



**PEMILIHAN MODEL REGRESI LINEAR BERGANDA TERBAIK PADA
KASUS MULTIKOLINEARITAS BERDASARKAN METODE *PRINCIPAL
COMPONENT ANALYSIS* (PCA) DAN METODE *STEPWISE***

Skripsi

Disusun sebagai salah satu syarat
untuk memperoleh gelar Sarjana Sains
Program Studi Matematika

oleh
Sri Pujilestari
UNNES
UNIVERSITAS NEGERI SEMARANG
4111412017

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG**

2016

PERNYATAAN

Saya menyatakan bahwa skripsi ini bebas plagiat, kecuali yang secara tertulis dirujuk dalam skripsi ini dan disebutkan dalam daftar pustaka. Apabila dikemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.



Semarang, Agustus 2016



Sri Fujilestari

4111412017

UNNES
UNIVERSITAS NEGERI SEMARANG

PENGESAHAN

Skripsi yang berjudul

Pemilihan Model Regresi Linier Berganda Terbaik pada Kasus
Multikolinieritas Berdasarkan Metode *Principal Component Analysis* (PCA)
dan Metode *Stepwise*.

disusun oleh

Sri Pujilestari

4111412017

Telah dipertahankan di hadapan sidang Panitia Ujian Skripsi FMIPA UNNES
pada tanggal 22 Agustus 2016.



Ketua
Prof. Dr. Zaenuri S.E., M.Si, Akt
196412231988031001

Sekretaris

Drs. Arief Agoestanto, M.Si
196807221993031005

Ketua/Penguji

Drs. Arief Agoestanto, M.Si
196807221993031005

Anggota Penguji /
Pembimbing II

Drs. Sugiman, M.Si
196401111989011001

Anggota Penguji/
Pembimbing I

Dr. Nurkaromah Dwidayati, M.Si
196605041990022001

MOTO DAN PERSEMBAHAN

MOTO

Life will always have a different plan for you. If you don't give up, you will eventually get to your destination. But towards the end of your life, you may look back and realize that it was never really about the destination; it was the journey that counted

~King Samuel Benson

The real happiness of life is to live modestly.

~ W.M.Thancheray

*Bermimpi adalah langkah pertama yang yang harus anda buat.
Sedangkan bertindak adalah langkah selanjutnya.*

PERSEMBAHAN

Karya ini saya persembahkan untuk:

- *Kedua orangtuaku, Abah Kamin dan Mama Suwatikha yang telah memberikan dukungan kasih sayang, doa dan semangatnya sehingga karya ini dapat terselesaikan tepat pada waktunya.*
- *Adikku tersayang Amalia Tyas Wulandari yang telah memotivasi, mendoakan dan memberikan semangat dalam penulisan karya ini.*
- *Teman-temanku Matematika 2012 yang aku sayangi yang telah menyemangati dan membantuku.*
- *Sahabatku yang telah memberikan semangatnya*

KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyusun dan menyelesaikan skripsi yang berjudul “Pemilihan Model Regresi Linier Berganda Terbaik pada Kasus Multikolinieritas Berdasarkan Metode *Principal Component Analysis* (PCA) dan Metode *Stepwise*”.

Skripsi ini disusun untuk menyelesaikan Studi S1 untuk meraih gelar Sarjana Sains. Penulis menyadari sepenuhnya bahwa skripsi ini selesai berkat bantuan, petunjuk, saran, bimbingan dan dorongan dari berbagai pihak. Oleh karena itu perkenankanlah penulis menyampaikan terima kasih kepada:

1. Prof. Dr. Fathur Rokhman, M.Hum., Rektor Universitas Negeri Semarang.
2. Prof. Dr. Zaenuri, S.E, M.Si,Akt., Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang
3. Drs. Arief Agoestanto, M.Si., Ketua Jurusan Matematika FMIPA Universitas Negeri Semarang.
4. Drs. Mashuri, M.Si., Ketua Prodi Matematika FMIPA Universitas Negeri Semarang.
5. Dr. Nurkaromah Dwidayati, M.Si., Dosen Pembimbing I yang telah dengan sabar memberikan bimbingan, petunjuk, motivasi dan saran yang sangat bermanfaat kepada penulis selama penyusunan skripsi ini.
6. Drs. Sugiman, M.Si., Dosen Pembimbing II yang telah dengan sabar memberikan bimbingan, petunjuk, motivasi dan saran yang sangat bermanfaat kepada penulis selama penyusunan skripsi ini.
7. Drs. Arief Agoestanto, M.Si., Dosen Penguji yang telah memberikan masukan dan saran yang membangun kepada penulis demi kesempurnaan penyusunan skripsi ini.

8. Prof. Dr. St. Budi Waluya, M.Si., Dosen Wali sekaligus sebagai inspirator dalam memberikan pencerahan dan dukungan untuk terus melangkah menyusun skripsi.
9. Seluruh Dosen Matematika yang telah membimbing dan memberikan ilmunya kepada penulis.
10. Keluargaku tercinta yang selalu memberi motivasi baik moral maupun material serta do'a restu dalam menyelesaikan skripsi ini.
11. Teman-teman Matematika angkatan 2012 yang telah memberikan semangat, motivasi dan bantuannya dalam menyelesaikan skripsi ini.
12. Semua pihak yang tidak dapat penulis sebutkan satu persatu.

Penulis berharap, semoga penelitian yang telah dilakukan ini bermanfaat bagi pembaca pada khususnya dan perkembangan pendidikan Indonesia pada umumnya.

Semarang, Agustus 2016

Penulis

UNNES
UNIVERSITAS NEGERI SEMARANG

ABSTRAK

Pujilestari, Sri. 2016. *Pemilihan Model Regresi Linier Berganda Terbaik pada Kasus Multikolinieritas Berdasarkan Metode Principal Component Analysis (PCA) dan Metode Stepwise.* Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Utama: Dr. Nurkaromah Dwidayati, M.Si dan Pembimbing Pendamping: Drs. Sugiman, M.Si.

Kata Kunci: Regresi Linier Berganda, Model Terbaik, Multikolinieritas, Metode *Principal Component Analysis* (PCA), Metode *Stepwise*

Analisis regresi merupakan salah satu metode statistik yang sering digunakan dalam kehidupan sehari-hari. Analisis regresi linier berganda digunakan untuk mengetahui sejauh mana hubungan sebuah variabel bebas dengan beberapa variabel tak bebas. Model regresi terbaik adalah model yang dapat menjelaskan perilaku peubah tak bebas dengan sebaik-baiknya dengan memilih peubah-peubah bebas dari sekian banyak peubah bebas yang tersedia dalam data dengan menggunakan kriteria pembandingan R^2 adjusted dan S^2 . Salah satu permasalahan asumsi pada model regresi linier berganda adalah seringnya terjadi korelasi antar variabel-variabel bebas pada model regresi linier berganda yang disebut sebagai multikolinieritas. Jika terdapat multikolinieritas maka kesimpulan yang dihasilkan tidak tepat. Pada penelitian ini metode yang digunakan untuk mencari model terbaik pada kasus multikolinieritas adalah metode *Stepwise* dan metode *Principal Component Analysis* (PCA). Tujuan utama dari penelitian ini yaitu mencari model terbaik dengan menggunakan metode *Stepwise* dan metode *Principal Component Analysis* (PCA).

Langkah pertama yang dilakukan yaitu, menguji asumsi regresi yaitu uji normalitas, uji kelinieran, uji keberartian, uji multikolinieritas, uji heterokedstisitas, dan uji autokorelasi. Kemudian mencari model terbaik dengan menggunakan metode *Stepwise* dan metode *Principal Component Analysis* (PCA) dengan pembandingan R^2 adjusted dan S^2 pada data *return* saham perusahaan dalam Indeks LQ 45 di BEI periode Juli – Desember 2015.

Berdasarkan penelitian hasil yang diperoleh untuk mencari model terbaik pada kasus multikolinieritas data *return* saham perusahaan dalam Indeks LQ 45 di BEI periode Juli – Desember 2015 dengan menggunakan metode *Principal Component Analysis* (PCA) $\hat{Y} = 11,992 + 2,179 F_1$ dengan nilai R^2 adjusted sebesar 0,050 dan nilai S^2 sebesar 63,049 sedangkan hasil yang diperoleh dengan metode *Stepwise* yaitu $\hat{Y} = 4,891 + 7,804X_6 + 0,144X_2$ dengan nilai R^2 adjusted sebesar 0,191 dan nilai S^2 sebesar 53,678. Dari hasil tersebut dapat disimpulkan bahwa metode *Stepwise* lebih cocok untuk mencari model terbaik pada kasus multikolinieritas.

DAFTAR ISI

Halaman

HALAMAN JUDUL.....	i
PERNYATAAN KEASLIAN TULISAN	ii
HALAMAN PENGESAHAN.....	iii
MOTTO DAN PERSEMBAHAN	iv
KATA PENGANTAR	v
ABSTRAK	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiv
DAFTAR SIMBOL.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	6
1.3 Batasan Masalah	6
1.4 Rumusan Masalah	7
1.5 Tujuan Penelitian	7
1.6 Manfaat Penelitian	8
1.7 Penegasan Masalah	8
1.8 Sistematika Penulisan Skripsi	9
BAB II TINJAUAN PUSTAKA.....	11

2.1 Matriks	11
2.2 Regresi	13
2.3 Asumsi Klasik	15
2.4 Multikolinieritas	26
2.5 Program SPSS.16 for Windows	28
2.6 Metode <i>Principal Component Analysis</i> (PCA)	32
2.7 Metode <i>Stepwise</i>	35
2.8 Pendeteksi Kelayakan Model Regresi Terbaik	39
2.9 Indeks LQ 45	41
2.10 Definisi Variabel	43
2.11 Penelitian Terdahulu	46
2.12 Kerangka Berpikir	50
BAB III METODE PENELITIAN	54
3.1 Penemuan Masalah	54
3.2 Kajian Pustaka	54
3.3 Analisis dan Pemecahan Masalah	55
3.4 Penarikan Kesimpulan	65
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	66
4.1 Hasil Penelitian	66
4.2 Pembahasan	85
BAB V PENUTUP	88
5.1 Simpulan	88
5.2 Saran	89

DAFTAR PUSTAKA	90
LAMPIRAN.....	93



DAFTAR TABEL

Tabel

Halaman

2.1	Kriteria Pengujian Autokorelasi dengan Durbin-Watson	25
2.2	Penelitian Terdahulu	49
4.1	VIF Varibel Bebas pada Data <i>Return</i> Saham	67
4.2	Hasil Uji Kolmogorov-Smirnov Data <i>Return</i> saham	68
4.3	Hasil Uji Linieriras	69
4.4	Hasil Uji Keberartian Simultan	70
4.5	Hasil Uji Multikolinearitas	71
4.6	Hasil Uji Heterokesdastisitas	72
4.7	Hasil Uji Autokorelasi	72
4.8	KMO dan <i>Barlett Test</i>	74
4.9	<i>Anti Image Matriks</i> (MSA)	74
4.10	KMO dan <i>Barlett Test</i> setelah reduksi	76
4.11	<i>Anti Image Matriks</i> (MSA) setelah reduksi	77
4.12	<i>Communalities</i>	77
4.13	<i>Total Variance Explained</i>	78
4.14a	<i>Component Matriks</i>	79
4.14b	<i>Component Score Coefficient Matriks</i>	80
4.15	<i>Coeffiscient Metode PCA</i>	81
4.16	<i>Variabel Ebtered.Removed</i>	82

4.17 ANOVA Metode <i>Stepwise</i>	83
4.18 <i>Coefficient</i> Metode <i>Stepwise</i>	84
4.19 Hasil Perhitungan Nilai R^2 adjusted dan S^2	84



DAFTAR GAMBAR

Gambar

Halaman

2.1	Kerangka Berpikir	53
3.1	Diagram Alir langkah-langkah Penyelesaian Masalah.....	64



DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data <i>Return</i> Saham LQ 45	93
2. Output Uji Multikolinearitas	96
3. Output Tes Normalitas	97
4. Output Uji Kelinieran.....	98
5. Output Uji Keberartian Simultan	99
6. Output Uji Heterokedastisitas.....	100
7. Output Uji Autokorelasi.....	101
8. Output KMO dan <i>Barlett Test</i>	102
9. Output <i>Anti Image Matriks</i> (MSA)	103
10. Output KMO dan <i>Barlett Test</i> setelah Reduksi.....	104
11. Output <i>Anti Image Matriks</i> (MSA) setelah Reduksi	105
12. Output <i>Communalities</i>	106
13. Ouput <i>Total Variance Explained</i>	107
14. Output <i>Component Matriks</i> dan <i>Component Score Coefficient</i> <i>Matriks</i>	108
15. Ouput <i>Coefficient</i> Metode PCA	109
16. Ouput <i>Variabel Entered/Removed</i>	110
17. Ouput ANOVA	111
18. Ouput <i>Coefficient</i> Metode <i>Stepwise</i>	112
19. Ouput <i>Model Summary</i> dan ANOVA Metode <i>Stepwise</i>	113
20. Ouput <i>Model Summary</i> dan ANOVA Metode PCA	114

DAFTAR SIMBOL

Y	: Pengamatan ke-i variabel tak bebas
X_{ki}	: Pengamatan ke-i pada variabel bebas
β_0	: Parameter intercept (konstanta)
$\beta_1, \beta_2, \dots, \beta_k$: Parameter koefisien regresi variabel tak bebas
ε	: error (galat kesalahan)
k	: Banyaknya variabel bebas
R	: Korelasi
b_j	: Koefisien regresi
Sb_j	: Kesalahan baku koefisien regresi.
DW	: Nilai Durbin-Watson Test
e	: Nilai residual
e_{t-1}	: Nilai residual satu baris/periode sebelumnya
λ	: Nilai eigen

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan salah satu metode statistik yang sering digunakan dalam kehidupan sehari-hari. Tujuan analisis regresi yaitu untuk mengetahui sejauh mana hubungan sebuah variabel bebas dengan beberapa variabel tak bebas. Bila dalam analisisnya hanya melibatkan sebuah variabel bebas saja, maka analisis yang digunakan adalah analisis regresi linier sederhana. Sedangkan, bila analisisnya melibatkan lebih dari satu atau beberapa variabel bebas, maka analisis yang digunakan adalah analisis regresi linier berganda.

Dalam menentukan model regresi, variabel bebas dapat masuk dalam model secara bersama-sama atau satu persatu. Jika variabel bebas masuk dalam model secara bersama-sama maka perhitungan akan ringkas, akan tetapi tidak akan kelihatan apa yang terjadi dalam perhitungan tersebut karena setiap variabel bebas yang masuk memberikan pengaruh yang berbeda, tergantung pada urutan variabel bebas tersebut yang masuk dalam model. Namun tidak berarti semua variabel yang masuk dalam model regresi menjadikan model tersebut model yang terbaik (Sembiring, 1995).

Model regresi terbaik adalah model yang dapat menjelaskan perilaku peubah tak bebas dengan sebaik-baiknya dengan memilih peubah-peubah bebas dari sekian banyak peubah bebas yang tersedia dalam data. (Hanum. H, 2011).

Beberapa hal lain yang penting juga untuk dipahami dalam penggunaan analisis regresi linier ganda yaitu perlunya melakukan uji asumsi klasik atau uji persyaratan analisis regresi ganda sehingga persamaan garis regresi yang diperoleh benar-benar dapat digunakan untuk memprediksi variabel dependen atau kriterium. Uji persyaratan tersebut harus terpenuhi, apabila tidak maka akan menghasilkan garis regresi yang tidak cocok untuk memprediksi.

Sebelum membahas uji persyaratan perlu dipahami bahwa statistik sebagai alat analisis dikelompokkan menjadi dua bagian yang berbeda, yaitu kelompok statistik parametrik dan statistik non-parametrik. Pada statistik nonparametrik tidak memerlukan persyaratan tertentu sedangkan pada statistik parametrik memerlukan persyaratan yang harus dipenuhi. Oleh karena itu, dalam uji persyaratan regresi linier ganda yang harus dilakukan pada dasarnya juga dikelompokkan menjadi dua bagian, yaitu uji persyaratan untuk statistik parametrik dan uji persyaratan untuk menggunakan regresi linier ganda. Uji persyaratan statistik terdiri dari uji normalitas, uji kelinieran, dan uji keberartian, sedangkan persyaratan parametrik terdiri dari uji asumsi klasik.

Dalam statistika sebuah model regresi linier berganda dikatakan baik atau cocok, jika dipenuhi asumsi-asumsi ideal (klasik), yakni tidak adanya autokorelasi, heterokedstisitas dan multikolinieritas. Sehingga proses kontrol terhadap model perlu dilakukan untuk menelaah dipenuhi tidaknya asumsi tersebut. (Ifadah, Anna. 2011).

Salah satu permasalahan dari ketiga asumsi model regresi linier berganda adalah seringnya terjadi korelasi antar variabel-variabel bebas pada model regresi

linier berganda yang disebut sebagai multikolinieritas. Montgomery dan Hines (1990) menjelaskan bahwa dampak multikolinieritas dapat mengakibatkan koefisien regresi yang dihasilkan oleh analisis regresi linier berganda menjadi sangat lemah atau tidak dapat memberikan hasil analisis yang mewakili sifat atau pengaruh dari variabel bebas yang bersangkutan. Dalam banyak hal masalah multikolinieritas dapat menyebabkan uji T menjadi tidak signifikan padahal jika masing-masing variabel bebas diregresikan secara terpisah dengan variabel tak bebas (simple regression) uji T menunjukkan hasil yang signifikan. Hal tersebutlah yang sering kali membuat pusing para peneliti karena hasil analisis yang dilakukan pada regresi linier berganda tidaklah sejalan atau bahkan sangat bertentangan. Jika ada data yang terdapat multikolinieritas berarti salah satu asumsi klasik regresi linier dilanggar maka kesimpulan yang didapat dari hasil pengujian untuk model regresi linier berganda maupun untuk masing-masing peubah yang ada dalam model seringkali tidak tepat. Oleh sebab itu, masalah multikolinieritas harus dihindari.

Untuk itu maka perlu dilakukan penyembuhan multikolinieritas agar persyaratan asumsi klasik terpenuhi. Ada beberapa prosedur yang dapat digunakan untuk mengatasi masalah multikolinieritas. Apabila seleksi variabel diperbolehkan dan tidak mengubah teori yang ada maka cara yang paling mudah untuk mengatasi multikolinieritas adalah dengan mengeluarkan salah satu atau beberapa variabel bebas tak penting dalam model sehingga akan diperoleh estimator dengan varian lebih kecil. Namun, tidak semua permasalahan jika terjadi multikolinieritas dapat menggunakan metode tersebut dalam mengatasinya

karena dapat mempengaruhi variabel tak bebas. Selain metode kuadrat terkecil masih ada beberapa metode untuk mengatasi masalah multikolinieritas diantaranya ada metode *Partial Least Square* (PLS) merupakan proses pendugaan yang dilakukan secara iteratif dengan melibatkan struktur keragaman variabel bebas dan variabel tak bebas (Nurhasan, *et al.* 2012). Model yang dihasilkan oleh metode *Partial Least Square* (PLS) mengoptimalkan hubungan antara dua kelompok variabel. Metode *Ridge Regression* ditujukan untuk mengatasi kondisi buruk (*ill conditioned*) yang diakibatkan oleh korelasi yang tinggi antara beberapa peubah penjelas di dalam model. Hal ini menyebabkan matriks $X'X$ hampir singular, sehingga menghasilkan nilai dugaan parameter model yang tidak stabil (Drapper dan Smith, 1992), metode *Ridge Regression* dapat digunakan untuk mengatasi masalah multikolinieritas sedang. Salah satu kesulitan menggunakan *Ridge Regression* adalah dalam menentukan nilai θ yang tepat. Metode lain dalam penyembuhan kasus multikolinieritas adalah metode *Principal Component Analysis* (PCA) ini digunakan untuk meminimumkan masalah multikolinieritas tanpa harus mengeluarkan variabel bebas yang terlibat hubungan kolinier. Tujuan metode ini adalah untuk menyederhanakan variabel yang diamati dengan cara menyusutkan (mereduksi) dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi di antara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali atau yang biasa disebut dengan komponen utama. *Principal Component Analysis* (PCA) biasanya digunakan untuk mengatasi masalah multikolinieritas kuat. Sedangkan metode *Stepwise* memilih peubah berdasarkan korelasi terbesar dengan peubah yang

sudah terdapat pada model (Hanum, 2011). Regresi *Stepwise* merupakan pengembangan dari regresi komponen utama. Keunggulan metode ini adalah dalam hal pemilihan variabel yang masuk dalam model jika dibandingkan dengan metode regresi komponen utama. Pemilihan variabel pada *Stepwise* didasarkan pada jumlah kuadrat galat (JKG), sedangkan pada komponen utama menggunakan matriks korelasi dari variabel-variabel penjelas. Prosedur awal dari metode *Stepwise* yaitu mentransformasi variabel penjelas menjadi variabel komponen utama. Tujuan metode ini adalah pemilihan variabel terbaik yang masuk dalam model. Metode *Stepwise* biasanya digunakan untuk mengatasi masalah multikolinieritas kuat. Untuk pemilihan model regresi linier berganda terbaik ada beberapa metode yang bisa digunakan. Diantaranya *Stepwise* dan *Best Subset Regression*, kedua metode ini memulai pemilihan dengan model paling sederhana yaitu model dengan satu peubah selanjutnya disusupkan peubah lain satu persatu sampai didapat model yang memenuhi kriteria terbaik.

Berdasarkan uraian di atas maka peneliti ingin membandingkan metode *Principal Component Analysis* (PCA) dan metode *Stepwise* untuk mencari model regresi linier berganda terbaik pada kasus multikolinieritas kuat, dengan kriteria pembandingan yang digunakan untuk kedua metode yaitu R^2 Adjusted dan S^2 yang bertujuan untuk menentukan proporsi atau persentase total variasi dalam variabel terikat yang diterangkan oleh variabel bebas dengan menggunakan data yang mengandung multikolinieritas.

Berdasarkan penjelasan di atas, peneliti tertarik untuk membandingkan metode regresi *Principal Component Analysis* (PCA) dengan *Stepwise* untuk

pemilihan model terbaik pada kasus *multikolinieritas*. Oleh karena itu penulis mengangkat judul untuk penelitian ini yaitu “*Pemilihan Model Regresi Linier Berganda Terbaik Pada Kasus Multikolinieritas Berdasarkan Metode Principal Component Analysis (PCA) dan Metode Stepwise*”.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka peneliti mengidentifikasi masalah sebagai berikut:

1. Memahami analisis regresi linier berganda.
2. Mengidentifikasi kriteria model terbaik pada regresi linier berganda.
3. Mengidentifikasi masalah multikolinieritas yang sering terjadi pada analisis regresi linier berganda.
4. Dalam banyak hal masalah multikolinieritas dapat menyebabkan uji T menjadi tidak signifikan padahal jika masing-masing variabel bebas diregresikan secara terpisah dengan variabel tak bebas (simple regression) uji T menunjukkan hasil yang signifikan.
5. Memilih model regresi terbaik dengan menggunakan metode *Principal Component Analysis (PCA)* dan metode *Stepwise* berdasarkan nilai $R^2_{adjusted}$ dan S^2

1.3 Batasan Masalah

Pembatasan masalah dilakukan dengan tujuan agar pokok permasalahan yang diteliti tidak terlalu meluas dari yang telah ditetapkan. Batasan permasalahan pada penulisan skripsi ini adalah:

1. Analisis regresi linier berganda adalah analisis regresi yang melibatkan lebih dari satu atau beberapa variabel bebas.
2. Pemilihan model regresi linier berganda terbaik pada kasus multikolinieritas dengan metode *Principal Component Analysis* (PCA) dan metode *Stepwise* dengan bantuan SPSS. 16.
3. Data yang akan diteliti adalah data sekunder.

1.4 Rumusan Masalah

Berdasarkan penjelasan latar belakang di atas, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana prosedur pemilihan model regresi linier berganda pada kasus multikolinieritas dengan Metode *Principal Component Analysis* (PCA) dan metode *Stepwise*?
2. Metode manakah antara *Principal Component Analysis* (PCA) dan *Stepwise* yang cocok untuk mencari model terbaik pada kasus multikolinieritas?

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan peneliti proposal ini sebagai berikut:

1. Untuk mengetahui prosedur pemilihan model regresi linier berganda pada kasus multikolinieritas dengan metode *Principal Component Analysis* (PCA) dan metode *Stepwise*.
2. Untuk mengetahui metode yang terbaik antara metode *Principal Component Analysis* (PCA) dan metode *Stepwise* dalam pemilihan model regresi linier berganda pada kasus multikolinieritas.

1.6 Manfaat Penelitian

Manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut:

1. Bagi penulis

Selain dapat mengaplikasikan teori yang telah didapat di tempat perkuliahan dengan permasalahan yang nyata terjadi, juga akan menambah pengetahuan akan masalah-masalah yang terjadi dalam regresi linier.

2. Bagi akademik

Menambah tambahan ilmu dan wawasan yang baru tentang pemilihan model regresi linier berganda terbaik pada kasus multikolinieritas dengan menggunakan metode *Principal Component Analysis* (PCA) dan metode *Stepwise*.

3. Bagi pembaca

Diharapkan agar hasil penelitian ini dapat menambah pengetahuan pembaca mengenai topik yang terkait dengan penulisan ini.

1.7 Penegasan Masalah

Untuk menghindari salah pengertian atau tafsiran dalam pemakaian istilah yang berkaitan dengan judul penelitian ini maka diperlukan penjelasan pengertian istilah sebagai berikut:

1. Model Regresi Terbaik

Model regresi terbaik adalah model yang dapat menjelaskan perilaku peubah tak bebas dengan sebaik-baiknya dengan memilih peubah-peubah bebas dari sekian banyak peubah bebas yang tersedia dalam data. Dalam penelitian kriteria untuk mencari model terbaik menggunakan $R^2_{adjusted}$ dan S^2 .

2. Analisis Regresi Berganda 6 variabel.

Analisis regresi berganda 6 variabel pada penelitian ini artinya bahwa peneliti hanya menggunakan variabel bebas atau variabel preditor dengan 6 macam.

1.8 Sistematika Penulisan Skripsi

Secara garis besar skripsi ini dibagi menjadi tiga bagian (bab) yaitu bagian awal skripsi, bagian isi skripsi, dan bagian akhir skripsi. Berikut ini dijelaskan masing-masing bagian skripsi.

Bagian awal skripsi meliputi halaman judul, pernyataan keaslian tulisan, pengesahan, motto dan persembahan, kata pengantar, abstrak, daftar isi, daftar gambar, daftar tabel, dan daftar lampiran.

Bagian isi skripsi secara garis besar terdiri dari lima bab, yaitu: (1) Bab 1 Pendahuluan. Bab ini berisi mengenai latar belakang, identifikasi masalah, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, penegasan masalah, dan sistematika penulisan skripsi; (2) Bab 2 Tinjauan Pustaka. Bab ini berisi kajian teori yang mendasari dan berhubungan dengan pemecahan masalah. Teori-teori tersebut digunakan untuk memecahkan masalah yang diangkat dalam skripsi ini. Teori yang digunakan adalah *Matriks*, Regresi Linier Berganda, *Multikolinieritas*, Asumsi Klasik, *Principal Component Analysis* (PCA), *Stepwise*, SPSS 16.0, penelitian terdahulu, dan kerangka berfikir; (3) Bab 3 Metode Penelitian. Bab ini mengulas metode yang digunakan dalam penelitian yang berisi langkah-langkah yang dilakukan untuk memecahkan masalah yaitu penemuan masalah, kajian pustaka, analisis dan pemecahan masalah, dan penarikan kesimpulan; (4) Bab 4 Hasil Penelitian Dan Pembahasan. Bab ini berisi

mengenai penyelesaian dari permasalahan yang diungkapkan; (5) Bab Penutup. Bab ini berisi tentang simpulan dari pembahasan dan saran yang berkaitan dengan simpulan.

Bagian akhir skripsi meliputi daftar pustaka yang memberikan informasi tentang buku sumber serta literatur yang digunakan dan lampiran-lampiran yang mendukung skripsi.



BAB 2

TINJAUAN PUSTAKA

2.1 Matriks

2.1.1 Definisi Matriks

Sebuah matriks adalah sebuah susunan segi empat dari bilangan-bilangan. Bilangan-bilangan di dalam susunan tersebut dinamakan entri di dalam matriks (Anton, 1994: 22). Matriks dapat ditulis sebagai berikut:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Susunan di atas disebut matriks m kali n (ditulis $m \times n$), karena memiliki m baris dan n kolom.

2.1.2 Penjumlahan Matriks

Jika A dan B adalah sebarang dua matriks ukurannya sama, maka jumlah $A + B$ adalah matriks yang diperoleh dengan menambahkan bersama-sama entri yang bersesuaian dalam kedua matriks tersebut (Anton, 1994: 23).

2.1.3 Perkalian Matriks

Jika A adalah suatu matriks dan c adalah suatu skalar, maka hasil kali (*product*) cA adalah matriks yang diperoleh dengan mengalikan masing-masing entri dari A oleh c (Anton, 1994: 24).

Jika A adalah matriks $m \times r$ dan B adalah matriks $r \times n$, maka hasil kali AB adalah matriks $m \times n$ yang entri-entrinya ditentukan sebagai berikut. Untuk

mencari entri dalam baris i dan kolom j dari \mathbf{AB} , pilih baris i dari matriks \mathbf{A} dan kolom j dari matriks \mathbf{B} . Kalikanlah entri-entri yang bersesuaian dari baris dan kolom tersebut bersama-sama dan kemudian tambahkanlah hasil kali yang dihasilkan (Anton, 1994: 25).

2.1.4 Transpos Matriks

Jika \mathbf{A} adalah sebarang matriks $m \times n$, maka transpos \mathbf{A} dinyatakan oleh \mathbf{A}' dan didefinisikan dengan matriks $n \times m$ yang kolom pertamanya adalah baris pertama dari \mathbf{A} , kolom keduanya adalah baris kedua dari \mathbf{A} , demikian juga dengan kolom ketiga dari \mathbf{A} , dan seterusnya (Anton, 1994: 27). Jika ukuran matriks seperti operasi yang diberikan dapat dilakukan, maka:

- $(\mathbf{A}^{-1})' = \mathbf{A}'$
- $(\mathbf{A} \pm \mathbf{B})^{-1} = \mathbf{A}^{-1} \pm \mathbf{B}^{-1}$
- $(k\mathbf{A})^{-1} = k\mathbf{A}^{-1}$, dimana k adalah sebarang vektor
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (Anton, 1994: 37)

2.1.5 Invers Matriks

Jika \mathbf{A} dan \mathbf{B} matriks bujur sangkar $n \times n$ demikian sehingga $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, \mathbf{B} disebut invers \mathbf{A} , ($\mathbf{B} = \mathbf{A}^{-1}$) dan \mathbf{A} disebut invers \mathbf{B} , ($\mathbf{A} = \mathbf{B}^{-1}$). Untuk operasi baris tereduksi \mathbf{A} terhadap \mathbf{I}_n akan mereduksi \mathbf{I}_n pada \mathbf{A}^{-1} .

2.1.6 Trace Matriks

Jika \mathbf{A} adalah matriks persegi, maka jumlah unsur-unsur diagonal utamanya disebut *trace* atau telusur matriks itu. *Trace* matriks \mathbf{A} dinyatakan dengan tanda $tr(\mathbf{A})$. Jadi jika \mathbf{A} adalah matriks persegi, maka $tr(\mathbf{A}) = \sum a_{ii}$.

2.2 Regresi

Dalam beberapa masalah terdapat dua atau lebih variabel yang hubungannya tidak dapat dipisahkan, dan hal tersebut biasanya diselidiki sifat hubungannya. Analisis regresi adalah sebuah teknik statistik untuk membuat model dan menyelidiki hubungan antara dua variabel atau lebih. Salah satu tujuan dari analisis regresi adalah menentukan model regresi yang baik, sehingga dapat digunakan untuk menerangkan dan memprediksi hal-hal yang berhubungan dengan variabel-variabel yang terlibat di dalam model (Widianingsih, 2008: 15)

Analisis regresi merupakan alat analisis statistik yang berguna untuk mengetahui pengaruh variabel bebas terhadap variabel terikatnya. Pengaruh ini diwujudkan dari besarnya nilai pengaruh dalam bentuk persentase (%) (Ariyanto, 2005: 32).

Bentuk paling sederhana dari model regresi sering disebut dengan regresi linier sederhana yaitu hubungan antara satu variabel tak bebas dengan satu variabel bebas.

Bentuk hubungannya dapat dilihat dalam persamaan berikut:

$$Y_i = \beta_u + \beta_1 X_1 + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.1)$$

Persamaan diatas menyatakan bahwa rata-rata dari Y berkaitan linier dengan X , β_u , dan β_1 adalah parameter yang akan diduga nilainya dan ε adalah gangguan (*disturbance*) yang akan ikut mempengaruhi nilai Y , tetapi diabaikan dalam model.

Dalam persoalan penelitian yang menggunakan analisis regresi pada umumnya memerlukan lebih dari satu variabel bebas dalam regresinya. Oleh

karena itu, model sederhana tidak bisa dipakai, sehingga diperlukan model regresi yang mempunyai lebih dari satu variabel bebas yang disebut model regresi linier berganda (Widianingsih, 2008: 15).

2.2.1 Analisis Regresi Berganda

Menurut Sudjana (2002:10) mengemukakan bahwa: “analisis regresi adalah studi yang menyangkut hubungan yang pada umumnya dinyatakan dalam bentuk persamaan matematik yang menyatakan hubungan fungsional antara variabel-variabel.”

Regresi linier berganda adalah analisis regresi yang menjelaskan hubungan antara variabel terikat dengan faktor-faktor yang menjelaskan yang mempengaruhi lebih dari satu variabel bebas. Tujuan analisis regresi linier berganda adalah untuk memuat prediksi/perkiraan nilai Y atas X. Bentuk persamaan linier berganda adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \text{ (Sembiring, 1995: 143)} \quad (2.2)$$

Keterangan:

Y_i = pengamatan ke- i variabel tak bebas

X_{ji} = pengamatan ke- i pada variabel bebas dimana $j = 1, 2, \dots, k$ dan $i = 1, 2, \dots, n$.

β_0 = parameter intercept (konstanta)

$\beta_1, \beta_2, \dots, \beta_k$ = parameter koefisien regresi variabel tak bebas

ε_i = error (galat kesalahan), $\varepsilon_i \sim N(0; \sigma^2)$.

k = banyaknya variabel bebas

Suatu model regresi linier berganda dengan k variabel bebas dimana $\beta_i, i = 0, 1, \dots, k$ disebut bilangan pokok (koefisien) regresi. Parameter β_i mewakili perubahan yang diharapkan dalam variabel terikat Y di tiap unit berubah ke X_j ketika semua variabel bebas yang tersisa $X_i (i \neq j)$ tidak berubah.

Bentuk umum sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Atau

$$y = x\beta + \varepsilon \quad (\text{Sembiring: 1995}) \quad (2.3)$$

Keterangan:

y = vektor kolom $n \times 1$ dari variabel tak bebas Y

x = matriks $n \times (k + 1)$ dari variabel bebas X

β = vektor kolom $(k + 1) \times 1$ dari parameter yang tidak diketahui $\beta_0, \beta_1, \dots, \beta_k$

ε = vektor kolom $n \times 1$ dari gangguan (*disturbance*) ε_i .

Penambahan variabel bebas ini diharapkan dapat lebih menjelaskan karakteristik hubungan yang ada, walaupun masih saja ada variabel yang terabaikan.

2.3 Asumsi Klasik

2.3.1 Uji Persyaratan Statistik

Statistik parametrik memerlukan persyaratan yang harus dipenuhi. Oleh karena itu, dalam uji persyaratan regresi linier ganda yang harus dilakukan uji persyaratan untuk statistik parametrik terlebih dahulu sebelum uji persyaratan untuk menggunakan regresi linier ganda. Uji persyaratan statistik terdiri terdiri:

2.3.1.1 Uji Normalitas

Menurut Suliyanto (2008) uji normalitas dimaksudkan untuk mengetahui apakah residual yang telah distandardisasi berdistribusi normal atau tidak. Nilai residual dikatakan berdistribusi normal jika nilai residual tersebut sebagian besar mendekati nilai rata-ratanya sehingga bila residual tersebut berdistribusi normal maka jika digambarkan dalam bentuk kurva, kurva tersebut akan berbentuk lonceng (*ell-shaped curve*) yang kedua sisinya melebar sampai tidak terhingga. Melihat pengertian uji normalitas tersebut maka uji normalitas disini tidak dilakukan per variabel (*univariate*) tetapi hanya terhadap nilai residual terstandarisasinya saja (*multivariate*). Tidak terpenuhinya normalitas pada umumnya disebabkan karena distribusi data yang dianalisis tidak normal, karena terdapat *outlier* dalam data yang diambil. Nilai *outlier* ini dapat terjadi karena adanya kesalahan dalam pengambilan sampel, bahkan karena kesalahan dalam melakukan input data atau memang karenakarakteristik data tersebut memang aneh.

Untuk mendeteksi apakah nilai residual terstandarisasi berdistribusi normal atau tidak, dapat digunakan uji Kolmogorov-Smirnov (Suliyanto, 2008).

Hipotesis yang diuji:

H_0 : data berasal dari populasi berdistribusi normal

H_1 : data berasal dari populasi tidak berdistribusi normal

Uji ini dilakukan dengan menggunakan langkah-langkah:

1. Membuat persamaan regresi.
2. Mencari nilai prediksinya (\hat{Y}) .

3. Mencari nilai residualnya ($Y - \hat{Y}$).
4. Mengurutkan nilai residual terstandarisasi dari yang terkecil sampai yang terbesar.
5. Mencari nilai Z_r relatif kumulatif.
6. Mencari nilai Z_t teoritis berdasarkan Tabel Z.
7. Menghitung selisih nilai Z_r dengan Z_t dan diberi simbol K .
8. Mencari nilai K mutlak terbesar dan beri nama dengan K_{hitung} .
9. Bandingkan nilai K_{hitung} dengan Tabel Kolmogorov-Smirnov (K_{tabel}).
10. Menarik kesimpulan dengan kriteria jika $K_{hitung} < K_{tabel}$ maka residual terstandarisasi berdistribusi normal, atau jika $sig > 0,05$ (pada perhitungan menggunakan SPSS 16) maka residual berdistribusi normal.

Konsekuensi jika asumsi normalitas tidak terpenuhi adalah nilai prediksi yang diperoleh akan bias dan tidak konsisten. Untuk mengatasi jika asumsi normalitas tidak terpenuhi dapat digunakan beberapa metode berikut (Suliyanto, 2008):

1. Menambah jumlah data.
2. Melakukan transformasi data menjadi log atau LN atau bentuk lainnya.
3. Menghilangkan data yang dianggap sebagai penyebab data tidak normal.

2.3.1.2 Uji Kelinearan

Menurut Suliyanto (2008) pengujian linieritas perlu dilakukan untuk mengetahui model yang dibuktikan merupakan model linier atau tidak. Uji linieritas dilakukan agar diperoleh informasi apakah model empiris sebaiknya linier, kuadrat, atau kubik. Apabila salah dalam menentukan model regresi maka

nilai prediksi yang dihasilkan akan menyimpang jauh sehingga nilai prediksinya akan menjadi bias.

Menurut Suliyanto (2008) uji Lagrange Multiplier (LM-Test) merupakan salah satu metode yang digunakan untuk mengukur linieritas yang dikembangkan oleh Engle pada tahun 1982. Prinsip metode ini adalah membandingkan antara nilai X_{hitung}^2 dengan nilai X_{tabel}^2 dengan $df=(n, \alpha)$. Langkah-langkahnya adalah sebagai berikut:

1. Membuat persamaan regresinya.
2. Mencari nilai prediksinya (\hat{Y}).
3. Mencari nilai residualnya ($Y - \hat{Y}$).
4. Mengkuadratkan semua nilai variabel bebas.
5. Meregresikan kuadrat variabel bebas terhadap nilai residualnya.
6. Mencari nilai koefisien determinasinya (R^2).
7. Menghitung nilai $X_{hitung}^2 = (n \times R^2)$ seperti pada persamaan (2.4)

$$X_{hitung}^2 = (n \times R^2)$$

(2.4)

dimana n adalah jumlah pengamatan.

8. Menarik kesimpulan uji linieritas, dengan kriteria jika $X_{hitung}^2 < X_{tabel}^2$ dengan $df=(n, \alpha)$ maka model dinyatakan linier. Demikian juga sebaliknya.

Hipotesis yang diuji:

$H_0 : \rho = 0$ tidak ada hubungan yang linier antara X_1, X_2, \dots, X_k dengan Y

$H_1 : \rho \neq 0$ ada hubungan yang linier antara X_1, X_2, \dots, X_k dengan Y .

Kriteria pengujian dengan menggunakan taraf kesalahan $\alpha = 0,05$ dan $df:(n,\alpha)$ yaitu, jika nilai nilai $X_{hitung}^2 < X_{tabel}^2$ maka tolak H_0 , dan sebaliknya.

2.3.1.3 Uji Keberartian

Uji keberartian (signifikansi) koefisien regresi dimaksudkan untuk mengetahui apakah nilai koefisien regresi (nilai b dari persamaan $\hat{Y} = a + bX$) tersebut signifikan atau tidak.

Ada dua cara uji keberartian dalam regresi linier berganda yaitu:

1. Uji Keberartian Simultan

Menurut Suliyanto (2008) pengujian keberartian simultan digunakan untuk menguji ketepatan model. Uji signifikansi simultan sering disebut uji F, digunakan untuk menguji apakah variabel bebas yang digunakan dalam model secara simultan (bersama-sama) mampu menjelaskan perubahan nilai variabel tak bebas atau tidak.

Hipotesis yang diuji

$$H_0 : \beta_1, \beta_2, \dots, \beta_k = 0$$

$$H_1 : \text{Tidak semua } \beta_i = 0, i = 1, 2, \dots, k$$

Menurut Suliyanto (2008) uji signifikansi simultan menggunakan nilai F_{hitung} yang dapat diperoleh dengan perhitungan rumus seperti pada persamaan (2.5). Nilai F_{hitung} juga dapat diperoleh menggunakan SPSS 16. Kriteria pengujian keberartian simultan yaitu jika $F_{hitung} \geq F_{tabel}$ atau nilai $\text{sig}(\alpha) < 0,05$ maka variabel bebas yang digunakan dalam model secara simultan (bersama-sama) mampu menjelaskan perubahan nilai variabel tak bebas. Nilai F_{tabel} diperoleh dari Tabel distribusi F dengan $df: \alpha, (k-1), (n-k)$

$$F = \frac{R^2/(k-1)}{1-R^2/(n-1)}$$

(2.5)

2. Uji Keberartian Parsial

Pengujian keberartian parsial perlu dilakukan untuk mengetahui keberartian masing-masing variabel bebas terhadap variabel tak bebas.

Hipotesis yang diuji:

$$H_0 : \beta_1 = 0, X_1 \text{ tidak mempengaruhi } Y$$

$$H_1 : \beta_1 \neq 0, X_1 \text{ mempengaruhi } Y$$

∴

$$H_0 : \beta_k = 0, X_k \text{ tidak mempengaruhi } Y$$

$$H_1 : \beta_k \neq 0, X_k \text{ mempengaruhi } Y$$

Uji keberartian parsial menggunakan nilai t_{hitung} , dengan kriteria pengujian jika $t_{hitung} > t_{tabel}$ atau nilai $(\alpha) < 0,05$ maka variabel bebas memiliki pengaruh yang berarti terhadap variabel tak bebas. Nilai t_{tabel} diperoleh dengan df: $\alpha, (n-k)$. Nilai t_{hitung} dapat diperoleh menggunakan rumus dengan persamaan (2.6) atau menggunakan SPSS 16.

$$t_{hitung} = \frac{b_j}{Sb_j} \quad (2.6)$$

Keterangan:

b_j = koefisien regresi

Sb_j = kesalahan baku koefisien regresi.

2.3.2 Uji Asumsi Klasik

Analisis regresi merupakan alat analisis yang termasuk statistik parametrik. Sebagai alat statistik parametrik analisis regresi membutuhkan asumsi yang perlu dipenuhi sebelum dilakukan analisis. Analisis ini dinamakan dengan uji asumsi klasik. Asumsi klasik tersebut dapat menghilangkan estimator liniertidak bias yang terbaik dari model regresi yang diperoleh dari metode kuadrat terkecil biasa. Dengan terpenuhinya asumsi tersebut, maka hasil yang diperoleh dapat lebih akurat dan mendekati atau sama dengan kenyataan.

Adapun model regresi linier berganda secara umum adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad i = 1, 2, \dots, n \quad (\text{Montgomery, 2006:67})$$

Dalam sebuah model regresi dikatakan baik, jika dipenuhi asumsi-asumsi sederhana yang sering disebut sebagai asumsi klasik yaitu:

1. Nilai rata-rata kesalahan pengguna nol, yaitu $E(\varepsilon_i) = 0$, untuk $i = 1, 2, \dots, n$.
2. Varian $(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2$, sama untuk semua kesalahan pengganggu (asumsi: homoskedastik).
3. Tidak ada autokorelasi antara kesalahan pengganggu, berarti kovarian $(\varepsilon_i, \varepsilon_j) = 0, i \neq j$.
4. Variabel Bebas X_1, X_2, \dots, X_k , konstanta dalam sampling yang terulang dan bebas terhadap kesalahan pengganggu ε_i .
5. Tidak ada kolinieritas ganda (*multicollinierty*) di antara variabel bebas X.
6. $\varepsilon_i \sim N(0; \sigma^2)$, artinya kesalahan pengganggu mengikuti distribusi normal dengan rata-rata nol dan varian σ^2 .

Uji asumsi klasik dalam regresi mencakup:

2.3.2.1 Uji Multikolinieritas

Menurut Suliyanto (2008) pengertian kolinieritas sering dibedakan dengan multikolinieritas. Kolinieritas berarti terjadi korelasi linier yang mendekatisempurna antara kedua variabel bebas. Sedangkan multikolinieritas berarti terjadikorelasi linier yang mendekati sempurna antara lebih dari dua variabel bebas.

Multikolinieritas bisa terjadi saat adanya kesalahan spesifikasi model (*spesificationmodel*). Hal ini dapat terjadi karena seorang peneliti memasukan variabel bebas yang seharusnya dikeluarkan dari model empiris. Dapat juga terjadi karena seorangpeneliti mengeluarkan variabel bebas yang seharusnya dimasukkan dalam modelempiris. Selain itu, adanya model yang berlebihan (an overdetermined model) jugadapat menyebabkan multikolinieritas. Hal ini terjadi ketika model empiris (jumlahvariabel bebas) yang digunakan melebihi jumlah data (observasi).

Untuk mendeteksi adanya masalah multikolinieritas, metode yang palingsering digunakan yaitu dengan menggunakan nilai VIF (*Variance Inflation Factor*). Untuk mengetahui ada tidaknya multikolinieritas antar variabel, salah satu caranyadengan melihat nilai *Variance Inflation Factor* (VIF) dari masing-masing variabelbebas terhadap variabel terikatnya. Menurut Gujarati (Suliyanto, 2008), jika nilaiVIF tidak lebih dari 10 maka model dikatakan tidak mengandung multikolinieritas.

Beberapa akibat yang timbul jika hasil estimasi model empiris mengalamimasalah multikolinieritas diantaranya (Suliyanto, 2008):

1. Penaksir OLS tidak bisa ditentukan (*indeterminate*) meskipun hasil estimasi yang dihasilkan masih BLUE (*Best Linier Unbiased Estimator*).
2. Interval kepercayaan cenderung meningkat lebih besar sehingga mendorong untuk menerima hipotesis nol (antara lain koefisien populasi adalah nol).
3. Nilai t-statistik koefisien dari satu atau beberapa variabel bebas secara statistik tidak signifikan sehingga dapat menyebabkan dikeluarkannya suatu variabel bebas dalam model regresi, padahal variabel bebas tersebut memiliki peran yang sangat penting dalam menjelaskan variabel terikat.
4. Penaksir-penaksir OLS dan kesalahan bakunya cenderung tidak stabil dan sangat sensitif bila terjadi perubahan data, meskipun perubahan itu sangat kecil.
5. Jika multikolinieritas sangat tinggi maka mungkin R^2 bisa tinggi namun sangat sedikit taksiran koefisien regresi yang signifikan secara statistik.

Menurut Suliyanto (2008) beberapa cara untuk mengatasi multikolinier diantaranya memperbesar ukuran sampel, menghilangkan salah satu atau lebih variabel bebas, menggabungkan data *time series* dan data *cross section*, atau melakukan transformasi data.

2.3.2.2 Uji Heteroskedastisitas

Menurut Suliyanto (2008) heteroskedastisitas berarti ada varians variabel dalam model yang tidak sama (konstan). Sebaliknya jika varians variabel dalam model memiliki nilai yang sama (konstan) disebut sebagai homoskedastisitas. Untuk menguji adanya masalah heteroskedastisitas dapat dilakukan dengan menggunakan metode Glejser.

Uji Glejser (Suliyanto, 2008) dilakukan dengan meregresikan semua variabel bebas terhadap nilai mutlak residualnya. Jika terdapat pengaruh variabel bebas yang signifikan terhadap nilai mutlak residualnya ($sig < 0,05$) maka dalam model terdapat masalah heteroskedastisitas.

Hipotesis yang diuji:

H_0 : tidak terdapat heteroskedastisitas

H_1 : terdapat heteroskedastisitas

Menurut Gujarati (Suliyanto, 2008) ada beberapa konsekuensi sebagai akibat dari adanya masalah heteroskedastisitas dalam model persamaan regresi diantaranya:

1. Walaupun penaksir OLS masih linier dan masih tak bias, tetapi akan mempunyai varian yang tidak minimum lagi serta tidak efisien dalam sampel kecil. Lebih lanjut penaksir OLS juga tidak efisien dalam sampel besar.
2. Formulasi untuk menaksir varian dari estimasi OLS secara umum adalah bias, dimana bila menaksir secara apriori, seorang peneliti tidak dapat mengatakan bahwa bias tersebut akan positif atau negatif. Akibatnya interval kepercayaan dan uji hipotesis yang didasarkan pada uji t dan nilai distribusi F tidak dapat dipercaya.
3. Prediksi yang didasarkan pada koefisien parameter variabel bebas dari data asli akan mempunyai varian yang tinggi sehingga prediksi tidak efisien.

Menurut Suliyanto (2008) perbaikan model apabila terjadi masalah heteroskedastisitas diantaranya melakukan transformasi model regresi dengan membagi model regresi dengan salah satu variabel independen yang

digunakan dalam model regresi tersebut atau melakukan transformasi logaritma dan LN.

2.3.2.3 Uji Autokorelasi

Menurut Suliyanto (2008) uji autokorelasi bertujuan untuk mengetahui apakah ada korelasi antara anggota serangkaian data observasi yang diuraikan menurut waktu (*time series*) atau ruang (*cross section*).

Menurut Gujarati (Suliyanto, 2008) ada beberapa cara untuk mendeteksi adanya masalah autokorelasi salah satunya yaitu Uji Durbin Watson (Uji DW). Uji DW pertama kali diperkenalkan oleh J. Durbin dan G. S. Watson tahun 1951.

Rumus yang digunakan untuk uji DW adalah

$$DW = \frac{\sum(e - e_{t-1})^2}{\sum e^2} \quad (2.7)$$

Keterangan:

DW = Nilai Durbin-Watson Test

e = Nilai residual

e_{t-1} = Nilai residual satu baris/periode sebelumnya

Dengan kriteria pengujian seperti pada Tabel 2.1 di bawah:

Tabel 2.1 Kriteria Pengujian Autokorelasi dengan Durbin-Watson

Range	Keputusan
-------	-----------

$0 < dw < dL$	Terjadi masalah autokorelasi yang positif yang perlu perbaikan.
$dL < dw < dU$	Ada autokorelasi positif tetapi lemah, dimana perbaikan akan lebih baik.
$-2 < dw < 2$	Tidak ada masalah autokorelasi.
$4-dU < dw < 4-dL$	Masalah autokorelasi lemah, dimana dengan perbaikan akan lebih baik.
$4-dL < dw$	Masalah autokorelasi serius.

Sumber : Durbin-Watson Table

Menurut Gujarati (Suliyanto, 2008) menyebutkan beberapa konsekuensi dari munculnya masalah autokorelasi dalam analisis regresi bahwa penaksir OLS *unbiased* dalam penyampelan berulang dan konsisten, tetapi sebagaimana dalam kasus heteroskedastisitas, penaksir OLS tidak lagi efisien (mempunyai varian minimum), baik dalam sampel kecil maupun sampel besar.

Menurut Suliyanto (2008) untuk memperbaiki autokorelasi dapat dilakukan dengan cara diantaranya dengan membuat persamaan perbedaan yang digeneralisasikan atau dengan metode perbedaan pertama.

2.4 Multikolinieritas

2.4.1 Pengertian Multikolinieritas

Menurut Suliyanto (2008) pengertian kolinieritas sering dibedakan dengan multikolinieritas. Kolinieritas berarti terjadi korelasi linier yang mendekatisempurna antara kedua variabel bebas. Sedangkan multikolinieritas berarti terjadikorelasi linier yang mendekati sempurna antara lebih dari dua variabel bebas.

Multikolinieritas bisa terjadi saat adanya kesalahan spesifikasi model (*spesificationmodel*). Hal ini dapat terjadi karena seorang peneliti memasukan variabel bebas yang seharusnya dikeluarkan dari model empiris. Dapat juga terjadi

karena seorang peneliti mengeluarkan variabel bebas yang seharusnya dimasukkan dalam model empiris. Selain itu, adanya model yang berlebihan (an overdetermined model) juga dapat menyebabkan multikolinieritas. Hal ini terjadi ketika model empiris (jumlah variabel bebas) yang digunakan melebihi jumlah data (observasi).

2.4.2 Penyebab Terjadinya Multikolinieritas

Masalah multikolinieritas dapat timbul karena berbagai sebab. Sebagai contoh karena sifat-sifat yang terkandung dalam kebanyakan variabel ekonomi berubah bersama-sama sepanjang waktu. Besaran-besaran ekonomi dipengaruhi oleh faktor-faktor yang sama. Oleh karena itu, sekali faktor-faktor yang mempengaruhi ini menjadi operatif, maka seluruh variabel akan cenderung berubah dalam satu arah. Dalam data *time series*, pertumbuhan dan faktor-faktor kecenderungan merupakan penyebab utama adanya multikolinieritas. Kedua, penggunaan nilai lag (*lagged values*) dari variabel-variabel bebas tertentu dalam model regresi.

Mengingat sifat yang sangat mendasar dari data, multikolinieritas diperkirakan terdapat pada sebagian besar hubungan-hubungan ekonomi. Oleh karena itu, perhatian sesungguhnya bukan lagi terletak pada ada atau tidaknya multikolinieritas, tetapi lebih pada akibat-akibat yang ditimbulkan oleh adanya multikolinieritas dalam sampel (Sumodiningrat, 1996: 281-282).

2.4.3 Konsekuensi Multikolinieritas

Beberapa akibat yang timbul jika hasil estimasi model empiris mengalami masalah multikolinieritas diantaranya (Suliyanto, 2008):

1. Penaksir OLS tidak bisa ditentukan (*indeterminate*) meskipun hasil estimasi yang dihasilkan masih BLUE (*Best Linier Unbiased Estimator*).
2. Interval kepercayaan cenderung meningkat lebih besar sehingga mendorong untuk menerima hipotesis nol (antara lain koefisien populasi adalah nol).
3. Nilai t-statistik koefisien dari satu atau beberapa variabel bebas secara statistik tidak signifikan sehingga dapat menyebabkan dikeluarkannya suatu variabel bebas dalam model regresi, padahal variabel bebas tersebut memiliki peran yang sangat penting dalam menjelaskan variabel terikat.
4. Penaksir-penaksir OLS dan kesalahan bakunya cenderung tidak stabil dan sangat sensitif bila terjadi perubahan data, meskipun perubahan itu sangat kecil.
5. Jika multikolinieritas sangat tinggi maka mungkin R^2 bisa tinggi namun sangat sedikit taksiran koefisien regresi yang signifikan secara statistik.

2.4.4 Cara Mendeteksi Multikolinieritas

Untuk mendeteksi adanya masalah multikolinieritas, metode yang palingsering digunakan yaitu dengan menggunakan nilai VIF (*Variance Inflation Factor*). Untuk mengetahui ada tidaknya multikolinieritas antar variabel, salah satu caranya dengan melihat nilai *Variance Inflation Factor* (VIF) dari masing-masing variabel bebas terhadap variabel terikatnya. Menurut Gujarati (Suliyanto, 2008), jika nilai VIF tidak lebih dari 10 maka model dikatakan tidak mengandung multikolinieritas.

2.5 Program SPSS 16.0 for Windows

2.5.1 Pengenalan Program SPSS

Program aplikasi SPSS (*Statistikan Package Social Science*) merupakan salah satu program yang relatif populer saat ini. Program ini terutama diperuntukkan bagi ilmu-ilmu sosial, sehingga fasilitas analisis lebih banyak variabel sosial. Program ini pada perkembangan sekarang SPSS sudah banyak digunakan oleh kalangan eksak pula. SPSS memuat perangkat-perangkat statistik dasar, sehingga cukup baik dipergunakan untuk memahami sifat-sifat suatu data dan pengolahan data secara sederhana (Sukestiyarno, 2008: 6).

Versi *Software* SPSS secara terus menerus mengalami perubahan. Saat sistem operasi komputer windows mulai populer, SPSS yang dahulu *under* DOS dan bernama SPSS PC, juga berubah menjadi *under windows* dan populer di Indonesia dengan nama SPSS Versi 6, kemudian versi 7.5, versi 9, versi 10, versi 11.5, versi 12, versi 13, versi 14, versi 15, versi 16 dan yang terakhir adalah SPSS versi 17. Selanjutnya penulis menggunakan SPSS versi 16 untuk keperluan analisis data.

2.5.2 Menu Analisis Faktor

Metode *Principal Component Analysis* (PCA) dan metode *Stepwise* diselesaikan dengan bantuan SPSS.16 dimana menu yang digunakan sebagai berikut:

1. *Analyze*

Analyze merupakan menu yang berfungsi untuk melakukan analisis data yang merupakan menu dimana anda melakukan analisis statistik mulai dari analisis deskriptif seperti menampilkan table atau grafik sampai analisis yang lebih kompleks.

2. *Regressions*

Regressions merupakan sebuah menu pada SPSS yang didalamnya berfungsi sebagai penguji dari variabel independen terhadap variabel dependen setelah diketahui ada hubungan antara variabel tersebut.

3. *Linier*

Linier merupakan sebuah sub menu pada regresi untuk menguji uji linier.

4. *Data Reduction*

Data Reduction merupakan menu pada SPSS yang berfungsi sebagai mereduksi sebuah data variabel.

5. *Factor*

Factor merupakan sub menu pada data reduction yang berfungsi mencari faktor sebagai pengganti variabel bebas sebuah penelitian.

6. *Kaiser-Meyer-Olkin (KMO) dan Barlett Test*

Mengenai layak atau tidaknya analisis faktor, maka perlu dilakukan uji *Kaiser-Meyer-Olkin (KMO)* dan *Barlett Test*. Apabila nilai KMO antara 0,5 sampai 1 maka dapat disimpulkan analisis faktor tepat digunakan (Bilson, 2005:123). Namun, jika nilai KMO kurang dari 0,5 maka analisis faktor tidak layak dilakukan. Sedangkan *Barlett Test* digunakan untuk menguji apakah benar variabel-variabel yang dilibatkan berkorelasi, sedangkan signifikansi *Bartlett's Test of Sphericity* ini kurang dari level signifikansi (α) yang digunakan dapat diartikan bahwa analisis faktor tepat digunakan.

7. *Anti Image Matriks*

Bagian *Anti Image Correlation*, khususnya pada angka korelasi yang bertanda a (arah diagonal dari kiri atas ke kanan bawah). Angka MSA (*Measure of Sampling Adequacy*) berkisar dari 0 sampai 1, dengan kriteria sebagai berikut (Santoso, 2010: 66)

- 1) $MSA = 1$, variabel tersebut dapat diprediksi tanpa kesalahan oleh variabel lain.
- 2) $MSA > 0,5$, variabel masih bisa diprediksi dan bisa dianalisis lebih lanjut.
- 3) $MSA < 0,5$, variabel tidak bisa diprediksi dan tidak bisa dianalisis lebih lanjut, atau dikeluarkan dari variabel lainnya.

8. *Communities*

Communities menunjukkan beberapa varians yang dapat dijelaskan oleh faktor yang terbentuk.

9. *Total Variance Explained*

Dalam analisis faktor terdapat beberapa komponen yang merupakan variabel. Setiap vektor mewakili variabel yang dianalisis. Kemampuan setiap faktor mewakili variabel yang dianalisis ditunjukkan oleh besarnya varians yang dijelaskan, yang disebut dengan *eigenvalue*. *Eigenvalue* menunjukkan kepentingan relatif masing-masing faktor dalam menghitung varians semua variabel yang dianalisis. Susunan *eigenvalue* selalu diurutkan dari yang terbesar sampai yang terkecil, dengan kriteria bahwa angka *eigenvalue* di bawah 1 tidak digunakan dalam menghitung jumlah faktor yang terbentuk.

10. *Component Matriks*

Componen Matriks merupakan tabel yang berisikan *faktor loading* (nilai korelasi) antara variabel-variabel analisis dengan faktor yang terbentuk. Proses penentuan variabel mana akan dimasukkan ke faktor yang mana, dilakukan dengan melakukan perbandingan besar korelasi setiap baris.

11. *Component Score Coefficient Matriks*

Setelah didapat faktor yang terbaik melalui proses reduksi, maka perlu dicari persamaan sehingga dapat dihitung skor setiap faktor secara manual. Persamaan yang dibuat mirip dengan regresi lenear berganda, hanya dalam persamaan faktornya tidak terdapat konstanta. Setelah komponen hasil PCA yang bebas mulikolinieritas diperoleh maka komponen-komponen tersebut diregresikan atau dianalisa pengaruhnya terhadap variabel tak bebas (Y) dengan menggunakan analisis regresi linier.

2.6 Metode *Principal Component Analysis (PCA)*

Metode PCA bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi antara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali. Setelah beberapa komponen hasil PCA yang bebas multikolinieritas diperoleh, maka komponen-komponen tersebut menjadi variabel bebas baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas (Y) dengan menggunakan analisis regresi. Keunggulan metode PCA diantaranya adalah dapat menghilangkan korelasi secara bersih tanpa harus mempengaruhi jumlah variabel asal.

Princial Component Analysis(PCA) merupakan salah satu analisis regresi yang menggunakan komponen utama untuk mengatasi masalah multikolinieritas pada regresi berganda (Dina *et al.*2009). Menurut Ul-Saufie *et al.*(2011) regresi komponen utama merupakan metode yang menggabungkan antara regresi linier dengan analisis komponen utama. *Principal Component Analysis* membentuk hubungan antara variabel terikat dengan komponen utama yang dipilih dari variabel bebas. Analisis komponen utama digunakan untuk menemukan suatu kombinasi linier dari kovarian yang saling tidak berkorelasi satu dengan yang lain, sehingga masalah multikolinieritas dapat dihindari. Selain itu analisis komponen utama dapat memastikan bahwa kombinasi linier yang dipilih memiliki varian maksimal.

Menurut Johnson dan Wichern (2007) Komponen utama merupakan kombinasi linier dari variabel random k (X_1, X_2, \dots, X_k). Analisis komponen utama tergantung pada matriks kovarian Σ atau matriks korelasi p . Menurut Draper dan Smith (1992), terdapat dua kriteria untuk menentukan model regresi komponen utama yaitu: mengambil akar ciri yang lebih dari 1 ($\lambda > 1$). Misalkan eigen $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq 0$ maka bentuk kombinasi linier sebagai berikut

$$W_1 = a'_1 X = a_{11}X_1 + a_{12}X_2 + \dots + a_{1k}X_k$$

(2.8)

$$W_2 = a'_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2k}$$

(2.9)

⋮

$$W_k = a'_k X = a_{k1}X_1 + a_{k2}X_2 + \dots + a_{kk}X_k \quad (2.10)$$

dengan $(W_j) = a'_j \Sigma a_j$ dan $Cov(W_j, W_1) = a'_j \Sigma a_1$ dimana $j, 1 = 1, 2, \dots, k$. Rumus untuk mencari proporsi dari varians populasi total yang dijelaskan oleh komponen utama ke-k adalah sebagai berikut.

$$\text{Proporsi dari total varians ke-k} = \frac{\lambda_j}{\lambda_1 + \lambda_2 + \dots + \lambda_k}; \quad j = 1, 2, \dots, k \quad (2.11)$$

Komponen utama dapat juga diperoleh dari variabel yang distandarkan yaitu:

$$\begin{aligned} Z_1 &= \frac{X_1 - \mu_1}{\sqrt{\sigma_{11}}} \\ Z_2 &= \frac{X_2 - \mu_2}{\sqrt{\sigma_{22}}} \\ &\vdots \\ Z_k &= \frac{X_k - \mu_k}{\sqrt{\sigma_{kk}}} \end{aligned}$$

dalam notasi matriks adalah sebagai berikut.

$$Z = \left(V^{\frac{1}{2}} \right)^{-1} (X - \mu)$$

$$\text{dengan } V^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{\sigma_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\sigma_{kk}} \end{bmatrix}$$

$$(2.12) \text{ dimana } E(Z) = 0 \text{ dan } Cov(Z) = \left(V^{\frac{1}{2}} \right)^{-1} \Sigma \left(V^{\frac{1}{2}} \right)^{-1} = \rho$$

Menurut Widiharini (2001) komponen utama ke-i dari variabel yang distandarkan

$Z' = [Z_1, Z_2, \dots, Z_k]$ dengan $Cov(Z) = \rho$ yaitu

$$W_1 = a'_1 Z = a_{11} Z_1 + a_{12} Z_2 + \dots + a_{1k} Z_k \quad (2.13)$$

$$W_2 = a'_2 Z = a_{21} Z_1 + a_{22} Z_2 + \dots + a_{2k} Z_k \quad (2.14)$$

\vdots

$$W_k = a'_k Z = a_{k1} Z_1 + a_{k2} Z_2 + \dots + a_{kk} Z_k \quad (2.15)$$

Menurut Johnson dan Wichern (2007) rumus untuk mencari proporsi dari varians populasi total yang dijelaskan oleh komponen utama ke-k adalah sebagai berikut

$$\text{Proporsi dari total varians ke-k} = \frac{\lambda_j}{k} \quad (2.16)$$

dengan λ_j adalah nilai eigen dari ρ ; $j=1,2,\dots,k$.

Menurut Johnson dan Wichern (2007) jumlah komponen utama dapat ditentukan dengan melihat persentase total varians ketika $j(j < k)$ buah komponen yang dipilih mampu menerangkan varians sekitar 80% sampai 90%. Komponen yang diambil tersebut sudah dapat menggantikan variabel-variabelnya tanpa banyak kehilangan informasi. Jumlah komponen utama juga dapat diketahui dengan menggunakan *scree plot*. *Scree plot* adalah plot antara λ_j dengan j besarnya nilai eigen. Untuk menentukan jumlah komponen utamaitudenganmelihat tikungan tajam pada *scree plot*. Jumlah komponen yang diambil adalah yang nilai eigen relatif kecil dan semuanya berukuran sama.

Secara umum, menurut Montgomery dan Peck (1991) bentuk persamaan dari model regresi komponen utamaitu:

$$y = Qa + \varepsilon \quad (2.17)$$

Keterangan:

y = Pengamatan variabel tak bebas

Q = Konstanta

a = Komponen yang dihasilkan

ε = error

Adapun algoritma *Principal Component Analysis* (PCA) sebagai berikut:

1. Menghitung *eigen value* dan *eigen vector* dari matriks korelasi atau kovarians
2. Menentukan skor komponen utama
3. Mengambil akar ciri (nilai eigen) yang lebih dari satu ($\lambda > 1$)
4. Melakukan analisis regresi linier berganda antara variabel terikat dengan komponen utama yang terpilih.

2.7 Metode *Stepwise*

Metode *Stepwise* merupakan sebuah metode yang dapat digunakan dalam menentukan model regresi berganda. Prosedur yang dilakukan yaitu dengan menyusupkan satu demi satu variabel bebas sampai diperoleh persamaan regresi yang memuaskan. Urutan penyisipan ditentukan dari koefisien korelasi parsial (r) (Draper & Smith, 1992). Jadi, variabel bebas yang mempunyai korelasi terbesar dengan variabel terikat akan masuk pertama kali dalam model.

Apabila salah satu peubah telah dimasukkan ke dalam model regresi, maka peubah lainnya tidak perlu masuk lagi ke dalam model regresi karena pengaruhnya telah diwakili oleh peubah yang sudah masuk di dalam model regresi. Sehingga tidak terdapat multikolinieritas pada model regresi yang dihasilkan (Sembiring, 1995).

Tahap penyelesaian pada metode *Stepwise* yaitu membentuk matriks koefisien korelasi. Koefisien korelasi yang dicari adalah koefisien korelasi linier sederhana Y dan X_i dengan rumus (Pakpahan, DF. L, et al):

$$r_{yx_i} = \frac{\sum(X_{ij} - \bar{X}_i)(Y_j - \bar{Y})}{\sqrt{\sum(X_{ij} - \bar{X}_i)^2 \sum(Y_j - \bar{Y})^2}}$$

(2.18)

Oleh karena itu dapat dibentuk matriks koefisien korelasi linier sederhana antara Y dan X_i :

$$\begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix}$$

Kemudian membentuk persamaan regresi linier yang pertama. Variabel yang pertama diregresikan adalah variabel yang mempunyai harga mutlak koefisien korelasi yang terbesar antara Y dan X_i , misalnya X_i .

$$X = \begin{pmatrix} 1 & X_{h1} \\ 1 & X_{h2} \\ \vdots & \vdots \\ 1 & X_{hn} \end{pmatrix} (X^T X)^{-1} = \begin{pmatrix} n & \sum X_h \\ \sum X_h & \sum X_h^2 \end{pmatrix}$$

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} X^T Y = \begin{pmatrix} \sum Y \\ \sum YX_1 \end{pmatrix}$$

Langkah selanjutnya yaitu seleksi variabel kedua diregresikan. Cara menyeleksi variabel yang kedua diregresikan adalah memilih parsial korelasi variabel sisa terbesar. Untuk menghitung harga masing-masing parsial korelasi sisa digunakan rumus:

$$r_{YX_h \cdot X_k} = \frac{r_{YX_h} - r_{YX_k} r_{X_h X_k}}{\sqrt{(1 - r_{YX_k}^2)(1 - r_{X_h X_k}^2)}}$$

(2.19)

Tahap selanjutnya adalah memasukkan variabel bebas yang memiliki korelasi terbesar kedua dengan variabel terikat. Tahap ini terus dilakukan sampai semua variabel masuk dalam model, kemudian membuat persamaan regresi kedua $Y = b_0 + b_1 X_1 + \dots + b_k X_k$. Namun setelah diperoleh satu variabel bebas yang

masuk dalam model perlu dilakukan pengujian keberartian variabel dalam model, kemudian dicek apakah koefisien regresi b_h signifikan, dengan hipotesis:

$H_0 : b_h = 0$, dimana variabel yang masuk dalam model tidak berarti

$H_1 : b_h \neq 0$, dimana variabel yang masuk dalam model berarti.

Untuk suatu $i = 1, 2, 3, \dots$

$$F_{hitung} = \left(\frac{b_h}{s(b_h)} \right)^2 \quad (2.20)$$

Pengujian dilakukan dengan uji-F dengan $F_{tabel} = F_{(1;n-p;\alpha)}$. Keputusan: bila $F_{hitung} < F_{tabel}$ terima H_0 artinya b_h dianggap sama dengan nol, maka proses dihentikan atau dikeluarkan dalam model dan persamaan terbaik $Y = b_0 + b_h X_h$. Bila $F_{hitung} \geq F_{tabel}$ tolak H_0 artinya b_h tidak sama dengan nol, maka variabel X_k tetap dalam model. Dimana α merupakan taraf nyata yang digunakan ($\alpha = 0.05$)

Dalam metode *Stepwise*, apabila ada dua variabel bebas yang saling berkorelasi, maka hanya ada satu variabel yang masuk dalam model. Pemilihan variabel yang masuk didasarkan pada koefisien korelasi terbesar terhadap variabel terikat (Iriawan, 2006).

Adapun algoritma metode *Stepwise* sebagai berikut :

1. Menentukan matriks koefisien korelasi antara variabel respon (Y) terhadap variabel bebas (X).
2. Pemilihan variabel yang pertama diregresikan yaitu variabel yang mempunyai harga mutlak koefisien korelasi terbesar terhadap responden (Y).
3. Pembentukan regresi pertama yaitu regresi sederhana untuk variabel terpilih pada langkah kedua dan menguji keberartian regresi.

4. Pemilihan variabel kedua diregresikan. Bila pada langkah ketigaternyata terima H1 maka dilakukan pemilihan variabel kedua untuk diregresikan selanjutnya. Variabel terpilih adalah variabelsisa (di luar regresi) yang mempunyai parsial korelasi terbesar.
5. Pembentukan regresi kedua yaitu merupakan regresi ganda dan menguji keberartian regresi. Bila tidak signifikan maka proses dihentikan sedangkan sebaliknya bila signifikan maka seluruh variabel tetap.
6. Pembentukan penduga apabila proses pemasukan variabel terhadap regresi sudah selesai, maka ditetapkan persamaan regresi yang menjadi penduga linier yang diinginkan yaitu merupakan persamaan regresi yang diperoleh terakhir.
7. Pembahasan pada penduga dan pembuatan kesimpulan.

2.8 Pendeteksi Kelayakan Model Regresi Terbaik

Salah satu tujuan di dalam analisis regresi adalah untuk mendapatkan model terbaik yang menjelaskan hubungan antara variabel terikat dengan variabel bebas, model terbaik adalah model yang seluruh koefisien regresinya berarti (significant) dan mempunyai kriteria model terbaik optimum.

Suatu model regresi dikatakan layak dan terbaik apabila model tersebut memenuhi kriteria sebagai berikut:

2.8.1 Nilai Adjusted R-square

Karena adanya kelemahan dalam perhitungan R^2 , banyak peneliti yang menyarankan untuk menggunakan Adjusted R-Square. Interpretasinya sama dengan R Square, akan tetapi nilai Adjusted R-Square dapat naik atau turun

dengan adanya penambahan variabel baru, tergantung dari korelasi antara variabel bebas tambahan tersebut dengan variabel terikatnya. Nilai Adjusted R-Square dapat bernilai negatif, sehingga jika nilainya negatif, maka nilai tersebut dianggap 0, atau variabel bebas sama sekali tidak mampu menjelaskan varians dari variabel terikatnya.

Suatu sifat penting R^2 adalah nilainya merupakan fungsi yang tidak pernah menurun dari banyaknya variabel bebas yang ada dalam model. Oleh karenanya, untuk membandingkan dua R^2 dari dua model, orang harus memperhitungkan banyaknya variabel bebas yang ada dalam model. Ini dapat dilakukan dengan menggunakan “adjusted R-square”. Istilah penyesuaian berarti nilai R^2 sudah disesuaikan dengan banyaknya variabel (derajat bebas) dalam model. Memang, R^2 yang disesuaikan ini juga akan meningkat bersamaan meningkatnya jumlah variabel, tetapi peningkatannya relatif kecil. Seringkali juga disarankan, jika variabel bebas lebih dari dua, sebaiknya menggunakan adjusted R square.

Asumsi – asumsi klasik yang melandasi regresi linier berganda adalah kenormalan sisaan, heterokedstisitas dan multikolinieritas antar variabel bebas. Asumsi yang sering tidak terpenuhi adalah asumsi nonmultikolinieritas berarti bahwa tidak terdapat hubungan linier antar variabel-variabel bebas dalam model regresi. Dalam bentuk matriks, multikolinieritas adalah kondisi buruk (*ill condition*) dari matrik $X'X$ yaitu kondisi yang tidak memenuhi asumsi klasik (Draper and Smith, 1992). Beberapa metode yang digunakan untuk mendeteksi adanya multikolinieritas adalah Koefisien korelasi antar variabel bebas, VIF dan bilangan kondisi.

Statistik R^2 yang sering disebut dengan Koefisien determinasi. Sedangkan R^2 adjusted memperhitungkan banyaknya variabel bebas dalam model serta telah disesuaikan terhadap derajat bebas masing-masing jumlah kuadrat. R^2 adjusted dirumuskan dengan :

$$R^2_{adjusted} = 1 - (1 - R^2) \frac{(N-1)}{(N-p)} \quad (\text{Sembiring, 1995})$$

(2.21)

Keterangan:

R^2 = Sampel R-square

p = Banyaknya parameter termasuk Y

N = Banyaknya pengamatan.

2.8.2 Nilai Rataan Kuadrat Sisa atau S^2

Salah satu patokan yang baik digunakan dalam memilih kecocokan suatu model dengan data adalah dengan melihat rataan kuadrat sisa (S^2), model yang baik memberikan S^2 yang kecil. Ukuran ini memperhitungkan banyaknya parameter dalam model melalui pembagian dengan derajat kebebasannya. Rataan kuadrat sisa (S^2) mungkin membesar bila penurunan dalam JK sisa akibat pemasukan suatu peubah tambahan kedalam model tidak dapat mengimbangi penurunan dalam derajat kebebasannya. Rataan kuadrat sisa dirumuskan:

$$S^2 = \frac{JKS}{n-p} \quad (\text{Sembiring, 1995: 236})$$

(2.22)

Keterangan:

JKS = Jumlah kuadrat sisa

n = banyaknya pengamatan

p = banyaknya parameter

2.9 Indeks LQ 45

2.9.1 Definisi Indeks LQ 45

Indeks LQ 45 adalah nilai kapitalisasi pasar dari 45 saham yang paling likuid dan memiliki nilai kapitalisasi yang besar hal itu merupakan indikator likuidasi. Indeks LQ 45, menggunakan 45 saham yang terpilih berdasarkan Likuiditas perdagangan saham dan disesuaikan setiap enam bulan (setiap awal bulan Februari dan Agustus). Dengan demikian saham yang terdapat dalam indeks tersebut akan selalu berubah.

Beberapa kriteria - kriteria seleksi untuk menentukan suatu emiten dapat masuk dalam perhitungan indeks LQ 45 adalah :

Kriteria yang pertama adalah :

1. Berada di TOP 95 % dari total rata – rata tahunan nilai transaksi saham di pasar reguler.
2. Berada di TOP 90 % dari rata – rata tahunan kapitalisasi pasar.

Kriteria yang kedua adalah :

1. Merupakan urutan tertinggi yang mewakili sektornya dalam klasifikasi industri BEJ sesuai dengan nilai kapitalisasi pasarnya.
2. Merupakan urutan tertinggi berdasarkan frekuensi transaksi (Tjiptono, 2001, p. 95-96).

Indeks LQ 45 hanya terdiri dari 45 saham yang telah terpilih melalui berbagai kriteria pemilihan, sehingga akan terdiri dari saham-saham dengan likuiditas dan

kapitalisasi pasar yang tinggi. Saham-saham pada indeks LQ 45 harus memenuhi kriteria dan melewati seleksi utama sebagai berikut :

1. Masuk dalam ranking 60 besar dari total transaksi saham di pasar reguler (rata-rata nilai transaksi selama 12 bulan terakhir).
2. Ranking berdasar kapitalisasi pasar (rata-rata kapitalisasi pasar selama 12 bulan terakhir).
3. Telah tercatat di BEJ minimum 3 bulan.
4. Keadaan keuangan perusahaan dan prospek pertumbuhannya, frekuensi dan jumlah hari perdagangan transaksi pasar reguler.

Saham-saham yang termasuk didalam LQ 45 terus dipantau dan setiap enam bulan akan diadakan review (awal Februari, dan Agustus). Apabila ada saham yang sudah tidak masuk kriteria maka akan diganti dengan saham lain yang memenuhi syarat. Pemilihan saham - saham LQ 45 harus wajar, oleh karena itu BEJ mempunyai komite penasehat yang terdiri dari para ahli di BAPEPAM, Universitas, dan Profesional di bidang pasar modal. (factbook 1997, Jakarta Stock Exchange).

Faktor –faktor yang berperan dalam pergerakan Indeks LQ 45 yaitu:

1. Tingkat suku bunga SBI sebagai patokan (benchmark) portofolio investasi di pasar keuangan Indonesia,
2. Tingkat toleransi investor terhadap risiko, dan
3. Saham – saham penggerak indeks (index mover stocks) yang notabene merupakan saham berkapitalisasi pasar besar di BEJ.

Faktor – faktor yang berpengaruh terhadap naiknya Indeks LQ 45 adalah:

1. Penguatan bursa global dan regional menyusul penurunan harga minyak mentah dunia.
2. Penguatan nilai tukar rupiah yang mampu mengangkat indeks LQ 45 ke zone positif.

Tujuan indeks LQ 45 adalah sebagai pelengkap IHSG dan khususnya untuk menyediakan sarana yang obyektif dan terpercaya bagi analisis keuangan, manajer investasi, investor dan pemerhati pasar modal lainnya dalam memonitor pergerakan harga dari saham-saham yang aktif diperdagangkan.

2.10 Definisi Variabel

2.10.1 Return saham

Return saham adalah tingkat yang dinikmati oleh pemodal atas suatu investasi yang dilakukannya (Akhamd Firdaus:2013). Setiap investasi baik jangka panjang maupun jangka pendek mempunyai tujuan utama untuk mendapatkan keuntungan yang disebut return, baik langsung maupun tidak langsung.

2.10.2 Net Profit Margin (NPM)

Net Profit Margin (NPM) adalah rasio yang digunakan untuk mengukur tingkat kembalian keuntungan bersih terhadap penjualan bersihnya. Semakin tinggi rasio *Net Profit Margin* (NPM) berarti laba yang dihasilkan oleh perusahaan juga semakin besar maka akan menarik minat investor untuk melakukan transaksi dengan perusahaan yang bersangkutan, karena secara teori jika kemampuan emiten dalam menghasilkan laba semakin besar maka harga saham perusahaan di pasar modal juga akan mengalami peningkatan (Yeye:2011).

2.10.3 Gross Profit Margin (GPM)

Rasio *Gross Profit Margin* (GPM) mencerminkan atau menggambarkan laba kotor yang dapat dicapai setiap rupiah penjualan, atau bila rasio ini dikurangkan terhadap angka 100% maka akan menunjukkan jumlah yang tersisa untuk menutup biaya operasi dan laba bersih. Data *Gross Profit Margin* (GPM) rasio dari beberapa periode akan dapat memberikan informasi tentang kecenderungan *Gross Profit Margin* (GPM) rasio yang diperoleh dan bila dibandingkan standar rasio akan diketahui apakah *margin* yang diperoleh perusahaan sudah tinggi atau sebaliknya.

2.10.4 *Current Ratio* (CR)

Current Ratio (CR) adalah rasio yang sangat berguna untuk mengukur kemampuan perusahaan dalam melunasi kewajiban-kewajiban jangka pendeknya, dimana dapat diketahui sampai seberapa jauh sebenarnya jumlah aktiva lancar perusahaan dapat menjamin utang lancarnya. Semakin tinggi rasio berarti terjamin utang-utang perusahaan kepada kreditur. Rasio ini digunakan untuk mengukur kemampuan perusahaan dalam membayar kewajiban jangka pendek dengan menggunakan aktiva lancar yang dimiliki. Berikut ini rumus untuk menghitung current rasio.

2.10.5 *Return On Asset* (ROA)

Return On Asset(ROA) merupakan salah satu rasio profitabilitas yang digunakan untuk mengukur efektifitas perusahaan dalam mendapatkan keuntungan dengan memanfaatkan semua aktiva yang dimilikinya. Bila kemampuan perusahaan dalam menghasilkan laba tergolong tinggi, maka harga saham juga akan mengalami peningkatan yang akan berdampak pada peningkatan

return saham di masa yang akan datang. Semakin tinggi ROA maka akan menunjukkan semakin efisien operasional dari suatu perusahaan, begitu pula sebaliknya, ROA yang rendah dapat disebabkan oleh banyaknya aset yang menganggur, kemudian investasi dalam persediaan yang terlalu banyak. Kelebihan uang kertas, aset tetap beroperasi dibawah normal, dan lain-lain. Semakin meningkatnya ROA maka kinerja perusahaan yang ditinjau dari profitabilitasnya akan semakin baik.

2.10.6 Return On Equality (ROE)

Return On Equality (ROE) merupakan ukuran kemampuan perusahaan (emiten) dalam menghasilkan keuntungan dengan menggunakan modal sendiri, sehingga ROE ini sering disebut sebagai rentabilitas modal sendiri. Sebagaimana ROA, maka semakin tinggi ROE juga menunjukkan kinerja perusahaan semakin baik dan berdampak pada meningkatnya harga saham perusahaan. Jika harga saham semakin meningkat maka return saham juga akan meningkat, maka secara teoritis sangat dimungkinkan ROE berpengaruh positif terhadap return saham.

2.10.7 Debt to Equity Ratio (DER)

Debt to Equity Ratio (DER) merupakan perbandingan antara seluruh hutang perusahaan, baik hutang jangka panjang maupun hutang jangka pendek dengan modal sendiri yang dimiliki perusahaan. Semakin tinggi DER maka akan menunjukkan komposisi total hutang yang semakin besar dibandingkan dengan total modal sendiri sehingga akan meningkatkan tingkat resiko investor karena hal tersebut akan berdampak pada menurunnya harga saham.

2.11 Penelitian Terdahulu

Sebagai referensi dan perbandingan dalam penelitian ini peneliti akan mengemukakan bahan penelitian terdahulu yang pembahasannya atau topik sesuai permasalahan dalam penelitian yang akan dilaksanakan. Adapun referensi yang ditulis oleh:

Pusparani, D. Dalam penelitiannya menjelaskan bahwa berdasarkan hasil perbandingan antara regresi *ridge* dan regresi *stepwise*, data yang memiliki tingkat multikolinieritas sedang dalam menangani masalah multikolinieritas lebih baik menggunakan regresi *ridge*. Sedangkan data yang memiliki tingkat multikolinieritas sangat kuat dalam menangani masalah multikolinieritas lebih baik menggunakan regresi *stepwise*.

Soemartini (2008) dalam penelitiannya menjelaskan bahwa berdasarkan hasil analisis dari contoh kasus sebelumnya, dapat disimpulkan bahwa metode *Principal Component Analysis* (PCA) terbukti dapat mengatasi masalah pelanggaran asumsi klasik multikolinieritas tanpa perlu membuang variabel bebas yang berkolinier tinggi. Sehingga setelah diperoleh variabel bebas baru dari hasil reduksi, kita dapat meramalkan pengaruh dari variabel bebas (pendapatan) terhadap variabel tak bebas (konsumsi) melalui analisis regresi linier. Dengan metode PCA, kita akan mendapatkan variabel bebas baru yang tidak berkorelasi, bebas satu sama lainnya, lebih sedikit jumlahnya daripada variabel asli, akan tetapi bisa menyerap sebagian besar informasi yang terkandung dalam variabel asli atau yang bisa memberikan kontribusi terhadap varian seluruh variabel.

Masruroh, Illati dalam penelitiannya menjelaskan bahwa hasil perbandingan metode Komponen Utama (*Principal Component Regression*) dan Regresi Gulud

(*Ridge Regression*) pada data yang tingkat multikolinieritas tinggi lebih efektif menggunakan metode regresi komponen utama.

Tazliqoh, A.Z. *et al* (2015) dalam penelitiannya menjelaskan bahwa berdasarkan perhitungan menggunakan nilai standar *error* menunjukkan bahwa penanganan multikolinieritas menggunakan PCA (*Principal Component Analysis*) lebih baik dibandingkan dengan *Ridge Regression* dalam analisis faktor-faktor Pendapatan Asli Daerah (PAD) Provinsi Jawa Tengah.

Nurhasanah *et al* (2012) dalam penelitiannya menjelaskan bahwa Metode *Partial Least Square* (PLS) memberikan hasil yang lebih baik jika dibandingkan dengan metode regresi komponen utama. Hal ini dapat disimpulkan dengan melihat nilai R^2 , *Mean Square Error Prediction* (MSEP), dan *Root Mean Square Error Prediction* (RMSEP). Metode *Partial Least Square* (PLS) mempunyai nilai R^2 yang lebih tinggi dan mempunyai nilai MSEP dan RMSEP yang lebih rendah jika dibandingkan terhadap metode regresi komponen utama atau PCA (*Principal Component Analysis*).

Hanum, H. (2011) dalam penelitiannya menjelaskan metode *Stepwise* dan *Best Subset Regression* tidak mempertimbangkan masalah multikolinieritas untuk mencari model terbaik sedangkan metode Fraksi lebih fokus untuk menghindari masalah multikolinieritas.

Berdasarkan penelitian terdahulu yang telah diuraikan di atas masing-masing metode memiliki kelebihan masing-masing diantaranya metode *Stepwise* dan metode *Principal Component Analysis* (PCA) memiliki kelebihan dalam mencari model terbalik regresi linier berganda pada kasus multikolinieritas yang memiliki

tingkat multikolinieritas sangat kuat. Metode *Principal Component Analysis* (PCA) dapat menyembuhkan kasus multikolinieritas secara bersih tanpa menghilangkan variabel bebas atau peubah, sedangkan metode *Stepwise* pemilihan variabel terbaik yang masuk dalam model dengan cara memilih peubah berdasarkan korelasi parsial terbesar dengan peubah yang sudah masuk. Metode *Partial Least Square* (PLS) menghasilkan model yang mengoptimalkan hubungan antara dua kelompok variabel. Sedangkan metode *Ridge Regression* ditujukan untuk mengatasi kondisi buruk (*ill conditioned*) yang diakibatkan oleh korelasi yang tinggi antara beberapa peubah penjelas di dalam model. Metode *Ridge Regression* dapat digunakan dalam penanganan kasus multikolinieritas yang memiliki tingkat multikolinieritas sedang.

Berdasarkan uraian penelitian terdahulu dapat ditulis dalam bentuk Tabel 2.2 seperti berikut:

Tabel 2.2 Penelitian Terdahulu

No	Peneliti	Metode	Sampel	Hasil Penelitian
1.	Pusparini, D.E	<i>Ridge Regression</i> dan <i>Stepwise</i>	Data 1: Ekspor kopi Indonesia periode 1975-1990. Data 2: Kekuatan tekana semen PT. Semen Gresik (Persero) Tbk. Periode Januari-Februari 1999. Data 3: Hasil produksi padi dan variabel yang mempengaruhinya.	Metode <i>Stepwise</i> untuk kasus multikolinieritas kuat sedangkan metode <i>Ridge Regression</i> untuk kasus multikolinieritas sedang.
2.	Soemartini (2008)	<i>Principal Component Analysis</i> (PCA)	Data keadaan ekonomi di Amerika Serikat tahun 1936-1952	Metode <i>Principal Component Analysis</i> dapat menyembuhkan multikolinieritas kuat.

3. Masruroh, Illati Metode *Principal Component Analysis* dan metode *Ridge Regression* Data 1: Mengenai kebugaran aerobik Data 2: Survey terhadap permintaan ayam di Amerika Serikat tahun 1982-2004. Metode *Principal Component Analysis* (PCA) lebih cocok untuk masalah multikolinieritas kuat.
4. Tazliqoh, A.Z., et al (2015) Metode *Principal Component Analysis* (PCA) dan metode *Ridge Regression* Pendapatan asli Daerah (PAD) Provinsi Jawa Tengah Metode *Principal Component Analysis* (PCA) lebih baik dibandingkan metode *Ridge Regression* untuk menganalisis faktor-faktor PAD.
5. Nursanah et al(2012) Metode *Partial Least Square* (PLS) dan metode *Principal Component Analysis* Data dengan 8 variabel bebas Metode *Partial Least Square* (PLS) lebih baik dibandingkan dengan metode *Principal Component Analysis* (PCA) karena mempunyai nilai R^2 lebih besar dan RMSEP lebih kecil.

No	Peneliti	Metode	Sampel	Hasil Penelitian
6.	Hanum, H. (2011)	Metode <i>Stepwise</i> , <i>Best Subset Regression</i> , dan Fraksi	Data dengan 6 variabel bebas dan 8 variabel bebas.	metode <i>Stepwise</i> dan <i>Best Subset Regression</i> tidak mempertimbangkan masalah multikolinieritas untuk mencari model terbaik sedangkan metode Fraksi lebih fokus untuk menghindari masalah multikolinieritas.

2.12

Kerangka Pemikiran

Menurut Uma Sekaran, dalam Sugiyono, (1997) mengemukakan bahwa kerangka pemikiran merupakan model konseptual tentang bagaimana teori berhubungan dengan berbagai faktor yang telah diidentifikasi sebagai masalah yang penting. Kerangka pemikiran yang baik akan menjelaskan secara teoritis hubungan antara variabel yang akan diteliti sampai menjawab pertanyaan secara teoritis.

Analisis regresi merupakan salah satu metode statistik yang sering digunakan dalam kehidupan sehari-hari. Tujuan analisis regresi yaitu untuk mengetahui sejauh mana hubungan sebuah variabel bebas dengan beberapa variabel tak bebas, bila analisisnya melibatkan lebih dari satu atau beberapa variabel bebas, maka analisis yang digunakan adalah analisis regresi linier berganda. Dalam menentukan model regresi, variabel bebas dapat masuk dalam model secara bersama-sama atau satu persatu. Jika variabel bebas masuk dalam model secara bersama-sama maka perhitungan akan ringkas, akan tetapi tidak akan kelihatan apa yang terjadi dalam perhitungan tersebut karena setiap variabel bebas yang masuk memberikan pengaruh yang berbeda, tergantung pada urutan variabel bebas tersebut yang masuk dalam model. Namun tidak berarti semua variabel yang masuk dalam model regresi menjadikan model tersebut model yang terbaik (Sembiring, 1995).

Model regresi terbaik adalah model yang dapat menjelaskan perilaku peubah tak bebas dengan sebaik-baiknya dengan memilih peubah-peubah bebas dari sekian banyak peubah bebas yang tersedia dalam data. Untuk menentukan peubah bebas mana yang akan dimasukkan ke dalam model regresi. Untuk mengetahui

model regresi yang baik maka perlu dilakukan uji asumsi klasik, namun pada uji asumsi klasik tak semuanya mulus dan didapat model yang baik. Apalagi jika model regresi pada regresi linier berganda, tentu banyak masalah yang menjadikan model regresi tidak baik atau signifikan salah satunya adalah masalah multikolinieritas. Jika ada masalah multikolinieritas maka kesimpulan yang didapat dari hasil pengujian untuk model regresi maupun masing-masing peubah yang ada dalam model seringkali tidak tepat. Oleh sebab itu masalah multikolinieritas harus dihindari. Untuk pemilihan model terbaik dapat menggunakan metode *Stepwise* yaitu metode dengan cara mengeluarkan variabel yang berkorelasi dan metode *Principal Component Analysis* (PCA) yaitu dengan cara menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali. Setelah beberapa komponen hasil PCA yang bebas multikolinieritas diperoleh, maka komponen-komponen tersebut menjadi variabel bebas baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas (Y) dengan menggunakan analisis regresi. Keunggulan metode PCA diantaranya adalah dapat menghilangkan korelasi secara bersih tanpa harus mengurangi jumlah variabel asal.

Setelah diperoleh model dari kedua metode tersebut maka untuk mencari metode terbaik dari kedua metode tersebut kita dapat membandingkan R^2 Adjusted dan S^2 yang bertujuan untuk menentukan proporsi atau persentase total variasi dalam variabel terikat yang diterangkan oleh variabel bebas dengan menggunakan data yang mengandung multikolinieritas kuat.

BAB 5

PENUTUP

5.1 Simpulan

Berdasarkan uraian yang telah dijelaskan pada bab sebelumnya, maka dapat disimpulkan sebagai berikut:

- 1) Metode *Principal Component Analysis* (PCA) digunakan untuk mengatasi multikolinieritas yang bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali. Setelah beberapa komponen hasil PCA yang bebas multikolinieritas diperoleh, maka komponen-komponen tersebut menjadi variabel bebas baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas (Y) dengan menggunakan analisis regresi. Metode PCA dilakukan dengan menggunakan bantuan analisis faktor dalam SPSS. Sedangkan metode *Stepwise* pada dasarnya digunakan untuk mencari model terbaik yaitu dengan cara memasukkan satu persatu variabel yang signifikan, namun pada data *return* saham Perusahaan dalam Indeks LQ 45 di BEI periode Juli –Desember 2015 yang mengandung multikolinieritas dapat disembuhkan dengan metode *stepwise* dengan bantuan SPSS.
- 2) Setelah dibandingkan antara metode *Principal Component Analysis* (PCA) dan *Stepwise* dengan membandingkan nilai S^2 atau nilai rata-rata kuadrat sisa

dan R^2 adjusted diperoleh hasil data yang disimulasikan dengan menggunakan metode *Stepwise* memiliki nilai S^2 lebih kecil dan R^2 adjusted lebih besar dibandingkan dengan metode *Principal Component Analysis* (PCA), sehingga dapat disimpulkan bahwa metode *Stepwise* lebih baik dibandingkan metode *Principal Component Analysis* (PCA) untuk mencari model terbaik pada kasus multikolinieritas.

5.2 Saran

Saran yang dapat peneliti sumbangkan sehubungan dengan hasil penelitian ini untuk penelitian selanjutnya adalah sebagai berikut:

- 1) Jika pada suatu model regresi terjadi penyimpangan asumsi multikolinieritas, maka harus dilakukan tindakan perbaikan untuk menghilangkan multikolinieritas tersebut.
- 2) Bila mencari model terbaik pada kasus multikolinieritas sebaiknya menggunakan metode *Stepwise* karena lebih efektif dibanding metode *Principal Component Analysis* (PCA).
- 3) Penelitian pada data *return* saham hanya menggunakan 6 variabel bebas penyelesaian dengan metode *Principal Component Analysis* (PCA) hanya menghasilkan 1 variabel faktor atau sering disebut variabel pengganti dari variabel bebas sebuah pengamatan. Sehingga pembaca dapat menggunakan variabel yang lebih banyak agar faktor yang terbentuk lebih dari satu.

DAFTAR PUSTAKA

- Anton, Howard. 1992. *Aljabar Linear Elementer*. Jakarta: Penerbit Erlangga Companies.
- Affinanda, A. 2015. *Analisis Pengaruh Rasio Keuangan Terhadap Return Saham Perusahaan Dalam Indeks Lq 45 Tahun 2010-2013*. Skripsi. Jurusan Akuntansi. FE. Universitas Diponegoro. Tersedia di eprints.undip.ac.id/45469/1/15_AFFINANDA.pdf [diakses 23-7-2016].
- Draper, N. R., & Smith, H. 1992. *Analisis Regresi Terapan*. Jakarta: Gramedia Pustaka Utama.
- Gujarati, D. N. 2004. *Basic Econometrics* (4th ed). New York: The McGraw-Hill.
- Herlina, H. 2011. Perbandingan Metode Stepwise, Best Subset Regression, dan Fraksi dalam Pemilihan Model Regresi Berganda Terbaik. *Jurusan Matematika, F.MIPA, Universitas Sriwijaya*. Tersedia di <http://statistik.studentjournal.ub.ac.id>. [diakses 20-4-2015].
- Ifadah, Ana. 2011. *Analisis Metode Principal Component Analysis (Komponen Utama) dan Regresi Ridge dalam Mengatasi Dampak Multikolinearitas dalam Analisis Regresi Linear Berganda*. *Jurusan Matematika, F.MIPA, Universitas Negeri Semarang*.
- LEMBANG, F.K. 2011. Analisis Regresi Berganda Dengan Metode Stepwise Pada Data Hbat. *Jurusan Matematika, F.MIPA, Unpatti*. Tersedia di ejournal.unpatti.ac.id [diakses 15-4-2015].
- Martin Charlton, dkk. 2010. *Principal Components Analysis: from Global to Local*. *University of Leicester*. Tersedia di <http://eprints.maynoothuniversity.ie>. [diakses 26-4-2015].
- Masruroh, Illati. *Pemilihan Model Regresi Linier Berganda Pada Kasus Multikolinieritas Dengan Menggunakan Metode Regresi Komponen Utama (Principal Component Analysis) Dan Regresi Gulud (ridge Regression)*.
- Pakpahan, DF. L, *et al.* 2013. Penggunaan Metode Stepwise Forward Untuk Menentukan Persamaan Regresi Linier Berganda. Tersedia di <http://jurnal.usu.ac.id/index.php/smatematika/article/view/4723>. [dikases 10 – 1 -2015]
- Parta, N. I, *et al.* Aplikasi Analisis Regresi Komponen Utama Terhadap Faktor-Faktor Yang Mempengaruhi Penyakit Diabetes Mellitus. *Universitas Negeri Malang*.

- Prasetyo, *et al.* Analisis Regresi Komponen Utama Untuk Mengatasi Masalah Multikolinieritas Dalam Analisis Regresi Linier Berganda. *Jurusan Matematika. Fmipa. Universitas Negeri Jakarta.*
- Purnomoadi, H., *et al.* 2010. Pemodelan *Statistical Downscaling* Luaran Gcm Dengan Pendekatan Regresi *Kontinum* Dan Pra-Pemrosesan PCA. *Jurusan Statistika, ITS Surabaya.* Tersedia di digilib.its.ac.id/public/ITS-Undergraduate-12368-Abstract_id.pdf [diakses 10-4-2015].
- Pusparani, D.E. 2014. Perbandingan Metode *Stepwise* Dan *Ridge Regression* Dalam Menentukan Model Regresi Berganda Terbaik Pada Kasus Multikolinieritas. *Jurusan Matematika, F.MIPA, Universitas Brawijaya.* Tersedia di <http://statistik.studentjournal.ub.ac.id/index.php/statistik/article/view/169> [diakses 10-4-105].
- Pusparini, D.E. Perbandingan Metode *Stepwise* Dan *Ridge Regression* Dalam Menentukan Model Regresi Berganda Terbaik Pada Kasus Multikolinearitas.
- Santoso, Singgih. 2010. *Statistika Multivariat: Konsep dan Aplikasi dengan SPSS.* Jakarta: Elex Media Komputindo.
- Sembiring, R.K. 1995. *Analisis Regresi.* Bandung: Penerbit ITB.
- Soemartini. 2008. Principal Component Analysis (PCA) Sebagai Salah Satu Metode Untuk Mengatasi Masalah Multikolinearitas. *Jurusan Statistika, F.MIPA, Universitas Padjadjaran.* Tersedia di <https://elmurobbie.files.wordpress.com/> [diakses 15-4-2015].
- Subekti, P. 2015. Perbandingan Metode Best Subset Dan Stepwise Untuk Mengetahui Pengaruh Tingkat Pendidikan Terhadap Pengangguran Di Jawa Timur. *Jurusan Teknologi dan Informatika. STIMIK Asia Malang.* Tersedia di www.scribd.com/doc/166097148 [diakses 4-4-2015].
- Tazliqoh, A. Z., *et al.* 2015. Perbandingan Regresi Komponen Utama Dengan Regresi Ridge Pada Analisis Faktor-Faktor Pendapatan Asli Daerah (PAD) Provinsi Jawa Tengah. *Jurusan Statistika. FSM. Universitas Diponegoro.* Tersedia di <http://ejournal-s1.undip.ac.id/index.php/gaussian> [diakses 4-4-2015].

Wiratmanto. 2014. Analisis Faktor Dan Penerapannya Dalam Mengidentifikasi Faktor – Faktor Yang Mempengaruhi Kepuasan Konsumen Terhadap Penjualan Media Pembelajaran. Skripsi. *Jurusan Matematika. FMIPA. Universitas Negeri Yogyakarta.* Tersedia di eprints.uny.ac.id/.../Skripsi_Analisis_Faktor_Wiratmanto_0730514... [diakses 23-7-2015].

