

**PEMODELAN REGRESI NONPARAMETRIK DATA LONGITUDINAL  
MENGUNAKAN POLINOMIAL LOKAL**

**(Studi Kasus: Harga Penutupan Saham pada Kelompok Harga Saham  
Periode Januari 2012 – April 2015)**



**SKRIPSI**

**Disusun oleh :**

**IZZUDDIN KHALID**

**24010211140098**

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG**

**2015**

**PEMODELAN REGRESI NONPARAMETRIK DATA LONGITUDINAL  
MENGUNAKAN POLINOMIAL LOKAL**

**(Studi Kasus: Harga Penutupan Saham pada Kelompok Harga Saham  
Periode Januari 2012 – April 2015)**

**Oleh :**

**IZZUDDIN KHALID**

**24010211140098**

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar**

**Sarjana Sains pada Jurusan Statistika**

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG**

**2015**

## HALAMAN PENGESAHAN I

Judul : **Pemodelan Regresi Nonparametrik Data Longitudinal Menggunakan Polinomial Lokal (Studi Kasus: Harga Penutupan Saham pada Kelompok Harga Saham Periode Januari 2012 – April 2015)**

Nama : Izzuddin Khalid

NIM : 24010211140098

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 22 Juni 2015 dan dinyatakan lulus pada tanggal 25 Juni 2015

Semarang, 25 Juni 2015

Panitia Penguji Ujian Tugas Akhir  
Ketua,



Yuciana Wilandari, S.Si, M.Si.  
NIP. 197005191998022001



Mengetahui,  
Ketua Jurusan Statistika  
Fakultas Sains dan Matematika

Dra. Dwi Ispriyanti, M.Si  
NIP. 195709141986032001

## HALAMAN PENGESAHAN II

Judul : **Pemodelan Regresi Nonparametrik Data Longitudinal Menggunakan Polinomial Lokal (Studi Kasus: Harga Penutupan Saham pada Kelompok Harga Saham Periode Januari 2012 – April 2015)**

Nama : Izzuddin Khalid

NIM : 24010211140098

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 22 Juni 2015

Semarang, 25 Juni 2015

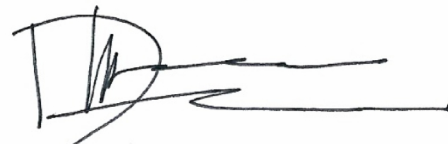
Pembimbing I



Dra. Suparti, M.Si.

NIP. 196509131990032001

Pembimbing II



Alan Prahutama, S.Si, M.Si.

NIP. 198804212014041002

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penulisan Tugas Akhir ini dengan judul “Pemodelan Regresi Nonparametrik Data Longitudinal Menggunakan Polinomial Lokal”.

Penulis menyadari tanpa bantuan dan dukungan dari berbagai pihak, tugas akhir ini tidak dapat diselesaikan dengan baik. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Ibu Dra. Dwi Ispriyanti, M.Si selaku Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Ibu Dra. Suparti, M.Si dan Bapak Alan Prahutama, S.Si, M.Si selaku dosen pembimbing I dan II yang telah memberikan bimbingan, arahan, dan petunjuk hingga tugas akhir ini selesai.
3. Bapak dan Ibu dosen Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
4. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu dalam penulisan tugas akhir ini.

Penulis menyadari bahwa penulisan tugas akhir ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang bersifat membangun dan semoga tugas akhir ini dapat bermanfaat bagi semua pihak.

Semarang, Juni 2015

Penulis

## ABSTRAK

Saham merupakan surat berharga yang dapat dibeli atau dijual oleh perorangan atau lembaga sebagai tanda penyertaan atau kepemilikan ke dalam suatu perusahaan sebesar proporsi saham yang dimilikinya. Dilihat dari nilai kapitalisasi pasar, saham dibagi menjadi 3 kelompok yaitu kapitalisasi besar (*Big-Cap*), kapitalisasi sedang (*Mid-Cap*), dan kapitalisasi kecil (*Small-Cap*). Data longitudinal merupakan pengamatan yang dilakukan sebanyak  $n$  subjek yang saling independen dengan setiap subjek diamati secara berulang dalam kurun waktu berbeda yang saling dependen. Teknik *smoothing* yang digunakan untuk mengestimasi model regresi nonparametrik pada data longitudinal adalah estimator polinomial lokal. Estimator polinomial lokal dapat diperoleh dengan metode WLS (*Weighted Least Square*). Estimator polinomial lokal sangat bergantung pada *bandwidth* optimal. Penentuan *bandwidth* optimal dapat diperoleh dengan menggunakan metode GCV (*Generalized Cross Validation*). Diantara kernel Gaussian, kernel Segitiga, kernel Epanechnikov dan kernel *Biweight* diperoleh model terbaik dengan menggunakan kernel Gaussian. Berdasarkan penerapan model secara simultan diperoleh nilai koefisien determinasi sebesar 97,80174% dan nilai MSE sebesar 0,03053464. Dengan kernel Gaussian diperoleh nilai MAPE data *out sample* sebesar 11,74493%.

Kata Kunci: Data Longitudinal, Polinomial Lokal, Saham

## ABSTRACT

Stocks are securities that can be bought or sold by individuals or institutions as a sign of participating or possessing a company in the amount of its proportions. From the lens of market capitalization values, stocks are divided into 3 groups: large capitalization (Big-Cap), medium capitalization (Mid-Cap) and small capitalization (Small-Cap). Longitudinal data is observation which is conducted as n subjects that are independent to each subject observed repeatedly in different periods dependently. Smoothing technique used to estimate the nonparametric regression model in longitudinal data is local polynomial estimator. Local polynomial estimator can be obtained by WLS (Weighted Least Square) methods. Local polynomial estimator is very dependent on optimal bandwidth. Determination of the optimal bandwidth can be obtained by using GCV (Generalized Cross Validation) method. Among the Gaussian kernel, Triangle kernel, Epanechnikov kernel and Biweight kernel, it is obtained the best model using Gaussian kernel. Based on the application of the model simultaneously, it is obtained coefficient of determination of 97,80174% and MSE values of 0,03053464. Using Gaussian kernel, MAPE out sample of data is obtained as 11,74493%.

Keywords: Longitudinal Data, Local Polynomial, Stocks

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN I</b> .....	ii
<b>HALAMAN PENGESAHAN II</b> .....	iii
<b>KATA PENGANTAR</b> .....	iv
<b>ABSTRAK</b> .....	v
<b>ABSTRACT</b> .....	vi
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR GAMBAR</b> .....	x
<b>DAFTAR TABEL</b> .....	xi
<b>DAFTAR LAMPIRAN</b> .....	xii
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	4
1.4 Tujuan Penelitian .....	4
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Saham .....	5
2.2 Data Longitudinal .....	6
2.3 Estimator Kernel .....	7
2.4 Polinomial Lokal pada Data Longitudinal .....	9
2.5 Pemilihan <i>Bandwidth</i> .....	12
2.6 Koefisien Determinasi .....	13

2.7	Kriteria Pemilihan Model terbaik.....	14
2.8	Ketepatan Metode Peramalan.....	14
<b>BAB III METODOLOGI PENELITIAN</b>		
3.1	Sumber Data .....	16
3.2	Variabel Penelitian .....	16
3.3	Tahapan Analisis .....	16
3.4	Diagram Alir Penelitian ( <i>Flowchart</i> ) .....	18
<b>BAB IV HASIL DAN PEMBAHASAN</b>		
4.1	Deskripsi Data .....	19
4.2	Regresi Polinomial Lokal .....	22
4.2.1	Pemodelan Polinomial Lokal Menggunakan Kernel	
	Gaussian .....	22
4.2.1.1	Pemilihan pada Perusahaan Astra Otoparts Tbk .....	22
4.2.1.2	Pemilihan pada Perusahaan Astra Graphia Tbk .....	24
4.2.1.3	Pemilihan pada Perusahaan Mahaka Media Tbk.....	25
4.2.1.4	Pemodelan Polinomial Lokal Data Longitudinal .....	27
4.2.2	Pemodelan Polinomial Lokal Menggunakan Kernel	
	Segitiga.....	28
4.2.2.1	Pemilihan pada Perusahaan Astra Otoparts Tbk .....	28
4.2.2.2	Pemilihan pada Perusahaan Astra Graphia Tbk .....	30
4.2.2.3	Pemilihan pada Perusahaan Mahaka Media Tbk.....	31
4.2.2.4	Pemodelan Polinomial Lokal Data Longitudinal .....	33
4.2.3	Pemodelan Polinomial Lokal Menggunakan Kernel	
	Epanechnikov .....	34

4.2.3.1	Pemilihan pada Perusahaan Astra Otoparts Tbk .....	34
4.2.3.2	Pemilihan pada Perusahaan Astra Graphia Tbk .....	35
4.2.3.3	Pemilihan pada Perusahaan Mahaka Media Tbk.....	37
4.2.3.4	Pemodelan Polinomial Lokal Data Longitudinal .....	38
4.2.4	Pemodelan Polinomial Lokal Menggunakan Kernel	
	<i>Biweight</i> .....	40
4.2.4.1	Pemilihan pada Perusahaan Astra Otoparts Tbk .....	40
4.2.4.2	Pemilihan pada Perusahaan Astra Graphia Tbk .....	41
4.2.4.3	Pemilihan pada Perusahaan Mahaka Media Tbk.....	43
4.2.4.4	Pemodelan Polinomial Lokal Data Longitudinal .....	44
4.3	Pemilihan Model Polinomial Lokal Terbaik .....	46
4.4	Menghitung Ketepatan Peramalan Model Polinomial Lokal	
	Terbaik .....	47
<b>BAB V</b>	<b>KESIMPULAN</b> .....	49
	<b>DAFTAR PUSTAKA</b> .....	50
	<b>LAMPIRAN</b> .....	52

## DAFTAR GAMBAR

	Halaman
Gambar 1 Plot 4 Fungsi Kernel.....	9
Gambar 2 Diagram Alir Penelitian .....	18
Gambar 3 Plot Harga Penutupan Saham Bulanan .....	21
Gambar 4 Kurva Estimasi Kernel Gaussian Subjek 1 .....	23
Gambar 5 Kurva Estimasi Kernel Gaussian Subjek 2 .....	25
Gambar 6 Kurva Estimasi Kernel Gaussian Subjek 3 .....	26
Gambar 7 Kurva Estimasi Kernel Segitiga Subjek 1 .....	29
Gambar 8 Kurva Estimasi Kernel Segitiga Subjek 2 .....	31
Gambar 9 Kurva Estimasi Kernel Segitiga Subjek 3 .....	32
Gambar 10 Kurva Estimasi Kernel Epanechnikov Subjek 1 .....	35
Gambar 11 Kurva Estimasi Kernel Epanechnikov Subjek 2 .....	36
Gambar 12 Kurva Estimasi Kernel Epanechnikov Subjek 3 .....	38
Gambar 13 Kurva Estimasi Kernel <i>Biweight</i> Subjek 1 .....	41
Gambar 14 Kurva Estimasi Kernel <i>Biweight</i> Subjek 2.....	42
Gambar 15 Kurva Estimasi Kernel <i>Biweight</i> Subjek 3.....	44

## DAFTAR TABEL

	Halaman
Tabel 1 Deskripsi Harga Penutupan Saham Bulanan .....	20
Tabel 2 Kombinasi Optimal Kernel Gaussian Subjek 1 .....	23
Tabel 3 Kombinasi Optimal Kernel Gaussian Subjek 2 .....	24
Tabel 4 Kombinasi Optimal Kernel Gaussian Subjek 3 .....	26
Tabel 5 Nilai Optimal Kernel Gaussian pada Masing-Masing Subjek .....	27
Tabel 6 Estimasi Model Menggunakan Kernel Gaussian .....	27
Tabel 7 Kombinasi Optimal Kernel Segitiga Subjek 1 .....	29
Tabel 8 Kombinasi Optimal Kernel Segitiga Subjek 2 .....	30
Tabel 9 Kombinasi Optimal Kernel Segitiga Subjek 3 .....	32
Tabel 10 Nilai Optimal Kernel Segitiga pada Masing-Masing Subjek .....	33
Tabel 11 Estimasi Model Menggunakan Kernel Segitiga .....	33
Tabel 12 Kombinasi Optimal Kernel Epanechnikov Subjek 1 .....	34
Tabel 13 Kombinasi Optimal Kernel Epanechnikov Subjek 2 .....	36
Tabel 14 Kombinasi Optimal Kernel Epanechnikov Subjek 3 .....	37
Tabel 15 Nilai Optimal Kernel Epanechnikov pada Masing-Masing Subjek .....	39
Tabel 16 Estimasi Model Menggunakan Kernel Epanechnikov .....	39
Tabel 17 Kombinasi Optimal Kernel <i>Biweight</i> Subjek 1 .....	40
Tabel 18 Kombinasi Optimal Kernel <i>Biweight</i> Subjek 2 .....	42
Tabel 19 Kombinasi Optimal Kernel <i>Biweight</i> Subjek 3 .....	43
Tabel 20 Nilai Optimal Kernel <i>Biweight</i> pada Masing-Masing Subjek ...	45
Tabel 21 Estimasi Model Menggunakan Kernel <i>Biweight</i> .....	45
Tabel 22 Perbandingan Koefisien Determinasi dan GCV .....	46
Tabel 23 Hasil Estimasi Data <i>out sample</i> .....	48

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data <i>in sample</i> Harga Penutupan Saham Bulanan .....	52
Lampiran 2. Data <i>out sample</i> Harga Penutupan Saham Bulanan .....	53
Lampiran 3. Program Pemilihan Kombinasi optimal Kernel Gaussian .....	54
Lampiran 4. Program Grafik Kombinasi Optimal Kernel Gaussian.....	56
Lampiran 5. Program Estimasi Model Polinomial Lokal Kernel Gaussian..	58
Lampiran 6. Output <i>Bandwidth</i> , Titik Waktu dan Order Polinomial Kernel Gaussian .....	60
Lampiran 7. Output Estimasi Model Polinomial Lokal Kernel Gaussian ....	64
Lampiran 8. Output <i>Bandwidth</i> , Titik Waktu dan Order Polinomial Kernel Segitiga .....	65
Lampiran 9. Output Estimasi Model Polinomial Lokal Kernel Segitiga .....	69
Lampiran 10. Output <i>Bandwidth</i> , Titik Waktu dan Order Polinomial Kernel Epanechnikov .....	70
Lampiran 11. Output Estimasi Model Polinomial Lokal Kernel Epanechnikov .....	74
Lampiran 12. Output <i>Bandwidth</i> , Titik Waktu dan Order Polinomial Kernel <i>Biweight</i> .....	75
Lampiran 13. Output Estimasi Model Polinomial Lokal Kernel <i>Biweight</i> ....	79
Lampiran 14. Program Perhitungan Keakuratan Model .....	80
Lampiran 15. Output Perhitungan Keakuratan Model.....	81

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Analisis regresi merupakan salah satu teknik yang paling umum digunakan dalam statistik. Tujuan dari analisis ini adalah untuk mengeksplorasi hubungan antara variabel prediktor dan respon, yaitu untuk menilai kontribusi variabel prediktor dan untuk mengidentifikasi dampaknya terhadap variabel respon (Fan dan Gijbels, 1997). Terdapat tiga pendekatan regresi yang dapat digunakan yaitu pendekatan parametrik, pendekatan nonparametrik, dan pendekatan semiparametrik. Pendekatan secara parametrik memerlukan asumsi bahwa bentuk yang mendasari fungsi regresi diketahui, sehingga dibutuhkan informasi dari teori, pengalaman masa lalu atau sumber lain yang tersedia yang memberikan pengetahuan rinci tentang proses yang diteliti. Pada pendekatan nonparametrik tidak tergantung pada asumsi bentuk kurva tertentu sehingga memiliki fleksibilitas yang tinggi karena data diharapkan mencari sendiri bentuk kurva regresinya. Metode ini paling cocok untuk inferensi dalam situasi dimana tidak tersedia informasi sebelumnya tentang kurva regresi dari data (Eubank, 1999). Apabila diketahui hubungan variabel respon dengan salah satu variabel prediktor, tetapi dengan variabel prediktor yang lain tidak diketahui bentuk pola hubungannya maka dalam kondisi ini digunakan pendekatan secara semiparametrik (Ruppert, Wand dan Carrol, 2003). Pendekatan semiparametrik merupakan gabungan antara pendekatan parametrik dan pendekatan nonparametrik.

Data longitudinal merupakan pengamatan yang dilakukan sebanyak  $n$  subjek yang saling independen dengan setiap subjek diamati secara berulang dalam kurun waktu berbeda yang saling dependen (Wu dan Zhang, 2006). Menurut Weiss (2005) data longitudinal memiliki kelebihan yaitu lebih efisien dan hemat biaya, dapat digunakan untuk memprediksi, dan melihat perubahan dari individu pada periode waktu. Untuk memodelkan data longitudinal lebih tepat jika menggunakan pendekatan regresi nonparametrik karena lebih fleksibel dibandingkan dengan pendekatan parametrik (Wu dan Zang, 2006). Terdapat beberapa pendekatan regresi nonparametrik untuk mengestimasi kurva regresi, salah satunya yaitu dengan estimator polinomial lokal. Kelebihan dari polinomial lokal adalah kemampuannya dalam beradaptasi terhadap data yang artinya membagi data tersebut kedalam suatu wilayah kemudian melakukan estimasi terhadap wilayah yang sudah ditetapkan nilainya tersebut. Selain itu, polinomial lokal baik untuk menaksir data yang nilainya menyimpang jauh dibandingkan nilai data yang lain (Fan dan Gijbels, 1997).

Penelitian sebelumnya telah banyak dikembangkan tentang estimasi model regresi pada data longitudinal antara lain estimasi kurva regresi semiparametrik pada data longitudinal berdasarkan estimator polinomial lokal oleh Utami (2013), pemodelan trombosit penderita demam berdarah dengan pendekatan regresi nonparametrik pada data longitudinal berdasarkan estimator lokal linier oleh Tita, dkk. (2010), dan estimasi model regresi nonparametrik pada data longitudinal berdasarkan estimator polinomial lokal kernel GEE (*Generalized Estimating Equation*) oleh Utami (2010).

Kemajuan perekonomian suatu negara dapat ditandai dengan adanya pasar modal yang tumbuh dan berkembang dengan baik (Purnomo, 2013). Sebagai bentuk pasar, pasar modal merupakan wadah untuk mempertemukan antara penjual dan pembeli. Adapun penjual dan pembeli disini adalah penjualan dan pembelian instrumen keuangan dalam kerangka investasi (Hadi, 2013). Pasar modal memfasilitasi berbagai sarana dan prasarana kegiatan jual-beli surat-surat berharga dan kegiatan terkait lainnya. Umumnya surat berharga yang diperdagangkan dapat dibedakan menjadi surat berharga bersifat hutang dan surat berharga bersifat pemilikan. Surat berharga bersifat pemilikan dikenal dengan nama saham (Hadi, 2013). Saham merupakan surat berharga sebagai bukti penyertaan atau kepemilikan seseorang atau badan usaha dalam suatu perusahaan (Purnomo, 2013). Sejalan dengan pertumbuhan industri keuangan, saham juga mengalami pembagian saham ke dalam berbagai segi salah satunya berdasarkan nilai kapitalisasi pasar. Menurut Hadi (2013), saham dilihat dari nilai kapitalisasi pasar dibagi menjadi 3 kelompok yaitu kapitalisasi besar (*Big-Cap*), kapitalisasi sedang (*Mid-Cap*) dan kapitalisasi kecil (*Small-Cap*).

Berdasarkan uraian tersebut, penulis tertarik untuk membahas estimasi model regresi nonparametrik pada data longitudinal berdasarkan estimator Polinomial Lokal. Adapun data yang akan diestimasi yaitu data harga penutupan (*closing price*) saham beberapa perusahaan berdasarkan tingkat kapitalisasi pasar dengan mengambil satu subjek perusahaan pada tiap kelompok saham.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang tersebut, maka rumusan masalah dari penulisan ini yaitu cara menentukan model regresi nonparametrik terbaik pada data longitudinal

berdasarkan estimator Polinomial Lokal untuk data *closing price* saham bulan Januari 2012 sampai dengan September 2014.

### **1.3 Batasan Masalah**

Data yang digunakan pada penelitian ini adalah data harga penutupan (*closing price*) saham bulanan pada perusahaan Astra Otoparts Tbk., perusahaan Astra Graphia Tbk., perusahaan Mahaka Media Tbk. Adapun data yang digunakan mulai bulan Januari 2012 - September 2014 sebagai data *in sample* dan Oktober 2014 – April 2015 sebagai data *out sample*. Sedangkan fungsi kernel yang digunakan yaitu kernel Gaussian, kernel Segitiga, kernel Epanechnikov dan kernel *Biweight*.

### **1.4 Tujuan Penelitian**

Berdasarkan rumusan masalah yang telah dikemukakan, maka tujuan yang ingin dicapai dari penelitian ini yaitu dapat menentukan model regresi nonparametrik terbaik pada data longitudinal berdasarkan estimator Polinomial Lokal untuk data harga penutupan (*closing price*) saham bulan Januari 2012 sampai dengan September 2014.