana teori analisis regresi *penalized spline*?
2. Bagaimana memodelkan Indeks Harga Saham (IHSG) dengan pendekatan regresi nonparametrik *penalized spline*?

3. Bagaimana mencari model optimal *penalized spline* dan manakah model terbaik antara *Generalized Cross Validation* (GCV) dan *Akaike Information Criterion* (AIC)?

#### 1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dikembangkan, dapat diketahui bahwa tujuan dari penelitian ini adalah sebagai berikut :

1. Mengkaji regresi *penalized spline*.
2. Mengkaji optimasi *Generalized Cross Validation* (GCV) dan *Akaike Information Criterion* (AIC) dalam regresi *penalized spline*.
3. Mengaplikasikan model *penalized spline* pada data Indeks Harga Saham Gabungan (IHSG).

#### 1.5. Manfaat Penelitian

Manfaat yang diperoleh dalam penelitian ini antara lain sebagai berikut :

1. Bagi penulis  
Sebagai sarana tambahan wawasan, pengetahuan tentang *penalized spline* dan penerapan teori pada saham, khususnya Indeks Harga Saham Gabungan (IHSG), serta dapat mengaplikasikan ilmu yang telah diperoleh selama di perkuliahan.
2. Bagi pembaca  
Diharapkan dapat menambah wawasan dan sebagai bahan referensi dalam penelitian selanjutnya, serta dapat mengembangkannya lebih luas.
3. Bagi Fakultas Sains dan Teknologi  
Diharapkan dapat menjadi bahan tinjauan pustaka atau referensi yang bermanfaat tentang *penalized spline*.

#### 1.6. Tinjauan Pustaka

Dalam subbab ini meliputi kajian pustaka dari penelitian sebelumnya, yang meliputi buku, skripsi, jurnal matematika, maupun sumber lainnya yang berkaitan

dengan objek pembahasan.

Penelitian Susanti & Adji (2020) membahas mengenai *Analisis Peramalan IHSG dengan Time Series Moseling ARIMA*. Penelitian ini bertujuan untuk menguji model *time series Autoregressive Integrated Moving Average* (ARIMA) yang digunakan untuk memprediksi IHSG. Data yang digunakan merupakan data IHSG harian tanggal 2 Januari 2017 sampai dengan 3 Januari 2018. Hasil dari penelitian yaitu didapatkan model terbaik ARIMA(7,3,1) yang dilihat dari nilai AIC dan *Schwarz Criterion* (SC) minimum dan ketepatan model didasarkan pada nilai *Root Means Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *Mean Percentage Error* (MAPE).

Penelitian Rezaldi & Sugiman (2021) mengenai *Peramalan Metode ARIMA Data Saham PT Telekomunikasi Indonesia* bertujuan untuk mengetahui model yang tepat dengan metode ARIMA untuk meramalkan harga saham PT Telekomunikasi Indonesia selanjutnya. Data yang digunakan yaitu data *close* saham PT Telekomunikasi Indonesia pada bulan Juni 2020 sampai dengan Mei 2021. Diperoleh hasil penelitian bahwa model terbaik yaitu model ARIMA(0,2,1) dengan kriteria *Mean Squared Error* (MSE) minimum.

Penelitian Agustina et al. (2015) membahas mengenai *Pemodelan Data Indeks Harga Saham Gabungan Menggunakan Regresi Penalized Spline*. Pada penelitian ini menggunakan nilai *Generalized Cross Validation* (GCV) minimum untuk menentukan parameter penghalus  $\lambda$  optimal, sedangkan untuk ketepatan atau keakuratan model menggunakan *Mean Absolute Percentage Error* (MAPE) dan nilai koefisien determinasi ( $R^2$ ).

Dalam penelitian Kurniasari et al. (2019) membahas mengenai *Estimasi Parameter Regresi Spline dengan Metode Penalized Spline*. Tujuan daripada penelitian tersebut yaitu untuk mengestimasi parameter dan data yang digunakan tidak memiliki pola tertentu. Data penelitian yang digunakan yaitu data sekunder Badan Pusat Statistik (BPS) pada tahun 2015, diantaranya indeks pembangunan manusia, gini rasio, harapan lama sekolah, penduduk miskin, dan kepadatan penduduk. Pa-

da penelitian ini menggunakan metode Generalized Cross Validation (GCV) untuk menentukan parameter smoothing ( $\lambda$ ), algoritma Fullsearch untuk menentukan titik knot optimal, koefisien determinasi ( $R^2$ ), dan *Mean Absolute Percentage Error* (MAPE) sebagai metode untuk memilih model terbaik. Pengolahan data dibantu dengan program R.

Penelitian oleh Siregar & Jatmiko (2019) membahas mengenai *Perbandingan Regresi B-Splines dan P-Splines pada Hubungan Indeks Pembangunan Manusia dan Persentase Penduduk Miskin Kabupaten/Kota di Indonesia Pada Tahun 2017*. Hasil analisis pada penelitian ini menunjukkan bahwa sebaran data dengan *fitting curve* menghasilkan kurva yang halus menjangkau seluruh sebaran data. Perbandingan MSE untuk *B-Spline* dan *P-Spline* diperoleh model *B-Spline* dengan MSE terkecil yaitu 0.1928378.

Penelitian Zia et al. (2017) membahas mengenai *Pemodelan Regresi Spline Menggunakan Metode Penalized Spline Pada Data Longitudinal (Studi Kasus : Harga Penutupan Saham LQ45 Sektor Keuangan dengan Kurs USD terhadap Rupiah Periode Januari 2011-Januari 2016)*. Data yang digunakan adalah data sekunder yang diambil dari saham BRI, BCA, Mandiri dan sebanyak 61 data pengamatan in sample. Pada penelitian ini untuk menentukan parameter smoothing ( $\lambda$ ) dan titik knot menggunakan metode *Generalized Cross Validation* (GCV), sedangkan untuk memilih model terbaik menggunakan koefisien determinasi ( $R^2$ ) dan *Mean Absolute Percentage Error* (MAPE). Pengolahan dan analisis data menggunakan program R.

### **1.7. Sistematika Penulisan**

Untuk mempermudah dalam pemahaman penulisan pada penelitian ini, maka penulis akan menggambarkan sistematika penulisan sebagai berikut:

#### **BAB I : PENDAHULUAN**

Pada bab ini akan dibahas tentang latar belakang penelitian, rumusan masalah, batasan penelitian, tujuan penelitian, manfaat penelitian, tinjauan pustaka, dan siste-



matika penulisan.

## BAB II : DASAR TEORI

Pada bab ini menjelaskan tentang teori-teori yang menjadi acuan atau penunjang, berisi tentang landasan teori dan hipotesis

## BAB III : METODOLOGI PENELITIAN

Pada bab ini membahas tentang metode penelitian yang digunakan dalam memecahkan masalah, dimana berisi tentang jenis penelitian, metode pengumpulan data, variabel dan sumber data, *software* yang digunakan, dan metode analisis data.

## BAB IV : HASIL DAN PEMBAHASAN

Pada bab ini berisi tentang hasil penelitian yang dapat diambil dari pembahasan permasalahan yang ada.

## BAB V : KESIMPULAN DAN SARAN

Pada bab ini berisi tentang kesimpulan yang dapat diambil dari pembahasan pada bab-bab sebelumnya dan saran-saran yang berkaitan dengan penulisan sejenis dimasa yang akan datang.

## BAB V

### PENUTUP

Pada bab ini akan diberikan kesimpulan dan saran-saran yang dapat diambil berdasarkan materi-materi yang telah dibahas pada bab-bab sebelumnya.

#### 5.1. Kesimpulan

Kesimpulan yang dapat diambil penulis setelah menyelesaikan penelitian ini adalah :

1. *Penalized spline* merupakan salah satu regresi nonparametrik yang menggabungkan regresi *B-Spline* dengan penalti kekasaran diskrit pada koefisien yang berdekatan dengan *B-Spline* yang bertujuan untuk memperhalus kurva. Penalti pada *penalized spline* memberikan kontrol tambahan, kontinu terhadap kehalusan. *Penalized spline* menjadi alternatif regresi *B-Spline* yang sensitif terhadap knot. *Penalized spline* memperkenalkan penalti pada kekasaran saat menggunakan *B-Spline* dengan jumlah knot yang besar. Fungsi objektif *penalized spline* dapat dituliskan sebagai :

$$Q = \sum_{i=1}^n \left( y_i - \sum_{j=1}^{m+k} \beta_j B_j(x) \right)^2 + \lambda \sum_{j=m+1}^{m+k} (\Delta^m \beta_j)^2 \quad (5.1)$$

dimana  $\Delta^m$  adalah *difference operator* (membentuk perbedaan orde) pada koefisien *B-spline* yang berdekatan dan  $\lambda$  adalah berat penalty. Perbedaan orde ( $\Delta^m$ ) dapat direpresentasikan dalam bentuk matriks **D**. Persamaan (5.1) dapat ditulis dalam bentuk matriks, yaitu :

$$\mathbf{Q} = (\mathbf{Y} - \beta\mathbf{B})^T(\mathbf{Y} - \beta\mathbf{B}) + \lambda\beta^T\mathbf{P}\beta \quad (5.2)$$

Sedangkan untuk mencari estimasi parameter  $\beta$  dapat dituliskan sebagai :

$$\hat{\beta} = (\mathbf{B}^T\mathbf{B} + \lambda\mathbf{P})^{-1}\mathbf{B}^T\mathbf{Y} \quad (5.3)$$

dimana  $\mathbf{B}$  adalah matriks yang terdiri dari elemen basis *B-spline*, fungsi *B-spline* ke- $j$  dan  $\mathbf{P}$  adalah jumlah kuadrat selisih,  $\mathbf{P} = \mathbf{D}^T \mathbf{D}$ .

2. Kombinasi orde, jumlah knot, dan  $\lambda$  optimal untuk model GCV dan AIC diperoleh pada orde 2 dengan 1 knot yaitu 6829.922 dengan masing-masing  $\lambda$  sebesar  $2.828283 \times 10^{-4}$  pada nilai GCV minimum sebesar 15600.36, dan  $\lambda$  sebesar  $1 \times 10^{-16}$  pada nilai AIC minimum sebesar 1299.176.
3. Diperoleh model regresi *penalized spline* terbaik menggunakan metode GCV yaitu :

$$\hat{y} = 5813.572B_{-1,2}(x) + 6858.177B_{0,2}(x) + 7219.961B_{1,2}(x) \quad (5.4)$$

Sedangkan untuk model regresi *penalized spline* terbaik menggunakan metode AIC yaitu :

$$\hat{y} = 5813.654B_{-1,2}(x) + 6858.172B_{0,2}(x) + 7220.146B_{1,2}(x) \quad (5.5)$$

Pada model GCV diperoleh nilai MAPE sebesar 1.3169% dan model AIC diperoleh nilai MAPE sebesar 1.31688%. Dikarenakan model AIC memiliki nilai MAPE lebih kecil daripada model GCV, maka model AIC dinilai lebih baik dalam menguji ketepatan model untuk peramalan. Berdasarkan kriteria nilai MAPE, nilai MAPE yang diperoleh menunjukkan kinerja model yang sangat akurat sehingga model sangat layak digunakan untuk melakukan pemodelan data Indeks Harga Saham Gabungan (IHSG).

## 5.2. Saran

Setelah membahas dan mengimplementasikan regresi *Penalized Spline* pada data IHSG penulis ingin menyampaikan beberapa saran :

1. Penggunaan orde, jumlah knot, dan  $\lambda$  pada penelitian ini yang terbatas, yaitu menggunakan orde 2 sampai 4 dengan jumlah knot sebanyak 1 sampai 3, dan batas  $\lambda$  yang digunakan antara  $1 \times 10^{-16}$  sampai  $1 \times 10^{-3}$ , maka penulis berharap untuk penelitian selanjutnya dapat dikembangkan lagi dengan penambahan variasi orde, jumlah knot, dan  $\lambda$  yang digunakan.

2. Metode yang digunakan dalam optimasi regresi *penalized spline* dalam penelitian ini yaitu menggunakan metode *Generalized Cross Validation* (GCV) dan *Akaike Information Criterion* (AIC). Di luar kedua metode optimasi tersebut, masih banyak metode yang dapat digunakan, maka penulis berharap untuk penelitian selanjutnya dapat memvariasi optimasi yang digunakan dalam pembentukan model regresi *penalized spline* terbaik.



## DAFTAR PUSTAKA

- Agustina, N., Suparti, S., & Mukid, M. A. (2015). Pemodelan data indeks harga saham gabungan menggunakan regresi penalized spline. *Jurnal Gaussian*, 4(3):603–612.
- Anoraga, P. & Pakarti, P. (2006). *Pengantar Pasar Modal*. Rineka Cipta.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time Series Analysis Forecasting and Control Fifth Edition*. John Wiley and Sons.
- Burnham, K. P. & Anderson, D. R. (2004). Multimodel inference: understanding aic and bic in model selection. *Sociological methods & research*, 33(2):261–304.
- Eilers, P. H. & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, 11(2):89–121.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing (2nd ed.)*. Marcel Dekker Inc.
- Finance, Y. (2022). Idx composite. <https://finance.yahoo.com/quote/%5EJKSE?p=%5EJKSE&.tsrc=fin-srch>. Diakses pada tanggal 12 Desember pukul 10.35 WIB.
- Keuangan, O. J. (2022). Apa itu ihsg?yuk kenali indeks saham ihsg. <https://sikapiuangmu.ojk.go.id/>. Diakses pada tanggal 3 Maret 2023 pukul 19.35 WIB.
- Kurniasari, W., Kusnandar, D., & Sulistianingsih, E. (2019). Estimasi parameter regresi spline dengan metode penalized spline. *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, 8(2).
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting Second Edition*. John Wiley and Sons.

- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction To Linear Regression Analysis Fifth Edition*. John Wiley and Sons.
- Priadana, M. S. & Sunarsi, D. (2021). *Metode Penelitian Kuantitatif*. Pascal Books.
- Qudratullah, M. F. (2013). *Analisis Regresi Terapan: Teori, Contoh Kasus, dan Aplikasi dengan SPSS*. Yogyakarta: Andi.
- Rezaldi, D. A. & Sugiman, S. (2021). Peramalan metode arima data saham pt. telekomunikasi indonesia. In *PRISMA, Prosiding Seminar Nasional Matematika*, volume 4, pages 611–620.
- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Number 12. Cambridge university press.
- Siregar, R. S. K. & Jatmiko, Y. A. (2019). Perbandingan regresi b-splines dan p-splines pada hubungan indeks pembangunan manusia dan persentase penduduk miskin kabupaten/kota di indonesia: Comparison of b-splines and p-splines regression on the relationship between human development index and the percentage of poor districts/cities in indonesia. *Emasains: Jurnal Edukasi Matematika dan Sains*, 8(1):101–111.
- Situmorang, S. H., Muda, I., Doli, M., & Fadli, F. S. (2010). *Analisis data untuk riset manajemen dan bisnis*. USUpres.
- Supandi, E. D. (2020). *Statistika dan Terapannya*. UIN SUKA UNERSITY PRESS.
- Susanti, R. & Adji, A. R. (2020). Analisis peramalan ihsg dengan time series modeling arima. *Jurnal Manajemen Kewirausahaan*, 17(1):97–106.
- Zia, N. G., Suparti, S., & Safitri, D. (2017). Pemodelan regresi spline menggunakan metode penalized spline pada data longitudinal (studi kasus: Harga penutupan saham Iq45 sektor keuangan dengan kurs usd terhadap rupiah periode januari 2011-januari 2016). *Jurnal Gaussian*, 6(2):221–230.