



TUGAS AKHIR - KS184822

**PREDIKSI INDEKS BATUBARA DI PT XYZ
MENGUNAKAN *RIDGE REGRESSION* DAN
*SUPPORT VECTOR REGRESSION (SVR)***

**RIZKY AMALIA PUTRI
NRP 062116 4000 0040**

**Dosen Pembimbing
Dra. Wiwiek Setya Winahju, M.S.
Dr. Muhammad Mashuri, M.T.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**



TUGAS AKHIR - KS184822

**PREDIKSI INDEKS BATUBARA DI PT XYZ
MENGUNAKAN RIDGE REGRESSION DAN
SUPPORT VECTOR REGRESSION (SVR)**

**RIZKY AMALIA PUTRI
NRP 062116 4000 0040**

**Dosen Pembimbing
Dra. Wiwiek Setya Winahju, M.S.
Dr. Muhammad Mashuri, M.T.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN ANALITIKA DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**



FINAL PROJECT - KS184822

**COAL INDEX PREDICTION IN PT XYZ USING RIDGE
REGRESSION AND SUPPORT VECTOR REGRESSION
(SVR)**

**RIZKY AMALIA PUTRI
NRP 062116 4000 0040**

**Supervisors
Dra. Wiwiek Setya Winahju, M.S.
Dr. Muhammad Mashuri, M.T.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF SCIENCE AND DATA ANALYTICS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2020**

LEMBAR PENGESAHAN

PREDIKSI INDEKS BATUBARA DI PT XYZ MENGUNAKAN *RIDGE REGRESSION* DAN *SUPPORT VECTOR REGRESSION (SVR)*

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember

Oleh :

Rizky Amalia Putri
NRP. 062116 4000 0040

Disetujui oleh Pembimbing:

Dra. Wiwiek Setya Winahju, M.S.

NIP. 19560424 198303 2 001

Muhammad Mashuri, M.T.

NIP. 19620408 198701 1 001

Styanjiz
(*Ressing*)



Mengetahui,
Kepala Departemen Statistika

Dr. Dra. Kartika Fithriasari, M.Si.

NIP. 19691212 199303 2 002

SURABAYA, JANUARI 2020

(Halaman ini sengaja dikosongkan)

**PREDIKSI INDEKS BATUBARA DI PT XYZ
MENGUNAKAN RIDGE REGRESSION DAN *SUPPORT
VECTOR REGRESSION (SVR)***

Nama Mahasiswa : Rizky Amalia Putri
NRP : 062116 4000 0040
Departemen : Statistika
Dosen Pembimbing I : Dra. Wiwiek Setya Winahju, M.S.
Dosen Pembimbing II : Dr. Muhammad Mashuri, M.T.

Abstrak

Semen merupakan salah satu bahan baku yang amat penting dalam pembangunan infrastruktur. Salah satu perusahaan yang bergerak di bidang produksi semen adalah PT XYZ. Tahapan terpenting dalam proses pembuatan semen adalah pada tahap pembakaran batu kapur dan tanah liat (clinker). Dalam proses pembakaran clinker membutuhkan bahan bakar utama yaitu batubara. Semakin banyak jumlah produksi clinker yang dihasilkan dan semakin sedikit batubara yang digunakan dalam proses pembakaran, maka semakin efektif dan efisien proses produksi tersebut. Dalam penelitian ini dilakukan analisis untuk memprediksi indeks batubara dengan beberapa variabel yang diduga mempengaruhi yaitu kualitas batubara, bahan baku, dan operasional yang kemudian dilakukan estimasi terhadap indeks batubara. Metode yang digunakan untuk mengestimasi indeks batubara adalah metode Regresi Ridge dan metode Support Vector Regression (SVR). Sebelum dilakukan SVR, perlu dilakukan pemilihan variabel atau Feature Selection dengan metode Recursive Feature Elimination berbasis Random Forest. Model yang terbentuk dengan metode SVR dibandingkan dengan metode regresi ridge, kemudian dipilih model terbaiknya diantara kedua model yang terbentuk menggunakan nilai RMSE. Hasil analisis didapatkan metode terbaik dengan nilai RMSE terkecil yaitu Support Vector Regression (SVR) dengan complete feature dan menggunakan kernel-polynomial yang menghasilkan parameter sigma bernilai 0,100 dan nilai c sebesar 1 dengan nilai RMSE sebesar 0,619.

Kata kunci: *Batubara, Clinker, Indeks, Ridge, Support Vector Regression.*

(Halaman ini sengaja dikosongkan)

COAL INDEX PREDICTION IN PT XYZ USING RIDGE REGRESSION AND SUPPORT VECTOR REGRESSION (SVR)

Name : Rizky Amalia Putri
Student Number : 062116 4000 0040
Department : Statistics
Supervisor I : Dra. Wiwiek Setya Winahju, M.S.
Supervisor II : Dr. Muhammad Mashuri, M.T.

Abstract

Cement is one of the most important raw materials in infrastructure development. One of the companies engaged in cement production is PT XYZ. The most important stage in the process of making cement is at the stage of burning limestone and clay (clinker). In the process of burning clinker requires the main fuel, namely coal. The more clinker production is produced and the less coal used in the combustion process, the more effective and efficient the production process. In this study analysis will be conducted to predict the coal indeks with several variables that are thought to affect the quality of coal, raw materials, and operations which will then be estimated on the coal indeks. The method used to estimate the coal indeks is Ridge Regression method and Support Vector Regression (SVR) method. Prior to the SVR, a variable or Feature Selection will be conducted using the Random Forest-based Recursive Feature Elimination method. The model formed by the SVR method will be compared with the ridge regression method which will then be chosen the best model between the two models formed using the RMSE value. The analysis results obtained the best method with the smallest RMSE value, is Support Vector Regression (SVR) with complete features and using kernel-polynomial that produces a sigma parameter of 0.100 and a c value of 1 with an RMSE value of 0.619.

Kata kunci: *Batubara, Clinker, Indeks, Ridge, Support Vector Regression.*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “Prediksi Indeks Batubara di PT XYZ Menggunakan *Ridge Regression* dan *Support Vector Regression* (SVR)” dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Institut Teknologi Sepuluh Nopember, khususnya Departemen Statistika ITS atas segala ilmu pengetahuan dan pengalaman selama saya kuliah.
2. Ibu Dra. Kartika Fithriarsari, M.Si. selaku Ketua Departemen Statistika, Ibu Dr. Santi Wulan Purnami, S.Si., M.Si. selaku Sekretaris Departemen Statistika Bidang Akademik, dan Ibu Dr. Vita Ratnasari, S.Si., M.Si. selaku Sekretaris Departemen Statistika Bidang Keuangan yang telah memberikan fasilitas, sarana, dan prasarana.
3. Ibu Dra. Wiwiek Setya Winahju, M.S. dan Bapak Dr. Muhammad Mashuri, M.T. selaku dosen pembimbing yang telah meluangkan waktu dan dengan sangat sabar memberikan bimbingan, ilmu, saran, dukungan serta motivasi selama penyusunan Tugas Akhir saya.
4. Bapak Prof. Drs. Nur Iriawan, MIKom., Ph.D. dan Ibu Dr. Vita Ratnasari, S.Si., M.Si. selaku dosen penguji yang telah memberikan koreksi dan masukan dengan sangat sabar sehingga Tugas Akhir ini dapat lebih baik.
5. Bapak Dr. Sutikno, S.Si., M.Si. selaku dosen wali selama masa studi yang telah banyak memberikan saran, arahan dan semangat dalam menjalani perkuliahan.
6. Seluruh dosen Departemen Statistika ITS yang telah memberikan ilmu dan pengetahuan yang tak ternilai harganya, serta segenap karyawan Departemen Statistika ITS khususnya Pak Umam dan Pak Pendi yang selalu siap siaga dalam memenuhi keperluan administrasi selama kuliah.

7. PT XYZ selaku instansi sumber data Tugas Akhir saya dan Bapak Imron Gozali S.T. selaku pembimbing saya selama magang di PT XYZ yang telah banyak memberikan ilmu, pengalaman, dan saran untuk Tugas Akhir saya.
8. Bapak M. Chasan Z. dan Ibu Yayik Mahmudah tercinta yang sudah mendoakan tiada henti dan mendukung baik secara moril maupun materiil. Serta adik Arshy Ahsanal Khafidoh tersayang yang selalu memberikan semangat hingga saat ini.
9. Teman-teman Statistika ITS angkatan 2016 khususnya Erika Chance, Arin Berliana, Zevira Fitri, Cahya Idznii, Fransiska, Kinanthi, Inan, dan Thalia yang selalu memberikan dukungan dan saran selama hampir 4 tahun ini.
10. Marita Qori'atunnadyah, teman berjuang dalam menyelesaikan Tugas Akhir dan mencari dosen pembimbing. Sebagai tempat berbagi semangat, keluhan, dan saran.
11. Al'awwa Prima Nadia Putri, selaku partner Kerja Praktik saya di BMKG Malang, partner survey, dan teman dalam hal apapun.
12. Finola Trisnisa yang selalu memberikan dukungan kepada penulis selama ini, sekaligus teman jalan dan bercerita.
13. Teman-teman SMA Expelliarmus, khususnya Hirani Nur Nuzuli yang menjadi tempat berbagi semangat, cerita, saran, pengalaman, dan kajian-kajian akhir zaman. Semoga dilancarkan skripsinya.
14. Teman-teman dan sahabat yang sudah memberikan semangat dan doa hingga saat ini.

Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Januari 2020

Penulis

DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN.....	v
Abstrak.....	vii
Abstract.....	ix
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
DAFTAR GAMBAR.....	xv
DAFTAR TABEL.....	xvi
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Analisis Regresi Linier dan <i>Ridge</i>	7
2.1.1 Asumsi Regresi.....	8
2.1.2 Regresi <i>Ridge</i>	9
2.1.4 Pengujian Hipotesis.....	10
2.2 <i>Support Vector Regression</i> (SVR).....	11
2.2.1 Fungsi Kernel.....	18
2.2.2 <i>Recursive Feature Elimination</i> (RFE).....	20
2.3 Kebaikan Model.....	21
2.4 Proses Produksi Semen.....	21
2.4.1 Indeks Batubara.....	24
BAB III METODOLOGI PENELITIAN.....	25
3.1 Sumber Data.....	25
3.2 Variabel Penelitian.....	25
3.4 Struktur Data.....	27
3.5 Langkah Analisis.....	28
3.6 Diagram Alir Penelitian.....	29

BAB IV ANALISIS DAN PEMBAHASAN.....	31
4.1 Eksplorasi Indeks Batubara dengan Tiga Kelompok Prediktor	31
4.2 Pemodelan Menggunakan <i>Ridge Regression</i>	34
4.2.1 Tahapan Pemodelan Indeks Batubara Menggunakan OLS	34
4.3 <i>Support Vector Regression (SVR)</i>	42
4.3.1 SVR Dengan <i>Complete Feature</i>	43
4.3.2 SVR Dengan <i>Feature Selection</i>	45
4.4 Pemilihan Model Indeks Batubara Terbaik.....	49
BAB V KESIMPULAN DAN SARAN.....	53
5.1 Kesimpulan.....	53
5.2 Saran	54
DAFTAR PUSTAKA	55
LAMPIRAN	57

DAFTAR GAMBAR

Halaman

Gambar 2.1 Ilustrasi Support Vector Regression	12
Gambar 2.2 Insensitive Zone (a), Original Input Space, dan (b) Feature Space	12
Gambar 2.3 (a) SVR Output, dan (b) Insensitive Loss Function	14
Gambar 2.4 Tipe Lost Function: (a) Linier, (b) Kuadrat, dan (c) Huber	15
Gambar 2.5 Regresi Non Linier.....	18
Gambar 2.6 Proses Pembaaran Clinker di Kiln	23
Gambar 3.1 Diagram Alir	29
Gambar 4.1 Pola Data Indeks Batubara PT XYZ.....	31
Gambar 4.2 Asumsi IIDN Data Indeks Batubara	36
Gambar 4.3 Grafik VIF dengan Berbagai Nilai c	39
Gambar 4.4 Prediksi Indeks Batubara	50

DAFTAR TABEL

Halaman

Tabel 3.1 Variabel Penelitian.....	25
Tabel 3.2 Struktur Data.....	27
Tabel 4.1 Deskriptif Indeks Batubara dan Tiga Kelompok Prediktor	32
Tabel 4.2 Hasil Uji Serentak (ANOVA).....	34
Tabel 4.3 Deteksi Multikolinieritas	35
Tabel 4.4 Nilai VIF dengan Berbagai Nilai c	37
Tabel 4.5 Koefisien Regresi Ridge dengan Berbagai Nilai c.....	39
Tabel 4.6 Tuning Parameter Kernel-Polynomial	43
Tabel 4.7 Tuning Parameter Kernel-Polynomial (Lanjutan)	44
Tabel 4.8 Tuning Parameter Kernel-RBF.....	45
Tabel 4.9 Ranking RF-RFE	46

DAFTAR LAMPIRAN

Halaman

Lampiran 1. Data Indeks Batubara Beserta Variabel Prediktor.	57
Lampiran 2. Syntax Matlab Menghitung MSE dan SSE Regresi Ridge	76
Lampiran 3. Nilai VIF dengan Berbagai Nilai c.....	77
Lampiran 4. Koefisien Regresi Ridge dengan Berbagai Nilai c	80
Lampiran 5. Perhitungan Transformasi ke Data Awal	84
Lampiran 6. Syntax R Support Vector Regression (SVR)	87
Lampiran 7. Syntax R Recursive Feature Elimination (RFE) ...	87
Lampiran 8. Prediksi SVR Menggunakan Metode Terbaik	88
Lampiran 9. Rumus Statistika Deskriptif	89
Lampiran 10. Surat Permohonan Ijin Memperoleh Data.....	90
Lampiran 11. Surat Keterangan Permintaan Data	91

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Suatu ciri negara berkembang adalah pembangunan infrastruktur yang dilakukan terus menerus di seluruh wilayah negeri dengan tujuan sebagai bentuk acuan pertumbuhan ekonomi dan mengurangi kesenjangan pembangunan antar wilayah. Perkembangan bidang pembangunan juga selalu berkembang sejalan dengan semakin banyaknya penduduk yang berada di Indonesia. Kebutuhan pembangunan infrastruktur yang paling besar adalah semen. Semen merupakan salah satu bahan baku komoditas strategis yang amat penting dalam kehidupan pembangunan manusia modern. Pembangunan manusia modern identik dengan pembangunan infrastruktur seperti pelabuhan, jalan, jembatan, bendungan, rumah, sekolah, dan lainnya (Hidayat, 2009).

Salah satu perusahaan yang bergerak di bidang produksi semen adalah PT XYZ. PT XYZ merupakan Badan Usaha Milik Negara (BUMN) *Multinasional Company* di Indonesia. Perusahaan ini telah terdaftar di Bursa Efek Indonesia (BEI) hingga sekarang dan membawahi dari beberapa perusahaan yang merupakan jenis perusahaan manufaktur dimana perusahaan tersebut bergerak di bidang produksi semen, karena semen merupakan suatu produk yang sangat dibutuhkan dalam pembangunan infrastruktur, oleh karena itu dibutuhkan semen yang berkualitas baik dengan proses produksi yang efektif dan efisien

Proses pembuatan semen di PT XYZ terdiri dari lima tahapan, yaitu penyediaan bahan mentah, penggilingan bahan mentah (batu kapur dan tanah liat), pembakaran *clinker* (campuran dari batu kapur dan tanah liat yang telah digiling), penggilingan akhir, lalu yang terakhir adalah pengantongan sak semen/pengemasan. Tahap pembakaran batu kapur dan tanah liat (*clinker*) merupakan salah satu tahapan penting dalam proses produksi semen. Proses pembakaran *clinker* membutuhkan bahan bakar utama yaitu batubara. Batubara adalah salah satu bahan bakar fosil yang berasal dari

batuan sedimen yang dapat terbakar dan terbentuk dari endapan organik, utamanya adalah sisa-sisa tumbuhan dan terbentuk melalui proses pembatubaraan. Batubara merupakan salah satu sumber energi di Indonesia. Kementerian Energi dan Sumber Daya Mineral (ESDM) merilis data cadangan batubara Indonesia, yang kini mencapai 39,89 miliar ton. Angka ini naik dibanding sebelumnya yang sebesar 37 miliar ton pada awal 2018. Meski begitu, jumlah cadangan ini masih tergolong kecil di tingkat dunia karena masih di kisaran rata-rata 3%-4% cadangan dunia (CNBC Indonesia, 2019). Indeks batubara pada PT XYZ didapatkan dari jumlah produksi *clinker* dibagi dengan jumlah batubara dengan satuan ton, dimana semakin tinggi indeks batubara maka semakin efektif dan efisien proses pembakaran *clinker* atau dengan arti lain semakin banyak jumlah produksi *clinker* yang dihasilkan dan semakin sedikit batubara yang digunakan dalam proses pembakaran maka semakin efektif dan efisien penggunaan batubara pada proses produksi tersebut, sehingga dapat digunakan indeks batubara untuk mengukur keefektifan dan efisiensi proses produksi.

Dalam penelitian ini dilakukan analisis untuk memprediksi indeks batubara dengan variabel yang digunakan adalah variabel indeks batubara dan beberapa variabel yang diduga mempengaruhi yang kemudian dilakukan estimasi terhadap indeks batubara. Metode yang digunakan untuk mengestimasi indeks batubara adalah metode *Ridge Regression* dikarenakan adanya kasus multikolinieritas dan *Support Vector Regression* (SVR) yang merupakan pengembangan dari metode *Support Vector Machine* (SVM). *Support Vector Regression* (SVR) memiliki tujuan memetakan vektor input ke dalam dimensi yang lebih tinggi, karena seringkali teknik *machine learning* yang dikembangkan dengan asumsi kelinieran, sehingga algoritma yang dihasilkan terbatas untuk kasus yang non-linier. *Support Vector Regression* (SVR) juga digunakan karena beberapa proses produksi pada semen memiliki indikasi yang berhubungan pada variabel prediktor yang satu dengan yang lain, sedangkan pada regresi linier sederhana tidak diperbolehkan adanya hubungan antara variabel prediktor, namun pada metode SVR tidak

mempermasalahkan hal tersebut. Metode *Support Vector Regression* (SVR) juga dapat mengatasi *overfitting* dan kemampuan generalisasi yang sangat baik dengan akurasi prediksi yang tinggi. Variabel yang menunjang untuk mengestimasi ini adalah faktor operasional (*speed kiln, ILC exit temperature, SLC exit temperature, temperature stage 4 ILC, temperature stage 4 SLC, ILC coal, SLC coal, kiln coal*), bahan baku (LSF, SIM, ALM), dan kualitas batubara (AC, VM, FC, TS, TM, GHV). Sebelum dilakukan *Support Vector Regression* (SVR), dilakukan pemilihan variabel atau *Feature Selection* dengan metode *Recursive Feature Elimination* (RFE) berbasis *Random Forest*, hal ini digunakan karena terdapat banyak variabel prediktor yang diduga terdapat hubungan antara variabel prediktor satu dengan yang lainnya. Kemudian, model yang terbentuk dengan metode SVR dibandingkan dengan metode regresi *ridge* yang kemudian dipilih model terbaiknya diantara kedua model yang terbentuk.

Penelitian sebelumnya adalah Abdul (2018) yang mengangkat judul sistem pendukung keputusan pemberian bonus tetap memanfaatkan SVR bahwa pemilihan metode antara SVR dengan neural network didasari dengan nilai MSE, didapatkan bahwa nilai MSE pada SVR lebih rendah dibandingkan *neural network*. Penelitian Ita Fitriati & Muhammad Ghazali (2017) yang berjudul *Pemodelan Pengaruh Rata-rata Lama Sekolah Terhadap Indeks Kedalaman Kemiskinan di Indonesia Menggunakan Support Vector Regression*. Metode *Support Vector Regression* (SVR) digunakan untuk menaksir Indeks Kedalaman Kemiskinan (Y) dengan menggunakan variabel prediktor Rata-rata lama sekolah (X). Model *Support Vector Regression* (SVR) pada penelitian ini menggunakan beberapa kernel berbeda yaitu kernel linier, polynomial dan RBF. Kriteria model terbaik adalah model yang menghasilkan MSE yang terkecil dan R^2 terbesar. Model terbaik yang dihasilkan adalah model dengan menggunakan kernel RBF. Penelitian ini penting untuk menganalisa dengan lebih akurat terhadap faktor kemiskinan di Indonesia. Penelitian sebelumnya dengan menggunakan metode *Support Vector Regression* (SVR) pernah dilakukan

oleh Abuella dan Chowdhury (2016) mengenai peramalan tenaga surya secara bergilir selama 24 jam ke depan selama satu tahun penuh. Variabel prediktor yang digunakan berupa dua belas variabel cuaca dan variabel baru diekstraksi. Hasil peramalan dengan menggunakan (SVR) untuk selanjutnya dibandingkan dengan jaringan saraf tiruan, dan model regresi linier berganda. Berdasarkan nilai RMSE didapatkan bahwa dengan menggunakan jaringan saraf tiruan didapatkan hasil peramalan yang paling akurat, diikuti dengan SVR dan regresi linier berganda. Namun SVR memiliki keuntungan yang tidak dapat ditampilkan secara ilustratif bahwa SVR lebih kuat dibandingkan jaringan saraf tiruan karena SVR tidak mengalami masalah minimal lokal, selain itu juga SVR memiliki kelebihan dalam mengatasi masalah *overfitting*. Berdasarkan perbandingan hasil peramalan dari ketiga metode tersebut didapatkan bahwa model SVR lebih akurat dibandingkan dengan model regresi linier sederhana dan pada saat yang sama lebih kuat jika dibandingkan dengan jaringan saraf tiruan. Hendra (2014) yang menggunakan metode *feature selection* RFE menyimpulkan bahwa dengan RFE memiliki hasil RMSE yang lebih kecil dan *R-squared* lebih besar dibandingkan *complete feature*.

1.2 Rumusan Masalah

Indeks batubara adalah suatu ukuran untuk mengukur kinerja proses produksi (*clinker*) yang dihasilkan dari tahap pembakaran. Semakin banyak produksi *clinker* yang dihasilkan dan semakin sedikit batubara yang digunakan dalam proses pembakaran maka semakin efektif dan efisien proses produksi tersebut. Sebelumnya, belum pernah dilakukan penelitian terkait prediksi indeks batubara berdasarkan faktor-faktor yang diduga berpengaruh dalam proses produksi semen, maka dari itu pada penelitian ini dilakukan prediksi terhadap indeks batubara menggunakan *Ridge Regression* dan *Support Vector Regression* (SVR) karena setiap akhir bulan PT XYZ melakukan laporan bulanan untuk evaluasi proses produksi semen terutama pada indeks batubara.

1.3 Tujuan Penelitian

Beberapa tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Menganalisis karakteristik indeks batubara di Pabrik Semen PT XYZ pada Januari 2019 - Juli 2019.
2. Memodelkan indeks batubara dengan tiga kelompok prediktor, yaitu faktor operasional, bahan baku, dan kualitas batubara.

1.4 Manfaat Penelitian

Terdapat dua manfaat pada penelitian ini, yaitu manfaat bagi PT XYZ dan manfaat bagi peneliti. Manfaat yang ingin diberikan bagi PT XYZ melalui penelitian ini adalah hasil prediksi dapat menjadi tambahan referensi bagi PT XYZ untuk meningkatkan dan mengevaluasi efisiensi dari proses produksi dengan melihat prediksi indeks batubara pada satu bulan kedepan.

Adapun manfaat yang diperoleh bagi peneliti adalah dapat mengaplikasikan pemodelan *Ridge Regression* dan *Support Vector Regression* (SVR). Selain itu, dapat memberikan penyelesaian secara statistik terutama dalam hal pemodelan prediksi menggunakan *Ridge Regression* dan *Support Vector Regression* (SVR).

1.5 Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini yaitu data yang digunakan dalam penelitian ini adalah data indeks batubara yang diperoleh dari PT XYZ pada tanggal Januari 2019 hingga Juli 2019, dengan variabel dependen adalah indeks batubara serta 17 variabel prediktor yang terbagi menjadi variabel operasional, bahan baku, dan kualitas batubara. Metode konvensional yang digunakan dalam prediksi indeks batubara ini adalah regresi *Ridge* untuk menanggulangi indikasi multikolinieritas. *Machine Learning* yang digunakan adalah *Support Vector Regression* (SVR) dengan penggunaan tiga kernel yaitu Linier, Polynomial dan RBF. *Feature selection* yang digunakan adalah *Recursive Feature Elimination* (RFE) dengan berbasis *Random Forest* (RF).

(Halaman ini sengaja dikosongkan)

BAB II TINJAUAN PUSTAKA

2.1 Analisis Regresi Linier dan Ridge

Regresi berganda adalah model regresi atau prediksi yang melibatkan lebih dari satu variabel bebas atau prediktor. Analisis ini diawali dengan *pre-processing* data yang bertujuan untuk mengidentifikasi adanya *missing value* dan data *outlier* yang akan menyebabkan asumsi residual berdistribusi normal tidak terpenuhi. Sehingga, *pre-processing* data merupakan langkah penting sebelum masuk ke tahap pengolahan dan analisis data, karena keputusan-keputusan yang berkualitas harus didasarkan pada data yang berkualitas (Kumar & Chadha, 2012). Regresi linier adalah sebuah pendekatan untuk memodelkan hubungan antara variabel terikat y dan satu atau lebih variabel bebas yang disebut x . Istilah regresi berganda dapat disebut juga dengan istilah *multiple regression*. Kata *multiple* berarti jamak atau lebih dari satu variabel. Model dari regresi linier berganda adalah:

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, i=1,2,\dots,n$$

dengan

y_i = nilai variabel dependen pada observasi ke- i . (2.1)

x_i = nilai variabel independen pada observasi ke- i .

p = banyaknya variabel independen x yang berpengaruh variabel dependen y .

ε_i = komponen galat yang diasumsikan berdistribusi normal dengan *mean* 0 dan memiliki *variansi* σ^2 .

$\alpha, \beta_1, \dots, \beta_p$ = koefisien regresi.

Nilai koefisien regresi α merupakan *intercept* yang diartikan sebagai nilai rata-rata y jika x_1, x_2, \dots, x_p masing-masing sama dengan nol. Nilai koefisien β_1 merupakan *slope* pada variabel y terhadap x_1 dan menganggap x_2, \dots, x_p adalah konstan. Nilai koefisien β_2 merupakan *slope* pada variabel y terhadap x_2 dan menganggap x_1, \dots, x_p dan x_1 adalah konstan, dan seterusnya.

2.1.1 Asumsi Regresi

Analisis varians (ANOVA) terdapat sejumlah asumsi yang harus dipenuhi agar pengujian ANOVA tersebut dapat dilakukan, yaitu asumsi residual yang bersifat IIDN (Identik, Independen, Distribusi Normal). Bila asumsi tersebut tidak terpenuhi, maka kesimpulan dari ANOVA tidak bisa digunakan (Montgomery, 2012). Selain asumsi IIDN (Identik, Independen, Distribusi Normal) pada regresi linier berganda perlu dilakukan adanya pemeriksaan asumsi multikolinieritas.

Analisis regresi linier mengasumsikan bahwa residual e_i berdistribusi normal. Pada regresi linier diasumsikan bahwa setiap didistribusikan secara random dengan $e_i \sim N(0, \sigma^2)$ (Gujarati, 2003). Kedua, yaitu asumsi variansi residu (e_i) pada setiap variabel adalah homoskedastisitas (Gujarati, 2003). Tujuan dari pengujian ini adalah untuk mengetahui apakah variansi pada tiap residu e_i konstan. Jika variansi pada tiap residu e_i berbeda disebut heteroskedastisitas. Ketiga, yaitu asumsi tidak adanya autokorelasi antara serangkaian pengamatan yang diurutkan menurut waktu. Pendeteksian autokorelasi dapat dideteksi secara grafis yaitu dengan melihat *scatterplot* residu terhadap urutan waktu. Jika sebaran residu terhadap urutan waktu tidak membentuk pola tertentu atau bersifat acak maka dapat disimpulkan tidak ada autokorelasi antar residu. Kemudian, pemeriksaan asumsi multikolinieritas yang merupakan terjadinya hubungan linier antara variabel bebas dalam satu model regresi linier berganda. Adapun dampak terjadinya multikolinieritas dalam model regresi linier berganda adalah penaksir OLS masih bersifat BLUE, tetapi mempunyai variansi dan kovariansi yang besar sehingga sulit mendapatkan taksiran (estimasi) yang tepat. Akibat penaksir OLS mempunyai variansi dan kovariansi yang yang besar, menyebabkan interval estimasi akan cenderung lebih lebar dan nilai hitung statistik uji t akan kecil, sehingga membuat variabel bebas secara statistik tidak signifikan mempengaruhi variabel tidak bebas. Walaupun secara individu variabel bebas tidak berpengaruh terhadap variabel tidak bebas melalui uji t, tetapi nilai koefisien determinasi (R^2) masih bisa relatif tinggi (Gujarati, 2003). Untuk

megatasi adanya kasus multikolinieritas, perlu dilakukan pemodelan menggunakan regresi *ridge* untuk menstabilkan nilai koefisien regresi akibat adanya multikolinieritas.

2.1.2 Regresi *Ridge*

Suatu teknik yang dikembangkan untuk menstabilkan nilai koefisien regresi karena adanya multikolinieritas. Metode *Ridge* pertama kali dikemukakan oleh A.E Hoerl pada tahun 1962. Metode ini bertujuan untuk mengatasi kondisi yang tidak diinginkan yang disebabkan oleh korelasi atau hubungan yang tinggi antara beberapa variabel singular atau tunggal sehingga menghasilkan nilai dugaan parameter model regresi yang tidak stabil (Draper & Smith, 1981).

Persamaan regresi *ridge* dinyatakan dalam bentuk sebagai berikut:

$$\hat{\beta}_R = W\mathbf{X}^t\mathbf{Y} \quad (2.2)$$

$$W = (\mathbf{X}^t\mathbf{X} + c\mathbf{I})^{-1} \quad (2.3)$$

$$\begin{aligned} \hat{\beta}_R &= (\mathbf{X}^t\mathbf{X} + c\mathbf{I})^{-1} \mathbf{X}^t\mathbf{Y} \quad (2.4) \\ &= (\mathbf{X}^t\mathbf{X} + c\mathbf{I})^{-1} (\mathbf{X}^t\mathbf{X})(\mathbf{X}^t\mathbf{X})^{-1} \mathbf{X}^t\mathbf{X} \\ &= (\mathbf{X}^t\mathbf{X} + c\mathbf{I})^{-1} (\mathbf{X}^t\mathbf{X}) \hat{\beta} \end{aligned}$$

Salah satu metode untuk mendeteksi multikolinieritas dengan menggunakan *ridge trace* (*jarak ridge*). Salah satu kesulitan dalam menggunakan *ridge trace* adalah membentuk nilai c yang tepat (Hoerl & Kennard, 1970). Suatu acuan yang sering digunakan untuk melihat besarnya c adalah dengan melihat VIF dan melihat kecenderungan plot estimator *ridge trace*. Bila kenaikan VIF yang mendekati nilai satu menunjukkan bahwa variabel bebas tidak saling berkorelasi dengan variabel bebas lainnya. Setelah mendapatkan nilai c yang optimum yaitu ketika VIF seluruh variabel tidak sama dengan nol, selanjutnya dilakukan transformasi koefisien parameter dari setiap variabel. Transformasi koefisien pa-

parameter dari setiap variabel didapatkan dengan rumus sebagai berikut :

$$\tilde{\beta}_k = \left(\frac{S_y}{S_k} \right) \tilde{\beta}_k^* \quad (2.5)$$

$\tilde{\beta}_k$ = Koefisien Parameter

$\tilde{\beta}_k^*$ = Koefisien Parameter Sebelum di Transformasi

S_y = Simpangan Baku Indeks Batubara

S_k = Simpangan Baku Variabel Prediktor ke- k

$k = 1, 2, 3, \dots, 17$

2.1.4 Pengujian Hipotesis

Untuk mengetahui adanya hubungan linier antara variabel bebas dengan variabel terikat atau tidak, dilakukan pengujian signifikansi parameter menggunakan hipotesis :

$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$ (tidak ada hubungan linier antara variabel bebas dengan variabel terikat)

$H_1 : \exists_j, \beta_j \neq 0, j = 1, 2, \dots, k$ (ada hubungan linier antara variabel bebas dengan variabel terikat)

dengan tingkat signifikansi (α) tertentu pada uji statistik

dengan $F_{hitung} = \frac{JKR / k}{JKR / (n - k - 1)}$, (2.6)

maka keputusan hipotesis dapat dilakukan penolakan H_0 jika $F_{hitung} > F_{tabel}$ atau gagal tolak H_0 jika $F_{hitung} \leq F_{tabel}$

Untuk mengetahui koefisien yang diperoleh berarti atau tidak, dilakukan pengujian signifikansi parameter menggunakan hipotesis :

$H_0 : \beta_i = 0$ (koefisien regresi tidak signifikan)

$H_1 : \exists_i, \beta_i \neq 0, i = 1, 2, \dots, k$ (koefisien regresi signifikan)

dengan tingkat signifikansi (α) tertentu pada uji statistik

$$t_{hitung} = \frac{b_i}{sb_i}, \quad (2.7)$$

dimana

b_i = Koefisien regresi variabel $-i$,

sb_i = Standar *error* variabel $-i$,

maka keputusan hipotesis dapat dilakukan penolakan H_0 jika

$$|t_{hitung}| > t_{\left(\frac{\alpha}{2}, n-k-1\right)},$$

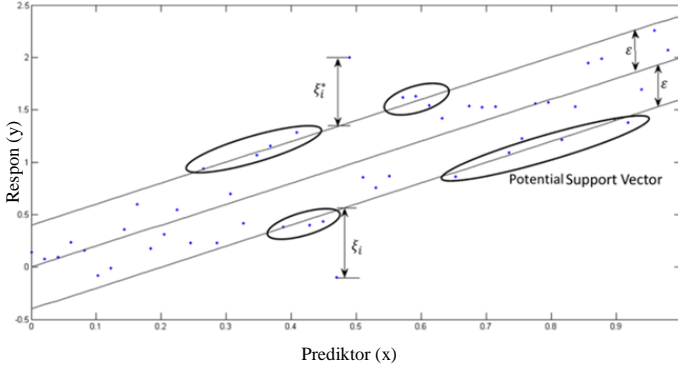
atau gagal tolak H_0 jika

$$|t_{hitung}| \leq t_{\left(\frac{\alpha}{2}, n-k-1\right)}.$$

2.2 Support Vector Regression (SVR)

Support Vector Regression (SVR) merupakan pengembangan *Support Vector Machine* (SVM) untuk kasus regresi. Tujuan dari *Support Vector Regression* (SVR) adalah untuk menemukan sebuah fungsi $f(x)$ sebagai suatu *hyperplane* (garis pemisah) berupa fungsi regresi yang mana sesuai dengan semua input data dengan sebuah *error* ε dan membuat ε setipis mungkin (Scholkopf & Smola, 2002). Sebelum melakukan pemodelan dengan SVR perlu dilakukan *pre-processing* data terlebih dahulu untuk mengidentifikasi adanya ketidaklengkapan data yang terjadi karena adanya atribut data yang tidak tersedia atau hilangnya nilai untuk beberapa data (atribut), hal itu bisa disebabkan karena pada saat pengumpulan data terdapat instrumen yang rusak karena kesalahan manusia (*human error*) ataupun kesalahan komputer (Han, Kamber, & Pei, 2012).

Perumusan masalah SVR seringkali berasal dari perspektif geometris, yaitu menggunakan Linier SVR satu dimensi pada Gambar 2.1.



Gambar 2.1 Ilustrasi *Support Vector Regression*

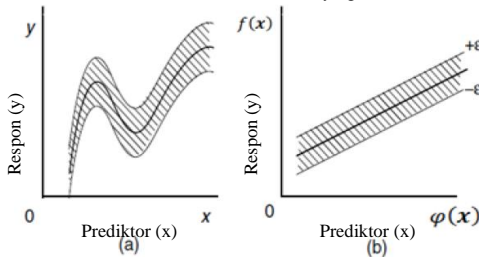
Ilustrasi Gambar 2.1 menjelaskan bahwa ϵ merupakan margin (jarak terjauh antara *hyperplane* dengan *support vector*), titik observasi diluar *hyperplane* akan dikenai pinalti berupa variabel *slack* ξ dan ξ^* . Untuk data multidimensi, perlu menambah x satu per satu dan memasukkan b ke dalam vektor w untuk notasi secara matematis.

$$y = f(x) = \langle w, x \rangle + b = \sum_{j=1}^M w_j x_j + b, y, b \in \mathbb{R}, x, w \in \mathbb{R}^M \quad (2.8)$$

$$f(x) = \begin{bmatrix} w \\ b \end{bmatrix}^T \begin{bmatrix} x \\ 1 \end{bmatrix} = \mathbf{w}^T \mathbf{x} + b \quad x, w \in \mathbb{R}^{M+1} \quad (2.9)$$

Pada regresi terdapat residual misalkan residual (r) didefinisikan dengan mengurangi *output* skalar y terhadap estimasi $f(x)$ yaitu $r = y - f(x)$ dengan :

$$E(r) = \begin{cases} 0 & \text{untuk } |r| \leq \epsilon \\ |r| - \epsilon & \text{untuk yang lain} \end{cases} \quad (2.10)$$



Gambar 2.2 *Insensitive Zone* (a), *Original Input Space*, dan (b) *Feature Space*

$D(x,y)=\pm\epsilon$ adalah jarak terjauh *support vector* dari *hyperplane* yang disebut margin. Memaksimalkan margin akan meningkatkan probabilitas data ke dalam radius $\pm\epsilon$. Jarak dari *hyperplane* $D(x,y)=0$ ke data (x,y) adalah $|D(x,y)|/\|\mathbf{w}\|$, dimana :

$$\mathbf{w} = (1 - \mathbf{w}^T)^T \quad (2.11)$$

Diasumsikan bahwa jarak maksimum data terhadap *hyperplane* adalah δ , maka estimasi yang ideal akan terpenuhi dengan :

$$\begin{aligned} \frac{|D(x,y)|}{\|\mathbf{w}\|} &\leq \delta \\ |D(x,y)| &\leq \delta \|\mathbf{w}\| \\ \delta \|\mathbf{w}\| &= \epsilon \end{aligned} \quad (2.12)$$

Oleh karena itu untuk memaksimalkan *margin* δ , diperlukan $\|\mathbf{w}\|$ yang minimum. Optimasi penyelesaian masalah dengan bentuk *Quadratic Programming* :

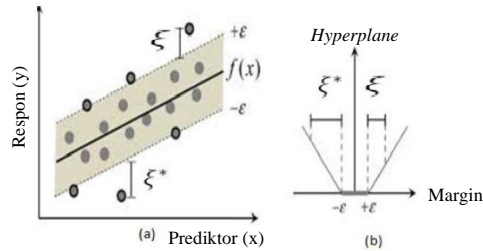
$$\min \frac{1}{2} \|\mathbf{w}\|^2 \quad (2.13)$$

dengan syarat :

$$y_i - \mathbf{w}^T \varphi(x_i) - b \leq \epsilon \text{ untuk } i = 1, 2, \dots, n$$

$$\mathbf{w}^T \varphi(x_i) - y_i + b \leq \epsilon \text{ untuk } i = 1, 2, \dots, n$$

Faktor $\|\mathbf{w}\|^2$ dinamakan regulasi. Meminimalkan $\|\mathbf{w}\|^2$ akan membuat suatu fungsi setipis (*flat*) mungkin, sehingga bisa mengontrol kapasitas fungsi (*function capacity*). Diasumsikan bahwa semua titik ada dalam rentang $f(x)\pm\epsilon$ (*feasible*), dalam hal ketidaklayakan (*infesibility*) ada beberapa titik yang mungkin keluar dari rentang $f(x)\pm\epsilon$ maka ditambahkan variabel *slack* ξ dan ξ^* untuk mengatasi masalah pembatasan yang tidak layak (*infeasible constraints*) dalam problem optimasi (Santosa, 2007). Posisi titik terhadap *hyperplane* ditampilkan pada gambar berikut :



Gambar 2.3 (a) SVR Output, dan (b) Insensitive Loss Function

Gambar 2.3 menampilkan bahwa terdapat titik diluar margin, titik-titik ini akan dikenai pinalti. Selanjutnya problem optimisasi di atas bisa diformulasikan menggunakan *loss function* untuk pinalti titik observasi yang berada diluar garis *support vector* dengan rumus sebagai berikut;

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n (\xi + \xi_i^*) \quad (2.14)$$

dengan syarat:

$$y_i - \mathbf{w}^T \varphi(x_i) - b - \xi_i \leq \varepsilon, \quad i=1,2,\dots,n$$

$$\mathbf{w}^T \varphi(x_i) - y_i + b - \xi_i^* \leq \varepsilon, \quad i=1,2,\dots,n$$

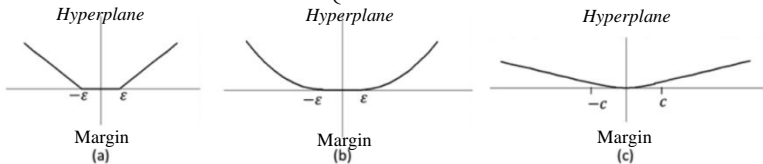
$$\xi_i, \xi_i^* \geq 0$$

Loss Function adalah fungsi yang menunjukkan hubungan antara *error* dengan bagaimana *error* ini dikenai pinalti. Perbedaan *loss function* akan menghasilkan formula SVR yang berbeda (Santosa, 2007). Beberapa *loss function* yang dapat digunakan adalah linier, kuadratik, dan Huber, seperti yang ditunjukkan pada persamaan (2.15), (2.16), dan (2.17). Seperti ditunjukkan pada Gambar 2.4, *lost function* Huber lebih halus daripada fungsi linier dan kuadratik. Pilihan *loss function* dipengaruhi oleh informasi apriori tentang distribusi kebisingan mempengaruhi sampel data (Huber, 1964).

$$L_{\varepsilon}(y, f(x, \mathbf{w})) = \begin{cases} 0 & |y-f(x, \mathbf{w})| \leq \varepsilon; \\ |y-f(x, \mathbf{w})| - \varepsilon & \text{lainnya,} \end{cases} \quad (2.15)$$

$$L_{\varepsilon}(y, f(x, \mathbf{w})) = \begin{cases} 0 & |y-f(x, \mathbf{w})| \leq \varepsilon; \\ (y-f(x, \mathbf{w}) - \varepsilon)^2 & \text{lainnya,} \end{cases} \quad (2.16)$$

$$L(y, f(x, \mathbf{w})) = \begin{cases} c|y-f(x, \mathbf{w})| - \frac{c^2}{2} & |y-f(x, \mathbf{w})| > c; \\ |y-f(x, \mathbf{w})| - \varepsilon & |y-f(x, \mathbf{w})| \leq c \end{cases} \quad (2.17)$$



Gambar 2.4 Tipe *Lost Function*: (a) Linier, (b) Kuadratik, dan (c) Huber

Dengan konstanta $c > 0$ menentukan (*trade off*) antara ketipisan fungsi (*flatness of function*) $f(x)$ dan batas atas deviasi yang lebih besar dari ε masih ditoleransi. Semua deviasi yang lebih besar dari ε akan dikenai pinalti sebesar c .

Pendekatan *soft margin* mirip dengan yang digunakan dalam SVM, variabel *slack* x, x^* dapat ditambahkan untuk menjaga terhadap adanya *outlier*. Parameter yang dapat disempurnakan yaitu yang memberikan bobot lebih untuk meminimalkan kerataan atau kesalahan untuk masalah optimisasi multi-tujuan ini. Sebagai contoh, c yang lebih besar memberikan bobot lebih untuk meminimalkan kesalahan.

Masalah optimisasi kuadratik terbatas ini dapat diselesaikan dengan menemukan *Lagrangian* pada persamaan (2.18). Penganda *Lagrange* atau variabel ganda adalah $\lambda, \lambda^*, \alpha, \alpha^*$ dan bilangan *real* non-negatif.

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi + \xi_i^* \quad (2.18)$$

dengan

$$\begin{aligned}
y_i - \mathbf{w}^T x_i &\leq \varepsilon + \xi_i^*, \quad i=1,2,\dots,n \\
\mathbf{w}^T x_i - y_i &\leq \varepsilon + \xi_i, \quad i=1,2,\dots,n \\
\xi_i, \xi_i^* &\geq 0 \quad i=1,2,\dots,n
\end{aligned}$$

Solusi optimasi dapat diselesaikan dengan fungsi *Lagrange* berikut :

$$\begin{aligned}
L(\mathbf{w}, \xi^*, \xi, \lambda, \lambda^*, \alpha, \alpha^*) &= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i + \xi_i^* - \sum_{i=1}^N \alpha_i^* (y_i - \mathbf{w}^T x_i - \varepsilon - \xi_i^*) \\
&\quad + \sum_{i=1}^N \alpha_i (-y_i + \mathbf{w}^T x_i - \varepsilon - \xi_i) - \sum_{i=1}^N \lambda_i \xi_i + \lambda_i^* \xi_i^*
\end{aligned} \quad (2.19)$$

Notasi $L(\mathbf{w}, \xi^*, \xi, \lambda, \lambda^*, \alpha, \alpha^*)$ dinamakan *Lagrangian*, dimana : $\lambda_i, \lambda_i^*, \alpha_i, \alpha_i^*$ adalah *Lagrange Multiplier*. Untuk mendapatkan solusi yang optimal, maka dilakukan turunan parsial dari L terhadap $w, \xi, \xi^*, \lambda, \lambda^*, \alpha, \alpha^*$. Persamaan (2.19) didapatkan dengan mengambil turunan parsial yang disama dengankan nol, berdasarkan kondisi Karush Kuhn Tucker (KKT). Derivatif parsial pada pengganda *Lagrange* mengembalikan kendala yang harus kurang dari atau sama dengan nol, seperti yang diilustrasikan dalam persamaan (2.20). Kondisi akhir KKT menyatakan bahwa pengganda *Lagrange* dan batasannya sama dengan nol pada persamaan (2.10). Pengganda *Lagrange* yang sama dengan nol sesuai dengan data di dalam tabung, sedangkan vektor dukungan memiliki pengganda *Lagrange* bernilai nol. Persamaan fungsi diwakili dalam persamaan (2.23) dengan mengganti persamaan (2.20) dalam persamaan (2.19), sehingga bentuk ganda dari masalah optimisasi dapat ditulis seperti yang ditunjukkan pada persamaan (2.25).

$$\frac{\delta L}{\delta \mathbf{w}} = \mathbf{w} - \sum_{i=1}^N (\alpha_i^* - \alpha_i) x_i = 0$$

$$\frac{\delta L}{\delta \xi_i^*} = C - \lambda_i^* - \alpha_i^* = 0$$

$$\frac{\delta L}{\delta \xi_i} = C - \lambda_i - \alpha_i = 0$$

$$\frac{\delta L}{\delta \lambda_i^*} = \sum_{i=1}^N \xi_i^* \leq 0 \quad (2.20)$$

$$\frac{\delta L}{\delta \lambda_i} = \sum_{i=1}^N \xi_i \leq 0$$

$$\frac{\delta L}{\delta \alpha_i^*} = y_i - \mathbf{w}^T x_i - \varepsilon - \xi_i^* \leq 0$$

$$\frac{\delta L}{\delta \alpha_i} = y_i + \mathbf{w}^T x_i - \varepsilon - \xi_i \leq 0$$

$$\alpha_i (-y_i + \mathbf{w}^T x_i - \varepsilon - \xi_i) = 0$$

$$\alpha_i^* (-y_i - \mathbf{w}^T x_i - \varepsilon - \xi_i^*) = 0 \quad \forall i \quad (2.21)$$

$$\lambda_i \xi_i = 0,$$

$$\lambda_i^* \xi_i^* = 0$$

$$\mathbf{w} = \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) x_i \quad (2.22)$$

$$f(x) = \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) x_i^T x, \alpha_i, \alpha_i^* \in [0, C] \quad (2.23)$$

$$\max_{\alpha, \alpha^*} -\varepsilon \sum_{i=1}^{N_{SV}} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) y_i - \frac{1}{2} \sum_{j=1}^{N_{SV}} \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) x_i^T x_j, \quad (2.24)$$

dengan

$$\sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) = 0, \alpha_i, \alpha_i^* \in [0, C]$$

Pada bagian ini, vektor bobot \mathbf{w} ditambah dengan skalar b dan derivasi dari formulasi matematis SVR dilakukan dengan mengabaikan perhitungan eksplisit dari b pada persamaan (2.9). Data *training* yang berada di luar batas tabung akan memiliki nilai bukan nol α atau α^* yang keduanya tidak boleh nol, maka *instance* (x_i, y_i) tidak mungkin milik batas bawah. Selanjutnya, karena titik tersebut tidak berada di luar tabung ξ_i , sehingga akan mengarah ke hasil dalam persamaan (2.26) ketika $\alpha \in (0, C)$. Persamaan (2.27)

menghitung b dengan melakukan analisis yang sama untuk α_i^* , sehingga didapatkan persamaan (2.28) dan (2.29).

$$y_i - \mathbf{w}^T x_i - b - \varepsilon - \xi_i = 0 \quad (2.25)$$

$$y_i - \mathbf{w}^T x_i - b - \varepsilon = 0 \quad (2.26)$$

$$b = y_i - \mathbf{w}^T x_i - \varepsilon \quad (2.27)$$

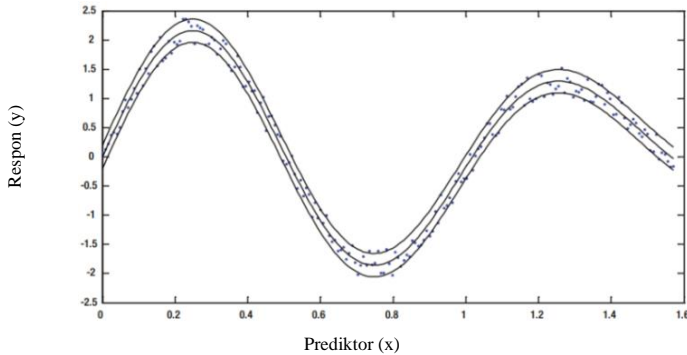
$$-y_i + \mathbf{w}^T x_i - b - \varepsilon = 0 \quad (2.28)$$

$$b = -y_i + \mathbf{w}^T x_i - \varepsilon \quad (2.29)$$

Selain menggunakan kondisi KKT, nilai b juga bisa dihitung dan menyelesaikan masalah optimisasi dengan menggunakan metode titik interior, yang dapat menyatu ke solusi optimal dalam waktu logaritmik dengan menavigasi di sepanjang jalur pusat wilayah. Jalur sentral ditentukan dengan memecahkan masalah optimisasi primal dan dual secara bersamaan.

2.2.1 Fungsi Kernel

Menurut Santosa (2007) banyak teknik *data mining* atau *machine learning* yang dikembangkan dengan asumsi kelinieran. Untuk fungsi non linier, data dapat dipetakan ke ruang dimensi yang lebih tinggi melalui φ , yang disebut ruang kernel untuk mencapai akurasi yang lebih tinggi dengan menggunakan kernel yang memenuhi kondisi Mercer pada Gambar 2.5.



Gambar 2.5 Regresi Non Linier

Oleh karena itu, perlu mengganti semua *instance* x dalam persamaan (2.7) sampai persamaan (2.28) dengan $k(x_i, x_j)$ menghasilkan formulasi primal yang ditunjukkan dalam persamaan (2.29), di mana $\varphi(\cdot)$ adalah transformasi dari fitur ke ruang kernel. Persamaan (2.30) menggambarkan vektor bobot baru dalam hal input yang ditransformasikan. Perkiraan fungsi $f(x)$ adalah dalam persamaan (2.32), di mana $k(\cdot, \cdot)$ adalah kernel yang diilustrasikan dalam persamaan (2.33).

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi + \xi_i^*, \quad (2.30)$$

dengan :

$$\begin{aligned} y_i - \mathbf{w}^T \varphi(x_i) &\leq \varepsilon + \xi_i^*, \quad i = 1, \dots, N \\ \mathbf{w}^T \varphi(x_i) - y_i &\leq \varepsilon + \xi_i, \quad i = 1, \dots, N \\ \xi_i, \xi_i^* &\geq 0 \quad i = 1, \dots, N \end{aligned}$$

$$\mathbf{w} = \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) \varphi(x_i) \quad (2.31)$$

$$\max_{\alpha, \alpha^*} -\varepsilon \sum_{i=1}^{N_{SV}} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) y_i - \frac{1}{2} \sum_{j=1}^{N_{SV}} \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) k(x_i - x_j) \quad (2.32)$$

$$\alpha_i, \alpha_i^* \in [0, C], i = 1, \dots, N_{SV}, \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) = 0$$

$$f(x) = \sum_{i=1}^{N_{SV}} (\alpha_i^* - \alpha_i) k(x_i, x) \quad (2.33)$$

$$k(x_i, x) = \varphi(x_i) \cdot \varphi(x) \quad (2.34)$$

Nilai $K(x_i, x_j)$ merupakan fungsi kernel yang menunjukkan pemetaan linier pada *feature space*. Nilai $K(x_i, x_j)$ tidak selalu bisa diekspresikan secara eksplisit sebagai kombinasi antara α , y dan $\varphi(x)$, karena dalam banyak kasus $\varphi(x)$ tidak diketahui atau sulit dihitung. Fungsi kernel yang digunakan dalam penelitian ini adalah:

1. Kernel Linier

$$\varphi(x) = K(x, x') = x^T x \quad (2.35)$$

2. Kernel Polynomial

$$\varphi(x) = K(x, x') = (\gamma(x^T x) + 1)^d \quad (2.36)$$

3. Radial Basis Function (RBF)

$$\varphi(x) = K(x, x') = \exp(-\gamma \|x - x_i\|^2) \quad (2.37)$$

2.2.2 Recursive Feature Elimination (RFE)

Feature selection adalah upaya untuk memilih fitur subset dari fitur asli yang paling berguna. *Feature selection* adalah upaya untuk memetakan semua fitur ke dalam fitur baru yang lebih sedikit. Kelebihan *feature selection* dibandingkan *feature extraction* adalah akuisisi data yang lebih cepat. Oleh karena itu, pengurangan fitur pada data akan lebih baik menggunakan *feature selection* (Casasent & Nakariyakul, 2004). Salah satu dari banyak *feature selection* yang tersedia adalah *Recursive Feature Elimination* (RFE).

Salah satu metode yang dapat digunakan dengan RFE ini adalah *Random Forest*. Kelebihan *Random Forest* antara lain dapat menghasilkan *error* yang lebih rendah, memberikan hasil yang bagus dalam klasifikasi, dapat mengatasi data *training* dalam jumlah yang sangat besar secara efisien dan metode yang efektif untuk mengestimasi *missing value* (Breiman, 2001).

Random Forest juga merupakan sebuah teknik *machine learning* yang dapat digunakan dengan dimensi data yang tinggi dan memperbolehkan hubungan non-linier yang berada di variabel prediktor, khususnya pada beberapa variabel prediktor yang memiliki hubungan sangat kuat. Tidak semua metode dapat digabungkan dengan *Recursive Feature Elimination* (RFE), terdapat beberapa metode yang memiliki keuntungan atau kerugian lebih, karena RFE membutuhkan model awal menggunakan set prediktor penuh, maka beberapa model tidak dapat digunakan ketika jumlah prediktor melebihi jumlah sampel. *Random Forest* merupakan salah satu model yang dapat digabungkan dengan RFE (RF-RFE) (Svetnik, 2003). Penggunaan *Random Forest* digunakan dengan RFE (RF-RFE) karena model ini memiliki metode internal yang mampu mengukur kepentingan fitur, di-

mana seluruh rangkaian prediktor digunakan untuk menghitung peringkat fitur.

2.3 Kebaikan Model

RMSE merupakan alat seleksi model berdasarkan pada *error* hasil estimasi. *Error* yang ada menunjukkan seberapa besar perbedaan hasil estimasi dengan nilai yang akan diestimasi. Nilai ini akan digunakan untuk menentukan model mana yang terbaik. Definisi RMSE dapat ditulis sebagai berikut. *Mean square error* digunakan untuk menghitung tingkat *error* dari dua buah hasil percobaan model, atau jika diimplementasikan dalam *data mining* maka untuk mengukur tingkat *error* dari hasil penghitungan sebuah analisis yang menggunakan metode tertentu antara data *training* dan data *testing*. Pengukuran RMSE dilakukan dengan rumus seperti berikut ini:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - Y_i)^2}{n}} \quad (2.38)$$

RMSE : *Root Mean Square Error*

n : Jumlah Sampel

y_i : Nilai Aktual

Y_i : Nilai Prediksi Sesuai Model

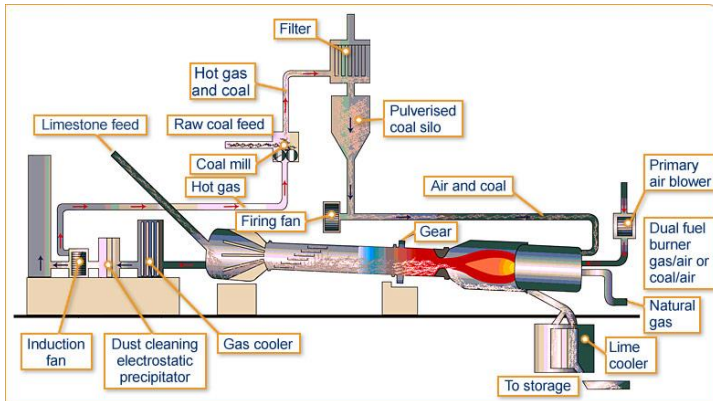
2.4 Proses Produksi Semen

PT XYZ merupakan produsen semen terbesar dan tercatat sebagai BUMN Multinasional pertama di Indonesia yang memiliki anak usaha di luar negeri yang hingga saat ini konsisten merealisasikan program pengembangan usaha memantapkan posisi di pasar, mendukung pertumbuhan ekonomi dalam membangun bangsa serta meningkatkan kesejahteraan para pemangku kepentingan.

Semen terbuat dari pengolahan empat *raw* material. Bahan baku semen sendiri terdiri atas batu kapur sebanyak 80%, tanah liat 15%, pasir silika 1-4%, dan pasir besi 1-4%. Proses pembuatan semen sendiri padadasarnya melalui lima tahapan, yaitu penye-

diaan bahan mentah, penggilingan bahan mentah (batu kapur dan tanah liat), pembakaran *clinker* (campuran dari batu kapur dan tanah liat yang telah digiling), penggilingan akhir, lalu yang terakhir adalah pengantongan sak semen/pengemasan. Unit pembakaran merupakan bagian terpenting karena terjadi pembentukan komponen utama semen. Unit ini terdapat *suspenser preheater*, *kiln* dan *great cooler*. Proses pembakaran untuk menghasilkan *clinker* diawali dengan menyiapkan bahan bakarnya terlebih dahulu baru kemudian melakukan pembakaran. Di dalam proses pembakaran pada *kiln* menggunakan bahan bakar utama batubara dengan kapasitas 15,40 ton per-jam untuk membakar material yang ada didalam *kiln*. Batubara digunakan sebagai bahan bakar dalam proses pembakaran karena biaya produksinya lebih murah sehingga dapat mengurangi *cost* dibandingkan menggunakan bahan bakar lain. Batubara yang telah disiapkan akan otomatis masuk ke dalam *kiln* sesuai dengan kebutuhan proses pembakaran.

Tujuan dari proses pembakaran ini ialah untuk menghasilkan *clinker* bermutu baik dengan pemakaian energi serendah mungkin dan operasi pembakaran berlangsung stabil dalam waktu yang lama. Salah satu faktor utama untuk mendapatkan hasil pembakaran yang baik ialah rancangan *kiln feed (raw mix design)* yaitu menentukan komposisi kimia dan ukuran partikel atau kehalusan dari *raw mix*. *Raw mix* dirancang untuk menghasilkan *clinker* bermutu baik (mempunyai senyawa *alite* C3S, *belite* C2S, *aluminat* C3A, *ferrite* C4AF dalam jumlah cukup dan mudah digiling) yang dapat diukur dengan perhitungan modulus LSF, SIM, dan ALM. Selain itu, batubara juga memiliki kualitas senyawa kimia yang terkandung didalamnya seperti AC (*Ash Content*), VM (*Volatile Matter*), FC (*Fixed Carbon*), TS (*Total Sulfur*), TM (*Total Moisture*), dan GHV (*Gross Heating Value*). Proses pada tahap ini meliputi pemanasan awal umpan baku di *preheater* (pengeringan, dehidrasi dan dekomposisi), pembakaran di *kiln* (klinkerisasi) dan pendinginan di *grate cooler (quenching)*. Selanjutnya *clinker* yang dihasilkan disimpan di *clinker silo*.



Gambar 2.6 Proses Pembaraan *Clinker* di *Kiln*

Pada saat material telah masuk ke *kiln*, terdapat empat zona proses pemanasan diantaranya:

1. *Calsinasi zone*, dimana pada proses ini material yang baru masuk kedalam *kiln*, material tersebut terkalsinasi dikarenakan mendapatkan panas yang lebih tinggi dari pada di dalam SP berkisar antara 1100-1200°C sehingga mengakibatkan perubahan bentuk pada material tersebut yang tadinya berupa serbuk-serbuk padat menjadi serbuk-serbuk yang mulai terlihat meleleh.
2. *Transisi zone*, dimana pada proses ini bahan material mendapatkan pemanasan yang lebih tinggi berkisar antara 1200-1300°C dimana pada proses ini material hampir mendekati cair.
3. *Burning zone*, dimana pada proses ini material benar-benar mendapatkan pemanasan secara penuh dari *kiln* hingga material tersebut mencair dan panasnya mencapai 1400–1600°C.
4. *Cooling zone*, pada proses ini material yang telah masuk ke *cooler* mendapatkan pendinginan secara cepat atau proses pendinginan yang dikagetkan karena pada *cooler* ini panas pada material harus lebih dingin dibandingkan didalam kiln dimaksudkan supaya *clinker* tersebut tidak lengket pada great plat dan panas pada *cooler* mencapai 150-200°C. Panas yang dihasilkan didalam tungku *kiln* tidak serta merta berimbas keluar dikarenakan pada dinding *kiln* dilapisi oleh bata tahan api yang mampu menahan panas yang sangat tinggi hingga

1600°C sehingga lingkungan yang disekitar *kiln* tidak terlalu panas pada saat kita berada disekitar area *kiln*.

2.4.1 Indeks Batubara

Indeks sejatinya adalah indikator ataupun ukuran atas sesuatu. Indeks batubara adalah suatu ukuran untuk mengukur kinerja proses produksi (*clinker*) yang dihasilkan dari tahap pembakaran. *Clinker* adalah hasil gilingan dari batu kapur dan tanah liat yang merupakan bahan baku utama dari semen dengan proses pengolahan menggunakan mesin (*kiln*) berbahan bakar utama batubara. Semakin banyak produksi *clinker* yang dihasilkan dan semakin sedikit batubara yang digunakan dalam proses pembakaran maka semakin efektif dan efisien proses produksi tersebut. Indeks batubara didapatkan melalui rumus sebagai berikut:

$$\text{Indeks batubara} = \frac{\text{clinker}}{\text{raw coal}} \quad (2.39)$$

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Pada penelitian ini sumber data yang akan digunakan merupakan data sekunder yang diperoleh pada produksi *clinker* dan pemakaian batubara bulan Januari 2019 - Juli 2019. Data diperoleh dari PT XYZ. Data yang digunakan sebanyak 213 observasi.

3.2 Variabel Penelitian

Variabel penelitian indeks batubara dan variabel prediktor-nya ditampilkan pada Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Simbol	Variabel	Satuan	Skala Data
Y	Indeks Batubara	ton/ton	Rasio
X ₁	<i>Speed Kiln</i>	Rpm	Rasio
X ₂	<i>ILC Exit Temperature</i>	°C	Rasio
X ₃	<i>SLC Exit Temperature</i>	°C	Rasio
X ₄	<i>Temperature Stage 4 ILC</i>	°C	Rasio
X ₅	<i>Temperature Stage 4 SLC</i>	°C	Rasio
X ₆	<i>ILC Coal</i>	Ton	Rasio
X ₇	<i>SLC Coal</i>	Ton	Rasio
X ₈	<i>Kiln Coal</i>	Ton	Rasio
X ₉	<i>LSF (Lime Saturation Factor)</i>	-	Rasio
X ₁₀	<i>SIM (Silica Modulus)</i>	-	Rasio
X ₁₁	<i>ALM (Alumina Modulus)</i>	-	Rasio
X ₁₂	<i>AC (Ash Content)</i>	%	Rasio
X ₁₃	<i>VM (Volatile Matter)</i>	%	Rasio
X ₁₄	<i>FC (Fixed Carbon)</i>	%	Rasio
X ₁₅	<i>TS (Total Sulfur)</i>	%	Rasio
X ₁₆	<i>TM (Total Moisture)</i>	%	Rasio
X ₁₇	<i>GHV (Gross Heating Value)</i>	cal/g	Rasio

Pengukuran variabel-variabel didasarkan pada pemeriksaan yang dilakukan oleh bagian laboratorium *Quality Control* dan *Production Planning and Evaluation*. Di tahap pembakaran *clinker*

(batu kapur dan tanah liat) terdiri dari delapan belas variabel yang diukur dan dicatat yaitu sebagai berikut.

1. Indeks batubara adalah suatu ukuran untuk mengukur kinerja proses produksi (*clinker*) yang dihasilkan dari tahap pembakaran. Semakin banyak produksi *clinker* yang dihasilkan dan semakin sedikit batubara yang digunakan dalam proses pembakaran maka semakin efektif dan efisien proses produksi tersebut.
2. *Speed kiln* adalah kecepatan mesin (*kiln*) dalam proses pembakaran bahan baku semen (batu kapur dan tanah liat).
3. ILC (*In-Line Calciner*) dan SLC (*Separate-Line Calciner*) *Exit Temperature* adalah dua jenis *temperature* yang dikeluarkan pada pemanas awal dan kalsiner yang terdapat pada hampir seluruh pemanas awal modern yang mempunyai dampak secara langsung terhadap konsumsi bahan bakar, emisi polutan dan kualitas akhir semen yang dihasilkan. Sistem ILC ini adalah aliran gas panas dari *kiln* langsung menuju kalsiner, sedangkan SLC adalah udara pembakaran pada kalsiner berasal dari *separate tertiary air duct*.
4. *Temperature stage 4* ILC (*In-Line Calciner*) dan SLC (*Separate-Line Calciner*) adalah dua jenis *temperature* pemanas awal pada *stage 4* dan kalsiner yang terdapat pada hampir seluruh pemanas awal.
5. ILC (*In-Line Calciner*) dan SLC (*Separate-Line Calciner*) *Coal* adalah jumlah batubara yang berada pada dua jenis sistem (ILC dan SLC) pemanas awal dan kalsiner yang terdapat pada hampir seluruh pemanas awal.
6. *Kiln coal* adalah jumlah batubara yang berada dalam *kiln*.
7. *Lime Saturation Factor* (LSF), *Silica Modulus* (SIM), dan *Alumina Modulus* (ALM) adalah standar kualitas unit *raw mill*. LSF, SIM, dan ALM adalah faktor dan modulus untuk kontrol kualitas yang terdiri dari beberapa senyawa dan tidak memiliki satuan (rasio). Hal ini dikarenakan pada unit proses pembuatan semen terjadi pencampuran senyawa menjadi senyawa lainnya.

8. *Ash Content* (AC) adalah bahan-bahan yang tidak terbakar setelah pembakaran *sample*. Hasil kadar abu (*ash content*) digunakan untuk mengukur kualitas batubara dan efisiensi proses pembersihan.
9. *Volatile Matter* (VM) adalah banyaknya zat yang hilang bila sampel batubara dipanaskan pada suhu dan waktu yang telah ditentukan (setelah dikoreksi oleh kadar *moisture*).
10. *Fixed Carbon* (FC) adalah zat yang tidak menguap dan tersisa setelah kandungan *moisture*, *volatile matter* (zat terbang) dan kadar abu dihilangkan.
11. *Total Sulfur* (TS) adalah jumlah kandungan sulfur yang terdapat dalam abu batubara (disebut pula *noncombustible* sulfur) dengan *combustible* sulfur. Atau definisi lainnya menyebutkan, total sulfur adalah jumlah *inorganic* sulfur dengan organik sulfur.
12. *Total Moisture* (TM) adalah kadar air yang terdapat pada permukaan luar batubara yang sangat dipengaruhi oleh iklim dan lingkungan dimana batubara itu berada.
13. *Gross Heating Value* (GHV) adalah harga yang dihasilkan oleh bahan bakar gas.

3.4 Struktur Data

Struktur data dari variabel indeks batubara dan variabel prediktornya ditampilkan pada Tabel 3.2.

Tabel 3.2 Struktur Data

I	Y_i	X_{1,i}	X_{2,i}	X_{3,i}	...	X_{17,i}
1	Y ₁	X _{1,1}	X _{2,1}	X _{3,1}	...	X _{17,1}
2	Y ₂	X _{1,2}	X _{2,2}	X _{3,2}	...	X _{17,2}
3	Y ₃	X _{1,3}	X _{2,3}	X _{3,3}	...	X _{17,3}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
213	Y ₂₁₃	X _{1,213}	X _{2,213}	X _{3,213}	...	X _{17,213}

Keterangan:

i = Pengamatan ke- i

n = Jumlah pengamatan

3.5 Langkah Analisis

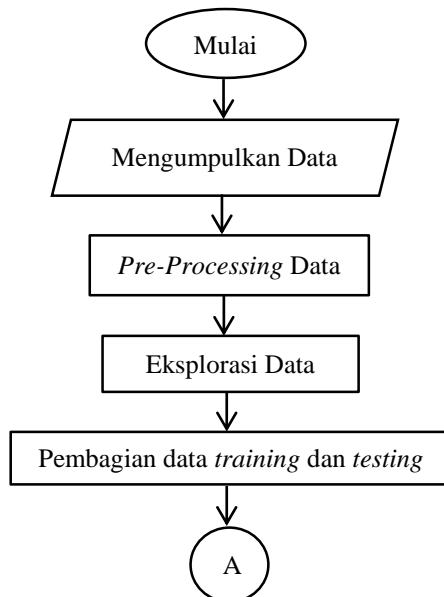
Langkah analisis yang akan dilakukan pada penelitian ini yaitu sebagai berikut.

1. Mengumpulkan data sekunder indeks batubara di PT XYZ.
2. Melakukan *pre-processing* pada data indeks batubara.
3. Melakukan eksplorasi data indeks batubara.
4. Melakukan pembagian data *training* sebesar 90% dan data *testing* sebesar 10%.
5. Melakukan pemilihan variabel dengan RFE dengan berbasis *Random Forest* (RF-RFE) untuk menentukan variabel yang akan dianalisis selanjutnya.
6. Pemodelan regresi linier menggunakan persamaan (2.1) dan pemodelan regresi *Ridge* menggunakan persamaan (2.4).
 - a. Pemodelan dengan regresi linier didahului dengan pemeriksaan hubungan antara semua variabel, jika terdapat hubungan antara variabel prediktor dengan respon dan telah diindikasikan terdapat kasus multikolinieritas maka pemodelan dilakukan menggunakan regresi *Ridge*.
 - b. Melakukan pemodelan regresi *Ridge*.
 - c. Pemeriksaan asumsi identik, independen, distribusi normal, dan asumsi non-multikolinieritas. Jika terdapat pelanggaran asumsi, maka dilakukan penanggulangan menggunakan transformasi.
 - d. Menghitung kebaikan model menggunakan nilai RMSE pada persamaan (2.38).
7. Melakukan pemodelan SVR dengan menggunakan persamaan (2.9) dengan fungsi kernel di persamaan (2.34).
 - a. Melakukan pemodelan dengan SVR menggunakan semua variabel (*complete feature*) dan beberapa variabel yang terpilih dari langkah 5 (*feature selection*).

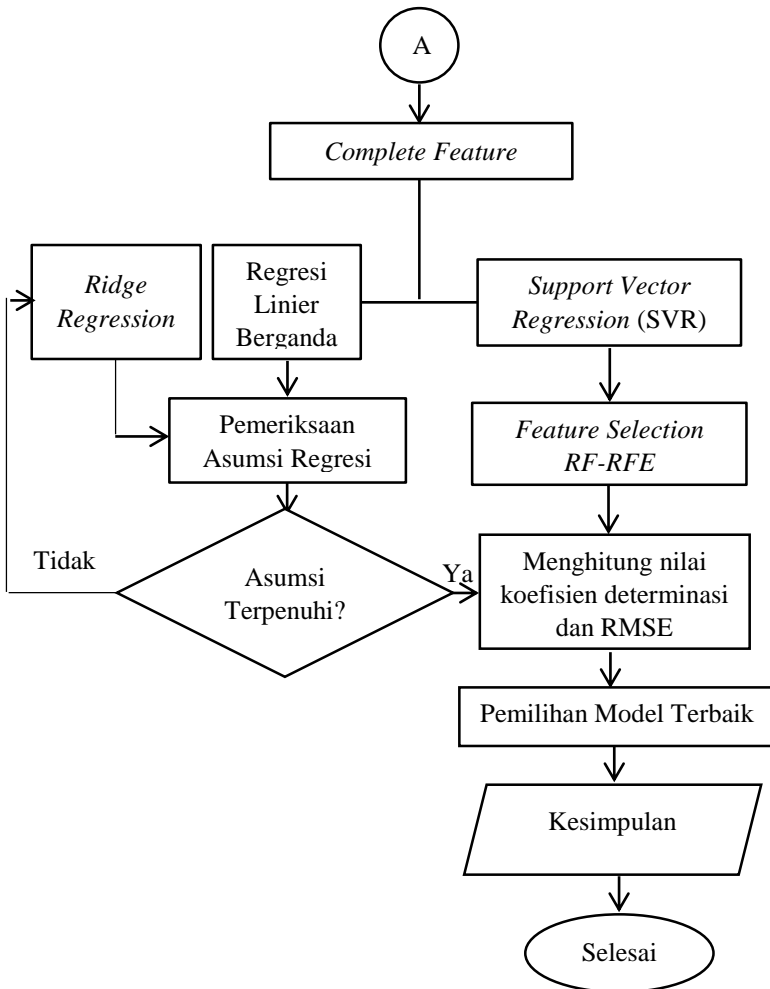
- b. Melakukan *tuning* parameter untuk mendapatkan parameter optimum dari ketiga kernel.
 - c. Menentukan range nilai parameter C , ε dan γ untuk optimasi *hyperplane* pada data *training*
 - d. Melakukan pemodelan dengan SVR berdasarkan *range* nilai parameter.
 - e. Mendapatkan model dan menghitung nilai RMSE dengan persamaan (2.38).
8. Pemilihan model terbaik menggunakan kriteria nilai RMSE di persamaan (2.38).
 9. Menarik kesimpulan dan saran.

3.6 Diagram Alir Penelitian

Berikut merupakan diagram alir yang dilakukan pada penelitian.



Gambar 3.1 Diagram Alir



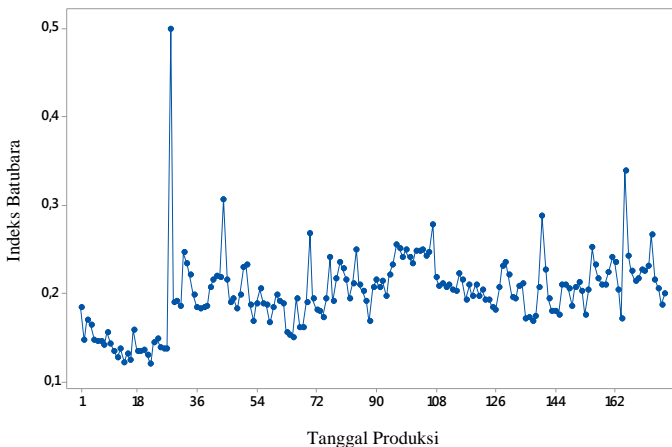
Gambar 3.1 Diagram Alir (Lanjutan)

BAB IV ANALISIS DAN PEMBAHASAN

Analisis dan pembahasan pada penelitian ini mencakup beberapa tahap yaitu eksplorasi data, *feature selection* dengan RFE, pembagian data *training* dan *testing*, analisis regresi linier berganda serta identifikasi asumsinya, kemudian penggunaan metode *Support Vector Regression* (SVR) serta pemilihan model terbaik. Tahapan prediksi akan dilakukan setelah mendapatkan model terbaik.

4.1 Eksplorasi Indeks Batubara dengan Tiga Kelompok Prediktor

Eksplorasi data bertujuan untuk mendapatkan karakteristik data secara visual. Secara visual dapat ditunjukkan dengan *time series* plot dan deskriptif. Berikut adalah pola indeks batubara dari bulan Januari 2019 hingga Juli 2019 di PT XYZ, dimana indeks batubara didapatkan dari jumlah produksi *clinker* dibagi dengan jumlah batubara dengan satuan ton.



Gambar 4.1 Pola Data Indeks Batubara PT XYZ

Hasil plot pada Gambar 4.1 menunjukkan pola Indeks batubara PT XYZ pada bulan Januari 2019 hingga Juli 2019 mengalami fluktuasi. Pada akhir bulan Januari indeks batubara di PT XYZ mengalami peningkatan yang sangat tinggi yaitu sebesar 0,5, hal ini disebabkan adanya *trouble* pada alat perekap data. Data *outlier* selanjutnya dibuang untuk mengatasi ketidak-normalan residual. Sebanyak 213 observasi dilakukan pembuangan *outlier*, dan jumlah observasi setelah dilakukan pembuangan *outlier* adalah sebanyak 177 observasi.

Pada bagian ini akan ditampilkan statistik deskriptif yang terdiri dari ukuran pemusatan dan penyebaran data. Ukuran pemusatan data meliputi rata-rata hitung, nilai minimum dan nilai maksimum. Ukuran penyebaran data dapat dilihat dari nilai varians. Varians adalah kuadrat simpangan nilai data terhadap *mean* dari data. Statistik deskriptif untuk setiap variabel berdasarkan rumus pada Lampiran 6 dapat dilihat pada Tabel 4.1 berikut.

Tabel 4.1 Deskriptif Indeks Batubara dan Tiga Kelompok Prediktor

Variabel	Mean	Variance	Minimum	Maximum
Indeks Batubara	0,19992	0,00183	0,1194	0,49929
Operasional				
Speed Kilen	56,334	145,974	3,36	66,782
ILC Exit Temp	9596,2	268475,8	4915,3	10747,6
SLC Exit Temp	9679,6	717559,8	3384	10663,7
Temp Stage 4 ILC	20093	1648717	9255	20657
Temp Stage 4 SLC	20321	3879282	8670	21294
ILC Coal	237,68	9926,59	18,34	408,33
SLC Coal	252,99	10665,47	29,43	535,35
Kiln Coal	464,35	6631,83	110,69	626,31
Bahan Baku				
LSF	2184,4	162114,4	387,4	3202,9
SIM	50,384	90,848	9,04	75,78
ALM	36,402	46,083	6,55	47,6

Tabel 4.1 Deskriptif Indeks Batubara dan Tiga Kelompok Prediktor (Lanjutan)

	Kualitas Batubara			
AC	3,2795	0,3082	2,01	4,62
VM	31,439	6,671	16,75	41,434
FC	30,557	6,563	18,34	34,81
TS	0,16211	0,00079	0,1	0,26
TM	34,723	6,041	31,36	60,53
GHV	4132,7	33942,3	2121	4285

Karakteristik indeks batubara dan variabel prediktornya dapat dilihat pada Tabel 4.1 dan dapat diketahui bahwa rata-rata indeks batubara PT XYZ pada bulan Januari 2019 hingga Juli 2019 adalah sebesar 0,199 dengan indeks batubara tertinggi sebesar 0,499 ton dan indeks batubara terendah yaitu sebesar 0,119. Rata-rata *speed kiln* atau kecepatan mesin *kiln* di PT XYZ pada bulan Januari 2019 hingga Juli 2019 adalah sebesar 56,33 Rpm dengan kecepatan tertinggi sebesar 66,78 Rpm dan kecepatan terendah yaitu sebesar 3,36 Rpm. Rata-rata ILC dan SLC *Exit temperature* secara berturut-turut adalah sebesar 9596,2°C dan 9679,6°C, sedangkan rata-rata *temperature stage 4* ILC dan SLC secara berturut-turut adalah sebesar 20093°C dan 20321°C. ILC, SLC, *kiln coal* memiliki rata-rata berat secara berturut-turut sebesar 237,68 ton, 252,99 ton, dan 464,35 ton. Rata-rata LSF (*Lime Saturation Factor*), SIM (*Silica Modulus*), ALM (*Alumina Modulus*) secara berturut-turut adalah sebesar 2184, 50, dan 36 dengan satuan dimensi. Sedangkan, untuk AC (*Ash Content*), VM (*Volatile Matter*), FC (*Fixed Carbon*), TS (*Total Sulfur*), TM (*Total Moisture*), dan GHV (*Gross Heating Value*) secara berturut-turut adalah sebesar 3,27%, 31,43%, 30,55%, 0,162%, 34,72%, dan 4132,7 cal/g.

4.2 Pemodelan Menggunakan *Ridge Regression*

Pada penelitian ini dilakukan pemodelan indeks batubara dengan semua prediktor yang terdeteksi berpengaruh menggunakan metode regresi linier berganda.

4.2.1 Tahapan Pemodelan Indeks Batubara Menggunakan OLS

Pengujian signifikansi model regresi secara serentak dilakukan untuk menguji variabel prediktor secara bersama-sama berpengaruh terhadap model. Hipotesis yang digunakan adalah:

$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{17} = 0$ (tidak ada hubungan linier antara variabel-variabel bebas dengan indeks batubara)

$H_0: \exists \beta_j \neq 0, j = 1, 2, 3, \dots, 17$ (minimal terdapat satu variabel yang berhubungan linier dengan indeks batubara),

Dengan tingkat signifikansi $\alpha = 0,05$, pengujian serentak model regresi dapat dilihat pada Tabel 4.2 yaitu dari nilai deviansi yang terbentuk. Berdasarkan hasil analisis diperoleh nilai $D(\hat{\beta})$ sebesar 236,12 dan nilai $\chi^2_{(0,05;17)} = 27,587$.

Tabel 4.2 Hasil Uji Serentak (ANOVA)

Model	Sum of Squares	df	Mean Square	F _h	P-Value	R ²
Regression	78,6319	17	4,62541	236,12	0,000	97,24%
Error	2,2331	114	0,01959			
Total	80,8651	131				

Keputusan pengujian ini tolak H_0 karena nilai $D(\hat{\beta}) > \chi^2_{(0,05;17)}$ dan *p-value* kurang dari 0,05 yang berarti bahwa seluruh para-meter secara bersama-sama mempunyai pengaruh dalam model.

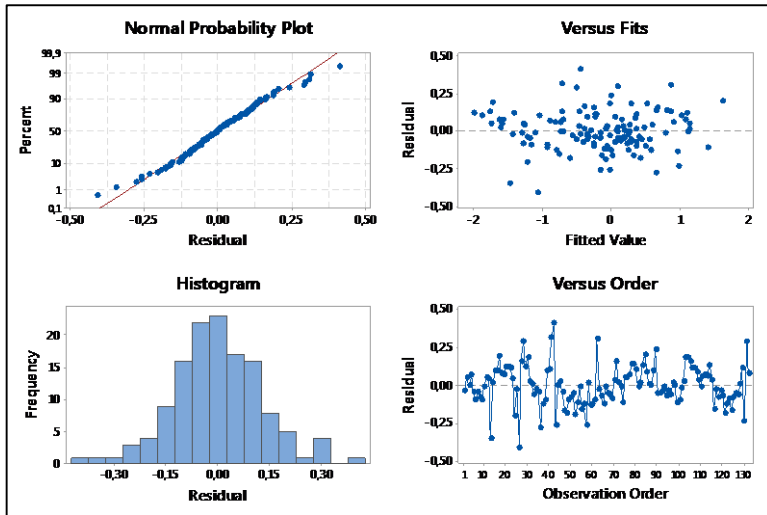
Dalam analisis regresi, kasus multikolinieritas memiliki pengaruh besar terhadap hasil estimasi parameter. Oleh sebab itu, sebelum melakukan analisis lebih lanjut dengan regresi linier berganda, terlebih dahulu akan dilakukan pemeriksaan terhadap kasus multikolinieritas antar variabel prediktor. Salah satu cara untuk memeriksa adanya kasus multikolinieritas adalah melihat nilai VIF (*Variance Inflation Factors*). Nilai VIF yang lebih dari 10 me-

rupakan indikasi adanya multikolinieritas. Berikut hasil pemeriksaan multikolinieritas disajikan pada Tabel 4.3.

Tabel 4.3 Deteksi Multikolinieritas

Model	B	Std. Error	T	P-Value	VIF
<i>Constant</i>	-0,07208	0,01643	-4,39	0	
X ₁	-0,61975	0,06059	-10,23	0	6,44
X ₂	0,11801	0,03537	3,34	0,001	2,32
X ₃	0,14379	0,04847	2,97	0,004	2,73
X ₄	-0,1155	0,146	-0,79	0,430	10,21
X ₅	-0,3117	0,1077	-2,89	0,005	10,28
X ₆	0,43739	0,02024	21,61	0	2,79
X ₇	0,47878	0,02255	21,23	0	2,86
X ₈	0,32082	0,03917	8,19	0	3,12
X ₉	-0,0359	0,1213	-0,3	0,768	21,03
X ₁₀	-0,0852	0,07906	-1,08	0,283	9,66
X ₁₁	0,08094	0,08822	0,92	0,361	13,09
X ₁₂	0,884	1,344	0,66	0,512	11560,36
X ₁₃	4,268	6,266	0,68	0,497	89353,17
X ₁₄	4,283	6,223	0,69	0,493	79331,7
X ₁₅	-0,01073	0,01619	-0,66	0,509	1,66
X ₁₆	3,966	5,957	0,67	0,507	78295,06
X ₁₇	-0,05659	0,06911	-0,82	0,415	10,82

Dapat dilihat pada Tabel 4.3 bahwa variabel yang mempengaruhi indeks batubara secara signifikan dengan α sebesar 0,05 adalah variabel *speed kiln*, *ILC Exit temperature*, *SLC Exit Temperature*, *Temperature Stage 4 SLC*, *ILC Coal*, *SLC Coal*, dan *Kiln Coal*. Tabel 4.3 juga menunjukkan bahwa terdapat variabel prediktor yang memiliki nilai VIF lebih dari 10, maka dapat disimpulkan bahwa terjadi kasus multikolinieritas. Artinya terdapat keterkaitan antara variabel prediktor yang akan digunakan untuk memodelkan indeks batubara PT XYZ. Kesimpulan model regresi OLS yang didapatkan kurang baik. Selanjutnya dilakukan identifikasi asumsi regresi, yaitu residual berdistribusi normal, heteroskedastisitas, dan autokorelasi dengan hasil plot yang ada pada Gambar 4.2



Gambar 4.2 Asumsi IIDN Data Indeks Batubara

Dapat dilihat secara visual dari Gambar 4.2 bahwa pada pendeteksian asumsi kenormalan dapat dilihat dari gambar *normal probability* plot yang terlihat bahwa plot mengikuti garis kenormalan (warna merah), sehingga secara visual telah mengikuti asumsi distribusi normal, hal ini didukung oleh gambar histogram yang membentuk distribusi normal. Gambar *versus fits* menjelaskan bahwa data tidak membentuk suatu pola dan tersebar merata sehingga secara visual dapat diputuskan bahwa data bersifat identik, sedangkan pada gambar *versus order* yang memperlihatkan secara visual bahwa data tersebar merata di atas maupun di bawah angka nol, sehingga secara visual dapat diputuskan bahwa telah memenuhi asumsi independen. Adanya asumsi multikolinieritas yang tidak terpenuhi yang akan memberikan berbagai dampak, maka kasus multikolinieritas perlu diatasi dengan menggunakan regresi *ridge*. Regresi *ridge* adalah metode yang digunakan untuk mengatasi masalah multikolinieritas, dimana metode ini merupakan modifikasi dari metode kuadrat terkecil yang menghasilkan penduga yang bias dengan menggunakan *ridge trace*. Nilai VIF untuk masing-masing nilai c ($0 \leq c \leq 1$) dengan menggunakan ma-

triks korelasi dari data indeks batubara PT XYZ pada bulan Januari 2019 hingga Juli 2019, maka diperoleh nilai VIF yang ditunjukkan pada Tabel 4.4, dengan hasil perhitungan VIF setiap prediktor yang terdapat pada Lampiran 3.

Tabel 4.4 Nilai VIF dengan Berbagai Nilai c

Nilai c	VIF β^*_1 (c)	VIF β^*_2 (c)	VIF β^*_3 (c)	VIF β^*_4 (c)	VIF β^*_5 (c)	VIF β^*_6 (c)
0	6,4422	2,3203	2,7316	10,213	10,275	2,7923
0,001	5,749	2,0516	2,6495	9,5	9,6934	2,5557
0,002	5,6485	2,0216	2,6233	9,1966	9,4096	2,5234
0,003	5,5514	1,9929	2,5979	8,9093	9,1394	2,4927
0,004	5,4576	1,9655	2,5732	8,6371	8,8817	2,4633
0,005	5,3668	1,9393	2,5492	8,3788	8,6358	2,435
0,006	5,279	1,9143	2,5259	8,1334	8,401	2,4079
0,007	5,1938	1,8902	2,5032	7,9	8,1766	2,3818
0,008	5,1113	1,8671	2,481	7,6779	7,9619	2,3566
0,009	5,0312	1,8449	2,4593	7,4662	7,7565	2,3323
0,01	4,9534	1,8236	2,4381	7,2643	7,5597	2,3087
0,02	4,2838	1,6459	2,2487	5,6673	5,9738	2,1074
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	0,1271	0,2227	0,1876	0,112	0,101	0,1935

Tabel 4.4 Nilai VIF dengan Berbagai Nilai c (Lanjutan)

Nilai c	VIF β^*_7 (c)	VIF β^*_8 (c)	VIF β^*_9 (c)	VIF β^*_{10} (c)	VIF β^*_{11} (c)	VIF β^*_{12} (c)
0	2,8638	3,1194	21,027	9,6571	13,089	11560,4
0,001	2,7807	2,978	18,897	9,2688	12,101	3,1254
0,002	2,7409	2,9447	17,787	8,9128	11,527	2,9077
0,003	2,7031	2,9126	16,776	8,5834	11,001	2,7999
0,004	2,6671	2,8817	15,852	8,2778	10,516	2,7126
0,005	2,6327	2,8519	15,005	7,9935	10,069	2,6342
0,006	2,5999	2,823	14,227	7,7283	9,6554	2,5617
0,007	2,5684	2,795	13,511	7,4803	9,2717	2,494
0,008	2,5381	2,7679	12,851	7,2479	8,9149	2,4302
0,009	2,5091	2,7415	12,239	7,0296	8,5825	2,37
0,01	2,4811	2,7159	11,673	6,8242	8,2721	2,3131

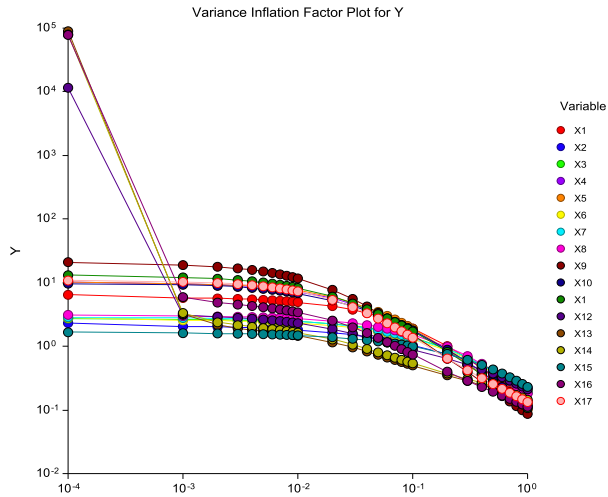
Tabel 4.4 Nilai VIF dengan Berbagai Nilai c (Lanjutan)

Nilai c	VIF $\beta^*_7(c)$	VIF $\beta^*_8(c)$	VIF $\beta^*_9(c)$	VIF $\beta^*_{10}(c)$	VIF $\beta^*_{11}(c)$	VIF $\beta^*_{12}(c)$
0,02	2,2465	2,4907	7,7257	5,2815	6,027	1,8759
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	0,1887	0,1805	0,0866	0,1166	0,1066	0,2107

Tabel 4.4 Nilai VIF dengan Berbagai Nilai c (Lanjutan)

Nilai c	VIF $\beta^*_{13}(c)$	VIF $\beta^*_{14}(c)$	VIF $\beta^*_{15}(c)$	VIF $\beta^*_{16}(c)$	VIF $\beta^*_{17}(c)$
0	89353,17	79331,7	1,6621	78295,06	10,8151
0,001	3,2174	3,3371	1,6093	5,9282	10,2925
0,002	2,1613	2,3795	1,5934	4,8644	9,8714
0,003	1,918	2,145	1,5782	4,5222	9,4773
0,004	1,7982	2,0216	1,5637	4,2969	9,108
0,005	1,7165	1,933	1,5498	4,1128	8,7614
0,006	1,6515	1,8604	1,5365	3,9505	8,4356
0,007	1,5958	1,797	1,5236	3,8027	8,1289
0,008	1,546	1,7398	1,5113	3,6662	7,8398
0,009	1,5007	1,6872	1,4994	3,539	7,567
0,01	1,4587	1,6385	1,4879	3,4199	7,3092
0,02	1,1522	1,2811	1,39	2,5338	5,3565
⋮	⋮	⋮	⋮	⋮	⋮
1	0,1355	0,1447	0,2301	0,112	0,1332

Tabel 4.4 menunjukkan bahwa nilai VIF semakin kecil jika ditambahkan tetapan bias (c), karena VIF yang optimal adalah bila nilai $VIF < 10$ maka didapatkan nilai c yang paling optimum adalah sebesar 0,02 karena memiliki nilai VIF seluruh variabel kurang dari 10, sehingga tidak ada gejala multikolinieritas pada setiap variabel prediktor dari indeks batubara. Maka digunakan koefisien bias (c) sebesar 0,02 untuk estimasi parameter untuk mendapatkan model yang tidak memiliki kasus multikolinieritas pada setiap variabel prediktor. Hal ini didukung dengan grafik VIF yang menunjukkan bahwa garis mulai cenderung konstan ketika berada pada koefisien bias (c) sebesar 0,02. Grafik VIF dengan berbagai nilai c ditunjukkan oleh gambar berikut :



Gambar 4.3 Grafik VIF dengan Berbagai Nilai c

Untuk mendapatkan nilai β^* yang sesuai dengan nilai $c=0,02$ yaitu dengan menggunakan persamaan regresi *ridge* atau dapat melihat tabel berikut :

Tabel 4.5 Koefisien Regresi Ridge dengan Berbagai Nilai c

Nilai c	β^*_1	β^*_2	β^*_3	β^*_4	β^*_5	β^*_6
0	-0,4041	0,0791	0,0763	-0,0394	-0,1444	0,5619
0,001	-0,394	0,0846	0,0778	-0,0466	-0,1379	0,5657
0,002	-0,392	0,085	0,0769	-0,0473	-0,1371	0,5648
0,003	-0,39	0,0853	0,076	-0,0479	-0,1363	0,5639
0,004	-0,3881	0,0857	0,0751	-0,0485	-0,1355	0,563
0,005	-0,3862	0,086	0,0741	-0,0491	-0,1347	0,5622
0,006	-0,3843	0,0863	0,0732	-0,0496	-0,1339	0,5613
0,007	-0,3824	0,0866	0,0723	-0,0501	-0,1331	0,5604
0,008	-0,3806	0,0869	0,0714	-0,0506	-0,1324	0,5596
0,009	-0,3788	0,0872	0,0705	-0,0511	-0,1317	0,5587
0,01	-0,377	0,0875	0,0696	-0,0516	-0,131	0,5579
0,02	-0,3604	0,0897	0,0606	-0,0553	-0,1244	0,5498
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	-0,0872	0,048	-0,1106	-0,0552	-0,0366	0,2958

Tabel 4.5 Koefisien Regresi Ridge dengan Berbagai Nilai c (Lanjutan)

Nilai c	β^*_7	β^*_8	β^*_9	β^*_{10}	β^*_{11}	β^*_{12}
0	0,5591	0,2251	-0,0211	-0,0521	0,0517	1,1002
0,001	0,5562	0,2211	-0,0321	-0,053	0,0588	-0,0293
0,002	0,5553	0,2206	-0,033	-0,0534	0,0596	-0,0315
0,003	0,5545	0,22	-0,0339	-0,0538	0,0603	-0,0322
0,004	0,5537	0,2195	-0,0347	-0,0542	0,061	-0,0326
0,005	0,5529	0,2189	-0,0354	-0,0546	0,0617	-0,0329
0,006	0,5521	0,2184	-0,0362	-0,055	0,0623	-0,0331
0,007	0,5512	0,2178	-0,0369	-0,0554	0,0629	-0,0332
0,008	0,5504	0,2173	-0,0375	-0,0558	0,0635	-0,0334
0,009	0,5496	0,2168	-0,0382	-0,0563	0,064	-0,0335
0,01	0,5489	0,2162	-0,0388	-0,0567	0,0645	-0,0336
0,02	0,5411	0,2113	-0,0436	-0,0606	0,0687	-0,0347
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	0,2668	0,1031	-0,0415	-0,0569	0,0566	-0,0588

Tabel 4.5 Koefisien Regresi Ridge dengan Berbagai Nilai c (Lanjutan)

Nilai c	β^*_{13}	β^*_{14}	β^*_{15}	β^*_{16}	β^*_{17}
0	3,1686	3,017	-0,0133	2,9	-0,0419
0,001	0,0276	0,0577	-0,0151	-0,0393	-0,0384
0,002	0,0212	0,0519	-0,015	-0,0445	-0,0378
0,003	0,0188	0,05	-0,0148	-0,046	-0,0372
0,004	0,0174	0,049	-0,0146	-0,0465	-0,0366
0,005	0,0165	0,0484	-0,0145	-0,0466	-0,0361
0,006	0,0158	0,048	-0,0143	-0,0466	-0,0355
0,007	0,0151	0,0477	-0,0142	-0,0465	-0,035
0,008	0,0146	0,0475	-0,014	-0,0464	-0,0346
0,009	0,0141	0,0474	-0,0138	-0,0462	-0,0341
0,01	0,0137	0,0473	-0,0137	-0,046	-0,0337
0,02	0,0105	0,0471	-0,0121	-0,0438	-0,0302
⋮	⋮	⋮	⋮	⋮	⋮
1	-0,0271	0,0557	0,0136	-0,0042	-0,0129

Tabel di atas model regresi *ridge* yang diperoleh jika nilai c sebesar 0,02 yaitu :

$$\begin{aligned}
 Y^* = & -0,3604 Z_1 + 0,0897 Z_2 + 0,0606 Z_3 - 0,0553 Z_4 - 0,1244 Z_5 \\
 & + 0,5498 Z_6 + 0,5411 Z_7 + 0,2113 Z_8 - 0,0436 Z_9 - 0,0606 Z_{10} \\
 & + 0,0687 Z_{11} - 0,0347 Z_{12} + 0,0105 Z_{13} + 0,0471 Z_{14} - 0,0121 \\
 & Z_{15} - 0,0438 Z_{16} - 0,0302 Z_{17}
 \end{aligned}$$

Setelah didapatkan model Y^* , selanjutnya dilakukan transformasi ke bentuk awal dengan persamaan (2.5), sehingga nilai koefisien regresi menjadi sebagai berikut.

Tabel 4.6 Koefisien Parameter Regresi *Ridge*

$\tilde{\beta}_k$	Koefisien Parameter
$\tilde{\beta}_1$	-0,5527
$\tilde{\beta}_2$	0,1338
$\tilde{\beta}_3$	0,1142
$\tilde{\beta}_4$	-0,1623
$\tilde{\beta}_5$	-0,2686
$\tilde{\beta}_6$	0,4279
$\tilde{\beta}_7$	0,4634
$\tilde{\beta}_8$	0,3011
$\tilde{\beta}_9$	-0,0741
$\tilde{\beta}_{10}$	-0,0990
$\tilde{\beta}_{11}$	0,1076
$\tilde{\beta}_{12}$	-0,0279
$\tilde{\beta}_{13}$	0,0141
$\tilde{\beta}_{14}$	0,0668
$\tilde{\beta}_{15}$	-0,0098
$\tilde{\beta}_{16}$	-0,0599
$\tilde{\beta}_{17}$	-0,0408

$$\begin{aligned}
\tilde{\beta}_0 &= \bar{Y} - \tilde{\beta}_1 \bar{X}_1 - \tilde{\beta}_2 \bar{X}_2 - \tilde{\beta}_3 \bar{X}_3 - \tilde{\beta}_4 \bar{X}_4 - \tilde{\beta}_5 \bar{X}_5 - \tilde{\beta}_6 \bar{X}_6 - \tilde{\beta}_7 \bar{X}_7 - \tilde{\beta}_8 \bar{X}_8 - \tilde{\beta}_9 \bar{X}_9 \\
&= -\tilde{\beta}_{10} \bar{X}_{10} - \tilde{\beta}_{11} \bar{X}_{11} - \tilde{\beta}_{12} \bar{X}_{12} - \tilde{\beta}_{13} \bar{X}_{13} - \tilde{\beta}_{14} \bar{X}_{14} - \tilde{\beta}_{15} \bar{X}_{15} - \tilde{\beta}_{16} \bar{X}_{16} - \tilde{\beta}_{17} \bar{X}_{17} \\
&= -0.1903 - (-0,5527)(0,2664) - (0,1338)(0,0135) - (0,1142)(0,1804) \\
&= -(-0,1623)(0,2072) - (-0,2686)(0,2572) - (0,4279)(-0,0335) \\
&= -(0,4634)(0,1452) - (0,3011)(0,2370) - (-0,0741)(0,2392) \\
&= -(-0,0990)(0,2182) - (0,1076)(0,2404) - (-0,0279)(0) \\
&= (0,0141)(-0,1348) - (0,0668)(0,105) - (-0,0098)(-0,1035) \\
&= (-0,0599)(0,0317) - (-0,0409)(0,0136) \\
&= -0,07726
\end{aligned}$$

Sehingga model regresi *ridge* yang diperoleh adalah

$$\begin{aligned}
Y_{RT} &= -0,0773 - 0,5527X_1 + 0,1338X_2 + 0,1142X_3 - 0,1623X_4 \\
&\quad - 0,2686X_5 + 0,4279X_6 + 0,4634X_7 + 0,3011X_8 - 0,0741X_9 \\
&\quad - 0,0990X_{10} + 0,1076X_{11} - 0,0279X_{12} + 0,0141X_{13} \\
&\quad + 0,0668X_{14} - 0,0098X_{15} - 0,0599X_{16} - 0,0408X_{17}
\end{aligned}$$

Selanjutnya, dilakukan perhitungan nilai RMSE dari model *Ridge Regression* dengan c sebesar 0,02, yang akan digunakan untuk dibandingkan dengan metode *Support Vector Regression (SVR) complete feature* dan *feature selection (RF-RFE)*. Hasil perhitungan RMSE menggunakan metode *Ridge Regression* dengan nilai c sebesar 0,02 adalah sebesar 2,088. Nilai RMSE tersebut akan dibandingkan dengan nilai RMSE metode *Support Vector Regression (SVR) complete feature* dan *feature selection (RF-RFE)*.

4.3 Support Vector Regression (SVR)

Selanjutnya, dilakukan pemodelan menggunakan *Support Vector Regression (SVR)* dengan *complete feature* dan *feature selection* menggunakan tiga kernel, yaitu kernel linier, polynomial, dan RBF. Hasil yang didapatkan akan dibandingkan menggunakan nilai RMSE untuk mendapatkan model dengan parameter yang optimum.

4.3.1 SVR Dengan *Complete Feature*

Support Vector Regression (SVR) dengan *complete feature* pada penelitian ini terdiri dari 17 variabel prediktor dan 1 variabel respon yaitu indeks batubara. Pembentukan model SVR disertai parameter *tuning* menggunakan metode *grid search* untuk mendapatkan prediksi parameter yang paling optimum. Untuk pemodelan kali ini menggunakan semua data indeks batubara dari bulan Januari 2019 hingga Juli 2019 dengan 17 variabel, pemodelan meliputi Kernel-Linier, Kernel-Polynomial dan Kernel-*Radial Basis Function (RBF)*.

Kernel Linier merupakan sebuah metode yang bila digunakan, data akan terpisah oleh sebuah garis linier yang disebut *hyperplane*. Hasil penggunaan kernel-linier dengan seluruh data indeks batubara mendapatkan nilai parameter *cost (C)* yaitu senilai 1, dengan nilai kriteria RMSE yaitu sebesar 0,6892864. Hal ini menyatakan bahwa dengan metode kernel-linier *complete feature*, model memiliki nilai akurasi prediksi sebesar 0,6892864, yang kemudian nilai ini akan dibandingkan dengan metode lainnya.

Kernel-polynomial merupakan kernel yang memetakan suatu data ke dimensi yang dinamakan *feature space*. Kernel-polynomial memiliki fungsi khusus untuk memetakan ke *feature space* yang biasanya berbentuk kurva parabola. Berikut adalah hasil *tuning* parameter untuk kernel-polynomial.

Tabel 4.6 *Tuning* Parameter Kernel-Polynomial

Degree	Sigma	C	RMSE
1	0,001	0,25	0,994
1	0,001	0,50	0,969
1	0,001	1,00	0,923
1	0,010	0,25	0,823
1	0,010	0,50	0,725
1	0,010	1,00	0,674
1	0,100	0,25	0,642
1	0,100	0,50	0,624
1	0,100	1,00	0,619
2	0,001	0,25	0,968

Tabel 4.7 *Tuning* Parameter Kernel-Polynomial (Lanjutan)

Degree	Sigma	C	RMSE
2	0,001	0,50	0,923
2	0,001	1,00	0,852
2	0,010	0,25	0,745
2	0,010	0,50	0,711
2	0,010	1,00	0,699
2	0,100	0,25	1,010
2	0,100	0,50	1,077
2	0,100	1,00	1,099
3	0,001	0,25	0,945
3	0,001	0,50	0,886
3	0,001	1,00	0,801
3	0,010	0,25	0,751
3	0,010	0,50	0,733
3	0,010	1,00	0,746
3	0,100	0,25	1,734
3	0,100	0,50	1,950
3	0,100	1,00	1,958

Hasil *tuning* parameter pada Tabel 4.7 untuk kernel-polynomial, didapatkan hasil bahwa terdapat dua parameter yang digunakan yaitu sigma dan *cost* (C) dengan kriteria terbaik yaitu RMSE sebesar 0,619, didapatkan parameter yang paling optimal adalah 0,100 (sigma) dan 1 (*cost*) dengan nilai *degree* bernilai 1. Dengan kernel-polynomial, akurasi prediksi menggunakan model dan parameternya memiliki nilai sebesar 0,619 yang kemudian dilakukan perbandingan terhadap semua model yang telah didapatkan.

Kernel-RBF (*Radial Basis Function*) merupakan salah satu dari beberapa kernel untuk kasus non-linier. Kernel ini memetakan suatu data ke dimensi yang lebih tinggi dan membentuk kurva yang fleksibel sehingga dapat mengikuti pola data yang digunakan. Berikut ini adalah hasil dari *tuning* parameter dengan metode kernel-RBF.

Tabel 4.8 *Tuning* Parameter Kernel-RBF

Sigma	C	RMSE
0,01127152	0,25	0,775
0,01127152	0,50	0,717
0,01127152	1,00	0,704
0,08480589	0,25	0,750
0,08480589	0,50	0,725
0,08480589	1,00	0,714
0,15834027	0,25	0,781
0,15834027	0,50	0,752
0,15834027	1,00	0,737

Tabel 4.8 menunjukkan hasil *tuning* parameter untuk SVR dengan kernel-RBF dengan keseluruhan data indeks batubara. Terdapat dua parameter untuk kernel-RBF yaitu sigma dan *cost* (C). Dengan kriteria RMSE sebesar 0,704 maka parameter yang terpilih dan paling optimal untuk data indeks batubara ini adalah sigma bernilai 0,01127152 dan nilai *cost* sebesar 1. Dengan model kernel-RBF ini, maka akurasi prediksi menggunakan model dan parameter terpilihnya memiliki nilai sebesar 0,704, yang kemudian akan dibandingkan.

4.3.2 SVR Dengan *Feature Selection*

Seleksi fitur digunakan untuk mereduksi variabel yang dalam suatu penelitian memiliki banyak fitur (variabel), yang bertujuan untuk mengurangi dimensi yang berlebihan. Seleksi fitur biasanya digunakan dalam *machine learning*. RFE merupakan salah satu *feature selection* yang menggunakan ranking sebagai tingkat pentingnya terhadap proses prediksi. Pada setiap iterasi, ranking pentingnya fitur diukur dan fitur yang kurang relevan dihilangkan. Fitur seleksi kali ini digabungkan dengan metode *Random Forest* yang metode tersebut cocok pada data yang berdimensi tinggi dan memiliki hubungan variabel prediktor didalamnya. Berikut adalah hasilnya.

Tabel 4.9 Ranking RF-RFE

<i>Ranking</i>	<i>Variabel</i>	<i>RMSE</i>
1	Speed Kiln	0,8773
2	ILC Exit Temp	0,6938
3	SLC Exit Temp	0,6050
4	Temp Stage 4 ILC	0,6009
5	FC	0,5824
6	ALM	0,5809
7	Kiln Coal	0,5768
8	AC	0,5758
9	LSF	0,5752
10	VM	0,5734
11	SIM	0,5722
12	GHV	0,5717
13	TM	0,5689
14	SLC Coal	0,5688
15	TS	0,5687
16	ILC Coal	0,5671
17	Temp Stage 4 SLC	0,5658

Hasil *Recursive Feature Elimination* pada Tabel 4.9 menunjukkan bahwa berdasarkan ranking RF-RFE, variabel Speed Kiln, ILC Exit Temp, SLC Exit Temp, Temp Stage 4 ILC, dan FC adalah merupakan variabel dengan ranking tertinggi yang menurut kode RF-RFE kurang relevan, sehingga kelima variabel tersebut akan dihilangkan (tidak digunakan).

Selanjutnya, dilakukan pemodelan SVR dengan *feature selection* (RF-RFE) yang merupakan suatu pembentuk model dengan cara *tuning* parameter untuk mendapatkan parameter paling optimal, dengan menggunakan tiga fungsi kernel yaitu linier, polynomial, dan RBF.

Kernel linier merupakan sebuah metode yang ketika digunakan, data akan terpisah oleh sebuah garis linier yang disebut *hyperplane*. Hasil penggunaan kernel-linier dengan seluruh data indeks batubara mendapatkan nilai parameter *cost* (C) yaitu senilai 1, dengan nilai kriteria RMSE yaitu sebesar 0,8902573. Hal ini menyatakan bahwa dengan metode kernel-linier *complete feature*, model memiliki nilai akurasi prediksi sebesar 0,8902573, yang kemudian nilai ini akan dibandingkan dengan metode lainnya.

Kernel-polynomial merupakan kernel yang bersifat non-linier. Kernel ini memetakan suatu data ke dimensi yang dinamakan *feature space*. Kernel-polynomial memiliki fungsi khusus untuk memetakan ke *feature space* yang biasanya berbentuk kurva parabola. Berikut adalah hasil tuning parameter untuk kernel-polynomial.

Tabel 4.10 *Tuning* Parameter Kernel-Polynomial RF-RFE

<i>Degree</i>	<i>Sigma</i>	<i>C</i>	<i>RMSE</i>
1	0,001	0,25	1,005
1	0,001	0,50	0,989
1	0,001	1,00	0,959
1	0,010	0,25	0,890
1	0,010	0,50	0,825
1	0,010	1,00	0,792
1	0,100	0,25	0,796
1	0,100	0,50	0,799
1	0,100	1,00	0,796
2	0,001	0,25	0,989
2	0,001	0,50	0,959
2	0,001	1,00	0,909
2	0,010	0,25	0,830
2	0,010	0,50	0,810
2	0,010	1,00	0,812
2	0,100	0,25	1,088
2	0,100	0,50	1,223
2	0,100	1,00	1,325
3	0,001	0,25	0,973
3	0,001	0,50	0,931

Tabel 4.10 *Tuning* Parameter Kernel-Polynomial RF-RFE (Lanjutan)

<i>Degree</i>	<i>Sigma</i>	<i>C</i>	<i>RMSE</i>
3	0,001	1,00	0,873
3	0,010	0,25	0,831
3	0,010	0,50	0,836
3	0,010	1,00	0,837
3	0,100	0,25	1,856
3	0,100	0,50	2,702
3	0,100	1,00	3,070

Tabel 4.10 merupakan tuning parameter untuk kernel-polynomial, didapatkan hasil bahwa terdapat dua parameter yang digunakan yaitu sigma dan *cost* (C). Dengan kriteria terbaik yaitu RMSE sebesar 0,792, didapatkan parameter yang paling optimal adalah 0,010 (sigma) dan 1 (*cost*) dengan nilai *degree* bernilai 1. Dengan kernel-polynomial, akurasi prediksi menggunakan model dan parameternya memiliki nilai sebesar 0,792 yang kemudian dilakukan perbandingan terhadap semua model yang telah didapatkan. Selanjutnya dilakukan perhitungan SVR *feature selection* menggunakan kernel-RBF.

Kernel-RBF (*Radial Basis Function*) merupakan salah satu dari beberapa kernel untuk kasus non-linier. Kernel ini memetakan suatu data ke dimensi yang lebih tinggi dan membentuk kurva yang fleksibel sehingga dapat mengikuti pola data yang digunakan. Berikut ini adalah hasil dari *tuning* parameter dengan metode kernel-RBF.

Tabel 4.11 *Tuning* Parameter Kernel-RBF RF-RFE

<i>Sigma</i>	<i>C</i>	<i>RMSE</i>
0,01493841	0,25	0,822
0,01493841	0,50	0,799
0,01493841	1,00	0,797
0,09646718	0,25	0,760
0,09646718	0,50	0,742
0,09646718	1,00	0,735
0,17799595	0,25	0,779

Tabel 4.11 *Tuning* Parameter Kernel-RBF RF-RFE (Lanjutan)

Sigma	C	RMSE
0,17799595	0,50	0,755
0,17799595	1,00	0,742

Hasil *tuning* parameter untuk SVR dengan kernel-RBF didapatkan nilai RMSE sebesar 0.735, dengan parameter sigma bernilai 0,09646718 dan nilai c sebesar 1. Dengan model kernel-RBF ini, maka akurasi prediksi menggunakan model dan parameter terpilihnya memiliki nilai sebesar 0,735.

Pada hasil pemodelan menggunakan *feature selection* yaitu RF-RFE, menunjukkan RMSE yang berbeda jika dibandingkan dengan SVR *complete feature*. Hasil yang lebih baik ditunjukkan oleh SVR menggunakan *complete feature* pada model kernel-linier, kernel-polynomial, dan kernel-RBF, maka dalam penelitian ini tidak diperlukan untuk analisis lebih lanjut menggunakan *feature selection* (RF-RFE).

4.4 Pemilihan Model Indeks Batubara Terbaik

Pemilihan model terbaik dalam penelitian ini menggunakan kriteria RMSE. Model-model yang terbentuk akan dibandingkan dengan kriteria tersebut. Model yang akan dibandingkan adalah Regresi *Ridge*, SVR *complete feature*, dan SVR *feature selection*. SVR terbagi menjadi tiga kernel yaitu kernel-linier, kernel-polynomial, dan kernel-RBF. Berikut adalah hasilnya.

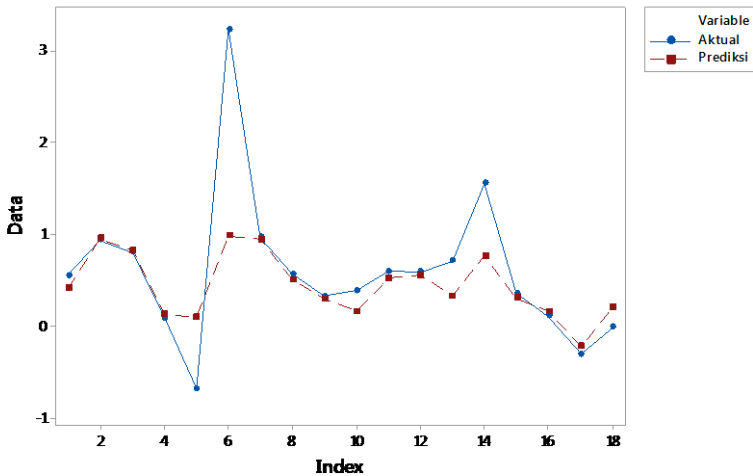
Tabel 4.12 Pemilihan Model Terbaik

No.	Metode	RMSE
1.	Regresi Ridge	2,088
	SVR – Linier	0,689
2.	Complete Polynomial	0,619
	Feature RBF	0,704
	SVR – Linier	0,890
3.	Feature Polynomial	0,792
	Selection RBF	0,735

Tabel 4.12 memberikan informasi bahwa, dari beberapa model yang terbentuk, kriteria model terbaik adalah pada metode

Support Vector Regression (SVR) Kernel-Polynomial dengan *Complete Feature* yang memiliki nilai kriteria RMSE bernilai 0,619, dengan parameter sigma bernilai 0,100 dan nilai *cost* sebesar 1.

Dengan model kernel-polynomial *complete feature* yang terpilih sebagai metode terbaik dalam penelitian ini disertai parameter pada Tabel 4.12, akan dilakukan prediksi indeks batubara, berikut adalah hasilnya.



Gambar 4.4 Prediksi Indeks Batubara

Gambar 4.4 memberikan informasi bahwa, terdapat 18 data indeks batubara. Secara visual nilai prediksi dengan model SVR kernel-polynomial telah mengikuti data aktual, namun pada observasi keenam dan kesepuluh cukup memiliki perbedaan yang jauh pada besaran indeks batubara, namun untuk observasi yang lain, telah mengikuti data aktual, dengan prediksi ini didapatkan nilai kesalahan RMSE adalah 0,619.

Tabel 4.13 Prediksi SVR Kernel Polynomial

Aktual	Prediksi
0.55370181	0.40630906
0.93818558	0.95195971

Tabel 4.13 Prediksi SVR Kernel Polynomial (Lanjutan)

Aktual	Prediksi
0.80308847	0.82027773
0.08717551	0.12605880
-0.68050335	0.09797804
3.23382710	0.98060407
0.96483638	0.94469045
0.56462345	0.50546523
0.32663371	0.29789054
0.38733326	0.15831093
0.59968584	0.51908278
0.58705272	0.55205516
0.70791839	0.32848158
1.55610059	0.76106988
0.35611551	0.30146852
0.10424052	0.15836224
-0.30456399	-0.22711545
-0.01114123	0.20457388

Nilai kesalahan RMSE pada metode SVR *complete feature* kernel polynomial ini lebih rendah dibandingkan dengan pemodelan *Ridge Regression*, SVR *Complete Feature* (kernel linier dan RBF), dan SVR *Feature Selection*.

(Halaman ini sengaja dikosongkan)

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan pada bab 4, maka diperoleh kesimpulan sebagai berikut.

1. Pola indeks batubara PT XYZ pada bulan Januari 2019 hingga Juli 2019 mengalami fluktuasi. Pada akhir bulan Januari indeks batubara di PT XYZ mengalami peningkatan yang sangat tinggi yaitu sebesar 0,5, hal ini disebabkan adanya *trouble* pada alat perekap data. Rata-rata indeks batubara PT XYZ pada bulan Januari 2019 hingga Juli 2019 adalah sebesar 0,199 dengan indeks batubara tertinggi sebesar 0,499 ton dan indeks batubara terendah yaitu sebesar 0,119.
2. Pada hasil analisis regresi linier teridentifikasi adanya multikolinieritas, sehingga perlu ditangani menggunakan metode *Ridge Regression*. Kemudian dilakukan pemodelan menggunakan metode *Support Vector Regression* (SVR) dengan fungsi kernel menghasilkan bahwa *complete feature* lebih memiliki hasil yang baik dibandingkan *feature selection* (RF-RFE). *Feature selection* berbasis *Random Forest* tidak cocok untuk digunakan penelitian ini karena dalam penelitian ini dibutuhkan variabel yang cukup banyak untuk dapat menjelaskan variasi indeks batubara. SVR *complete feature* juga menghasilkan model dan parameter yang lebih baik dibandingkan dengan *Ridge Regression*. Metode yang terpilih adalah kernel-polynomial dengan *complete feature*. Hasil prediksi menunjukkan bahwa dengan model kernel-polynomial telah memprediksi indeks batubara di bulan Juli 2019 dengan baik, terlihat secara visual bahwa secara menyeluruh telah memprediksi dengan baik, walaupun terdapat beberapa yang berbeda jauh antara nilai prediksi dan nilai aktual. Hasil prediksi menggunakan SVR *complete feature* dengan kernel-polynomial mendapatkan nilai RMSE sebesar 0,619 dengan parameter sigma bernilai 0,100 dan nilai c sebesar 1. Sedangkan, nilai RMSE menggunakan metode lainnya menghasil-

kan RMSE yang lebih tinggi, yaitu metode *ridge regression* menghasilkan RMSE sebesar 2,088, metode SVR *complete feature* menggunakan kernel-linier sebesar 0,689, metode SVR *complete feature* menggunakan kernel-RBF, metode SVR *feature selection* menggunakan kernel-linier sebesar 0,890, kernel-polynomial sebesar 0,792, dan kernel-RBF sebesar 0,735. Sehingga, metode terbaik dalam memprediksi indeks batubara PT XYZ adalah SVR *complete feature* dengan kernel-polynomial.

5.2 Saran

Berdasarkan kesimpulan yang diperoleh, dapat dirumuskan saran sebagai pertimbangan penelitian selanjutnya adalah sebagai berikut.

1. Untuk meningkatkan indeks batubara PT XYZ, perlu dilakukan peningkatan *speed kiln*, *ILC Exit temperature*, *SLC Exit Temperature*, *Temperature Stage 4 SLC*, *ILC Coal*, *SLC Coal*, dan *Kiln Coal*.
2. Pada pola *time series* plot pada akhir bulan Januari indeks batubara di PT XYZ mengalami peningkatan yang sangat tinggi yaitu sebesar 0,5, hal ini disebabkan adanya *trouble* pada alat perekap data, sehingga PT XYZ disarankan untuk lebih mengontrol kinerja dari alat perekap data untuk menghindari adanya *missing* dalam pencatatan observasi.
3. Dapat ditambahkan variabel yang diduga berpengaruh terhadap indeks batubara PT XYZ supaya informasi yang didapatkan lebih lengkap, misalnya faktor cuaca yang mempengaruhi kadar air dalam batubara.

DAFTAR PUSTAKA

- Casasent D, & Nakariyakul S. (2004, Oktober 26). *Hyperspectral ratio feature selection: agricultural*. Nondestructive Sensing for Food Safety, Quality, and Natural Resources 5587, *Proceeding of SPIE*, 133-143.
- CNBC Indonesia. (2019). Dipetik 10 3, 2019, dari <https://www.cnbcindonesia.com/news/20-190624120905-4-80160/cadangan-batu-bara-ri-naik-jadi-3989-miliar-ton>
- Draper, N. R., & Smith, H. (1981). *Applied Regression Analysis*. New York: Jhon Wiley & Sons.
- Gujarati, D. (2003). *Ekonometri Dasar*. Jakarta: Erlangga.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. USA: Elsevier Inc.
- Hidayat, A. (2009). *Metode Penelitian Keperawatan dan Teknik. Analisis Data*. Jakarta: Salemba Medika.
- Hoerl, A. E., & Kennard, R. W. (1970). *Ridge Regression: Biased Estimation for Nonorthogonal Problems*. *Technometrics*, 12, 55-67.
- Huber, P. J. (1964). *Robust estimation of a location parameter*. *Ann. Math. Statist.* 35, 73âˆA ,S101.
- Montgomery, D. C. (2012). *Intoduction Statistical Quality Control*. 3 ed. New. York: Jhon Willey and Son.
- Narayana, B. L., & Kumar, S. P. (2012). *Mining Association Rules in Student's Assessment Data*. Vol. 9, Gurgaon. Indiam
- Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Scholkopf, B., & Smola, A. (2002). *Support Vector Machines and Kernel Algorithms*. Articles authored or co-authored.
- Sudjana. (1996). *Teknik Analisis Regresi Dan Korelasi*. Bandung: Tarsito.

Svetnik. (2003). *Random forest: A classification and regression tool for compound classification*. *Journal of Chemical Information and Computer*, 43:1947–1958.

LAMPIRAN

Lampiran 1. Data Indeks Batubara Beserta Variabel Prediktor

Y	X ₁	X ₂	X ₃	X ₄	X ₅
0,183762	62,62772	9275,004	10327,71	20504,66	20751,04
0,146457	63,73441	9454,278	10298,84	20535,38	20897,3
0,169449	63,57896	9545,356	10357,61	20490,08	20803,09
0,163915	63,57186	9488,896	10480,37	20449,79	20816,18
0,146837	63,56538	9324,046	10393,99	20545,32	20791,91
0,145472	61,77358	9370,227	10263,59	20539,97	20730,93
0,144897	62,10265	9451,028	10193,09	20545,81	20780,4
0,140738	62,43691	9420,191	10089,77	20557,42	20673,96
0,155594	62,60358	9621,754	10090,48	20588,16	20729,37
0,142889	62,60618	9620,991	10165,36	20556,92	20750,31
0,133568	62,61092	9453,061	10228,83	20535,62	20819,04
0,126633	62,61274	9334,836	10125,15	20518,31	20862,46
0,136144	62,60804	9253,543	10229,61	20519,96	20821,6
0,12125	52,31956	10056,37	10240,39	20531,6	20995,97
0,131053	63,26339	9636,668	9987,025	20545,9	20794,79
0,123343	63,47523	9680,034	9794,116	20607,77	20878,36
0,158478	64,04567	9653,7	9786,4	20579,23	20781,89
0,133868	61,65748	9661,759	9972,701	20582,42	20906
0,132983	61,34757	9152,007	10567,25	20530,76	20854,36
0,134934	61,04657	9250,261	10410,62	20525,19	20688,3
0,129508	60,63665	9275,333	10202,51	20586,2	20883,41
0,1194	63,01744	9097,076	10520,43	20538,43	20924,42
0,14344	63,18211	9113,422	10663,72	20487,26	20927,25
0,147484	61,71533	9354,12	10368,21	20526,51	20920,02
0,138271	53,7055	10108,49	9907,245	20657,34	20865,91
0,137074	53,09678	9948,384	9910,719	20510,49	21069,87

0,136328	35,97214	9633,537	9306,232	18790,64	18332,85
0,49929	3,35994	7087,914	3383,978	13970,75	10168,9
0,189498	49,43503	10228,13	10410,9	20380,97	21015,94
0,19033	45,61764	10264,97	10079,3	20228,17	21081,57
0,184685	48,61465	10065,05	9706,446	20233,93	20978,33
0,246036	18,3467	10197,51	9131,913	18395,18	14379,95
0,232872	48,78863	10747,56	9412,994	20408,63	20847,99
0,219871	58,58589	9985,268	9656,949	20400,42	20896,67
0,197619	60,92061	9820,087	9643,307	20440,11	20955,98
0,183121	64,71694	9731,621	9604,886	20433	21040,72
0,181596	65,97537	9756,769	9697,111	20350,63	20989,12
0,183197	65,28145	9659,41	9612,675	20354,88	21052,16
0,184747	66,78194	9951,486	9462,349	20447,93	21102,4
0,206046	64,27445	9871,516	9843,858	20447,36	21061,34
0,215382	64,11074	9648,76	9906,537	20407,35	21063,35
0,219036	64,11546	9531,306	10045,23	20417,4	21034,47
0,216908	64,00639	9619,927	9944,411	20430,77	21034,16
0,306244	25,94823	10258,2	9990,596	19018,62	17068,23
0,215152	45,32125	10698,38	10094,65	20489,55	19135,67
0,188841	58,77276	10047,8	10154,77	20298,96	21008,06
0,194003	60,7486	10222,07	10318,52	20314,65	21065,74
0,18168	59,59943	9767,996	10209,18	20326,99	20917,91
0,198021	59,68662	9487,999	10118,31	20326,42	21009,05
0,228498	61,70646	9771,562	9974,617	20358,6	20980,73
0,231746	48,10947	9945,666	10173,63	20382,93	19385,77
0,187005	44,13882	10500,47	10083,89	20205,54	20638,72
0,168402	63,14335	9985,672	9331,487	20441,04	20882,37
0,187095	63,94557	9549,858	9579,245	20441,83	20936,39
0,204316	65,03683	9325,638	9961,588	20462,54	20932,33

0,187499	62,64914	9300,525	9973,418	20434	20980,44
0,186391	61,59784	9236,042	10055,7	20399,92	20992,36
0,166247	62,18459	9341,645	9787,964	20394,13	21005,25
0,184072	62,79402	9410,354	9938,196	20395,33	21066,61
0,198011	62,79763	9318,816	10304,56	20354,99	21160,81
0,190288	60,54755	9749,965	10003,19	20365,53	20958,4
0,18742	52,64998	10058,37	9858,606	20394,9	19800,74
0,154571	61,81029	9649,61	9879,95	20352,83	21133,53
0,151716	61,80212	9664,402	9807,078	20351,48	21074,82
0,149438	61,12385	9621,334	9804,272	20352,96	21006,52
0,193956	41,88755	10010,19	9630,752	20358,2	18807,02
0,160177	57,41859	9767,098	9761,212	20426,69	20945,17
0,16001	59,95294	9718,375	9785,669	20378,75	20928,34
0,18939	33,21996	9940,859	9230,888	18758,7	16430,49
0,26707	19,86388	10671,62	9059,408	19256,04	15677,03
0,193148	43,94468	9987,902	10198,37	20463,57	20826,53
0,181226	43,96062	10099,2	9989,369	20347,99	20811,3
0,179509	43,94947	10040,02	10024,47	20352,58	20864,31
0,172592	43,95092	10011,16	10049,58	20332,62	20914,84
0,193039	15,06368	9595,831	8425,51	15673,84	13083,9
0,239695	36,2385	10487,83	7448,516	20493,72	15217,9
0,190147	61,64356	9468,318	9843,331	20359,07	21086,85
0,215918	63,05071	9897,446	9765,249	20398,8	21106,95
0,234861	50,96603	10159,58	9737,142	20502,97	20195,96
0,227336	58,09086	9796,432	9843,876	20480,73	21185,68
0,214852	54,29059	9823,567	9859,929	20489,69	20752,13
0,193693	59,55846	9372,558	10020,72	20375,68	20999,37
0,210929	59,88603	9414,514	9863,431	20404,49	20980,87
0,249447	34,74586	9963,757	9582,189	19566,57	17860,01

0,209014	54,06924	9696,14	9924,511	20341,63	20937,45
0,20187	58,09669	9350,122	10009,26	20400,72	21064,36
0,19033	60,14183	9388,746	9818,64	20435,55	21140,7
0,167747	62,19277	9434,636	9703,935	20416,84	21172,54
0,205651	63,1384	9542,991	9602,223	20430,66	21083,9
0,215439	63,13849	9453,538	9626,39	20429,37	21025,59
0,205815	60,60124	9562,282	9597,849	20430,84	21075,34
0,212799	53,51598	9854,197	9584,746	20472,31	19999,88
0,195561	60,21534	9210,569	9178,729	20373,92	21119,98
0,220736	62,02424	9235,793	9192,706	20374,66	20950,96
0,232543	61,75409	9393,325	9193,832	20384,73	21037,01
0,254786	61,85436	9333,957	9422,38	20392,32	20966,09
0,250388	63,06391	9250,635	9607,087	20399,86	21013,01
0,240683	63,9226	9210,571	9568,034	20397,74	20991,34
0,248283	63,4463	9319,24	9472,072	20416,41	20993,89
0,240152	62,85219	9315,111	9587,556	20354,36	20951,46
0,233662	60,08381	9613,331	9261,197	20364,04	20965,22
0,247599	59,37334	9684,912	9550,478	20332,96	20970,8
0,246852	60,0176	9551,164	9517,829	20324,89	20900,56
0,249217	60,01045	9348,207	9421,067	20351,26	20912,69
0,242306	60,01421	9299,741	9459,136	20350	20967,59
0,246497	25,18393	9379,065	8139,433	16297,47	15760,12
0,277506	39,93552	9699,307	9050,84	19499,63	18921,41
0,218235	60,7903	9692,839	9800,537	20471,48	21055,89
0,207869	60,86725	9560,799	9697,884	20471,89	21045,25
0,210883	60,51732	9531,85	9797,967	20443,45	21126,89
0,205541	60,83951	9670,516	9928,246	20399,64	21237,59
0,209258	59,0931	9858,58	9887,56	20451,31	21294,05
0,203417	56,68374	9863,362	9863,52	20411,61	21128,59

0,202481	58,02642	9566,751	9701,593	20399,78	21020,6
0,221244	59,62532	9378,993	9693,331	20383,86	21016,57
0,215123	59,6418	9314,026	9782,498	20375,56	21111,61
0,192469	59,1145	9355,325	9802,751	20375,88	21135,46
0,208914	58,70409	9499,567	9864,349	20374,06	21089,62
0,196768	59,46697	9474,2	9817,123	20418,74	21061,22
0,209516	60,8708	9500,562	9790,845	20409,81	21024,72
0,196278	60,88452	9536,138	9789,827	20409,18	21094,74
0,203739	61,15082	9617,385	9859,376	20417,28	21170,04
0,192551	61,14372	9451,731	9848,484	20362,93	21084,29
0,192689	61,02983	9416,999	9720,176	20342,95	20985,36
0,183939	61,728	9427,569	9685,314	20404,22	21081,29
0,180942	60,85324	9507,584	9735,22	20377,26	21089,79
0,206157	60,85364	9637,464	9850,814	20352,72	21144,85
0,22981	60,49409	9545,682	9857,19	20349,51	21149,68
0,234774	59,70291	9388,561	10012,69	20367,05	21096,95
0,22011	59,68203	9598,937	9681,842	20400,33	20974,62
0,195111	59,68288	9689,767	9664,736	20351,36	20967,03
0,193021	59,68159	9735,297	9619,74	20303,53	20952,17
0,207875	58,81671	9847,701	9586,27	20305,4	20955,91
0,210731	57,74885	9943,981	9454,369	20353,61	20954,72
0,170325	59,00274	10022,59	9434,674	20351,49	20952,75
0,172489	59,18661	9892,065	9551,888	20326,86	20974,76
0,167322	59,19458	9883,305	9632,208	20303,52	21000,3
0,173953	59,1959	9964,334	9366,825	20303,17	20998,59
0,206044	14,413	4915,306	4664,818	9255,398	8670,135
0,286979	9,386857	8823,63	7222,812	16001	12650,4
0,226549	27,73483	9119,084	7692,668	17158,84	14568,6
0,193331	46,2389	9750,034	9642,176	19911,7	19712,48

0,179172	63,80361	9449,336	9795,724	20376,66	21048,8
0,178605	62,56254	9462,599	9775,009	20360,99	21049,07
0,174261	63,31465	9431,335	9798,449	20375,36	21044,6
0,208523	63,32477	9683,754	9815,862	20399,52	21001,25
0,209607	63,32945	9724,623	9757,549	20423,48	20924,42
0,204396	64,31698	9751,441	9753,878	20417,73	20893,27
0,185021	63,94682	9620,716	9697,094	20423,47	20953,59
0,20653	63,3375	9630,951	9948,735	20419,7	20991,85
0,211483	64,1448	9712,448	9865,656	20387,13	20963,49
0,201402	64,3583	9664,953	9737,358	20432,01	20903,74
0,175438	63,45194	9522,218	9873,876	20397,28	20989,83
0,202813	62,62954	9824,334	10042,63	20396,68	20920,8
0,251471	47,13808	10208,64	10114,39	20249,55	19539,5
0,231673	62,66909	9522,834	9951,353	20396,13	20921,14
0,216076	61,20525	9768,448	10091,37	20365,3	20908,91
0,209365	64,20992	9328,504	9908,403	20359,49	20916,9
0,209701	64,58093	9500,928	9856,441	20417,16	20857,32
0,22363	64,57821	9536,41	9899,662	20481,14	20717,44
0,240098	64,58471	9743,034	9940,619	20463,35	20875,36
0,234312	64,57727	9596,567	10040,14	20459,63	20862,9
0,203649	63,31606	9595,876	10042,88	20445,4	20857,7
0,170769	29,24016	9484,933	8857,696	14979,67	15531,51
0,338421	16,70025	8486,824	6374,541	16666,35	14274,69
0,241239	61,39102	9374,74	10172,52	20402,7	20906,37
0,224098	61,88409	9774,186	10052,62	20475,29	20891,66
0,213905	61,13663	9765,299	9852,604	20492,78	20911,43
0,216505	61,63285	9783,151	10078,91	20469,31	20844,05
0,2256	60,03392	9712,849	10048,23	20476,17	20868,72
0,225059	60,37601	9476,755	10193,16	20447,79	20941,82

0,230236	59,35522	9318,575	9975,641	20373,72	20897,07
0,266563	21,12817	9109,863	6333,189	16694,57	15042,42
0,215168	56,20871	9580,032	9733,775	20413,34	20876,77
0,20438	58,20055	9425,175	10003,8	20399,89	20863,39
0,186871	58,1746	9572,43	9896,886	20366,68	20860,23
0,199438	60,00738	9738,05	9563,401	20376,8	20850,49

Lampiran 1. Data Indeks Batubara Beserta Variabel Prediktor (lanjutan)

X_6	X_7	X_8	X_9	X_{10}	X_{11}
105,4457	391,5511	500,5829	2319,734	54,83736	36,39813
49,28401	258,5572	481,863	2345,8	53,99356	38,12839
71,64092	341,6174	491,5511	2347,307	53,09644	38,4165
41,8851	305,1522	527,948	2344,673	54,83423	37,59907
50,37505	261,7176	490,1144	2339,535	53,65164	38,31932
34,07595	275,4914	473,8084	2344,012	53,55056	37,94104
41,58552	267,8391	466,6448	2330,604	53,72003	36,81523
39,93086	268,6149	459,1346	2344,089	53,67011	36,88886
79,05161	305,8865	461,2697	2342,297	54,3596	36,83955
51,26098	251,7136	472,6442	2359,958	54,58583	37,18299
39,36543	222,2303	464,0692	2349,802	54,56474	35,35948
55,20519	171,3933	463,1539	2302,978	54,12006	37,13785
58,00649	225,0198	461,4236	2278,766	53,78563	44,35259
35,66171	155,7713	393,6385	2269,655	52,98349	36,7314
33,07444	206,1495	460,8478	2311,864	53,47587	36,69169
28,94478	177,019	459,6466	2320,023	53,50075	37,74313
62,38404	307,21	480,1693	2344,323	52,78096	38,14389
47,64093	177,9783	452,9575	2326,693	54,84452	36,82816
63,17957	172,8399	453,1108	2326,693	54,84452	36,82816

54,20632	171,7428	461,8039	2338,66	54,24288	38,25392
36,29275	157,7473	468,063	2313,887	54,67788	36,89441
31,74181	149,269	445,7172	2327,629	53,84301	38,51235
82,1198	209,7101	469,7846	2318,663	53,60181	36,9853
68,18911	240,4316	471,9927	1730,282	40,37286	27,42741
31,98575	196,2168	436,1592	1648,663	37,40464	27,566
31,99376	152,7795	433,7088	2319,7	51,75	39,3
18,34356	82,95712	291,7329	1652,49	36,61	28,36
43,43655	33,73942	142,0702	3202,95	75,78	47,6
205,4149	124,3117	445,2934	2232,07	50,9	35,91
178,1739	127,0514	408,724	2349,16	53,9	37,32
200,3211	155,4673	418,7882	2339,24	52,76	37,2
143,5442	62,27578	254,6835	871,57	19,19	14,84
260,4559	345,7957	393,2537	2035,92	44,51	34,62
289,1558	315,8684	449,5857	2344,3	51,3	40,39
281,1881	265,1074	468,665	2324,34	50,96	40,32
273,9644	234,8766	490,7294	2318,64	53,13	41,68
274,7145	264,8777	474,5	2330,04	53,17	40,24
234,147	272,3954	519,5738	2329,55	51	37,29
172,8566	325,7331	534,2017	2327,82	52,42	38,25
227,1532	358,708	551,8743	2343,22	51,78	40,84
240,2706	398,5302	557,2247	2361,01	50,73	41,8
251,5584	398,9448	567,5766	2368,07	51,44	41,24
252,5619	391,7419	562,1586	2342,64	51,46	38,29
201,2418	147,8091	322,0695	1374,38	30,59	23,02
238,3235	114,804	476,6478	1758,02	39,13	29,16
243,839	183,9757	475,1085	2370,61	50,9	39,54
227,354	206,6032	500,8645	2361,24	50,91	37,78
223,895	131,4479	484,4083	2346,99	51,53	37,11

259,3288	180,2867	477,9659	2386,19	50,25	38,73
336,8515	319,6366	477,6459	2365,08	48,48	35,6
325,1888	208,0527	432,2524	2156,35	43,14	34,58
219,8876	150,5094	379,7618	2247,18	47,04	35,94
245,0701	209,3495	476,3362	971,93	21,55	15,72
286,3373	267,849	496,9223	2332,99	53,27	37,04
340,5995	302,2845	501,5645	2330,85	53,43	37,56
324,0852	220,8016	499,6357	2370,46	51,52	37,87
305,694	230,5264	491,4487	2353,68	50,55	37,08
224,0375	212,5932	498,9854	2377,31	50,57	38,29
234,9443	269,8771	531,6572	2372,93	51,02	38,62
240,4246	277,5693	569,8207	2331,45	51,74	39,08
210,3831	238,3387	552,3911	2342,39	51,46	37,44
253,6095	168,4645	453,1446	2318,73	53,54	36,19
217,492	156,8999	486,3855	2317,02	53,43	36,08
204,811	146,1751	492,3885	2323,68	53,45	36,41
195,5439	145,0091	490,4956	2330,25	53,07	37,51
212,5063	126,1663	421,8244	2122,62	49,32	34,61
252,9271	139,5329	475,4894	2332,86	53,62	38,3
245,814	148,9587	495,2671	2325,96	53,67	38,68
160,3809	84,89738	322,6677	874,84	19,81	14,21
199,9136	68,63703	281,6904	1155,6	26,52	18,73
308,1246	117,2967	387,4133	2304,84	52,96	38,75
245,422	121,3513	395,1814	2301,41	53,5	39,2
248,9147	112,4428	394,9886	2308,58	53,77	38,9
238,577	92,19079	396,3604	2314,31	53,65	39,58
93,63492	29,42683	138,9119	962,15	22,64	16,62
283,985	105,0515	438,0107	1756,19	39,77	29,07
290,3962	195,8479	534,0845	2363,51	51,83	39,47

278,5001	293,0503	575,0275	2351,09	52,96	38,89
277,407	291,6688	514,7807	2032,15	46,1	33,98
297,4221	262,806	626,3146	2312,81	53,61	39,97
264,111	205,6838	612,3222	2029,97	47,44	34,23
290,5021	259,8927	494,8507	2315	53,79	38,43
306,0531	319,7692	512,5323	2332,97	53,63	38,06
207,7471	198,7435	370,7466	1272,55	28,89	20,7
259,9636	265,6524	487,5584	2153,67	48,25	36,51
323,2949	239,0747	475,6179	2354,7	52,11	40,49
284,3619	243,5149	503,2925	2338,52	53,83	39,18
240,2865	167,7558	508,3837	2329,51	54,16	39,33
256,3613	335,456	535,0402	2328,42	54,31	39,7
275,3284	365,8143	539,3741	2338,26	53,33	39,59
244,9165	326,6519	531,9521	2332,27	52,79	39,34
232,0385	267,1187	497,1131	2341,01	51,59	39,19
263,1683	279,2989	498,0853	2350,53	51,52	39,89
301,8561	406,4733	499,1986	2345,13	50,9	40,99
286,8028	460,5267	506,5901	2338,74	51,63	40,7
336,1564	535,35	503,7348	2332,89	50,03	41,71
342,5658	484,9547	510,801	2332,16	51,34	41,98
346,8449	465,6899	494,2647	2343,36	52,61	41,16
354,5183	488,2746	496,7706	2339,51	53,09	39,88
327,0574	453,4547	487,4082	2328,38	52,99	39,74
305,0814	430,9219	461,2484	2326,66	53,7	40,4
329,9373	446,226	467,9685	2321,06	53,67	40,73
318,1786	468,3915	497,3185	2322,61	54,4	39,94
325,3695	467,5953	506,6904	2342,55	51,97	41,93
294,2196	442,2135	516,961	2318,19	54,36	39,64
117,6016	165,1762	208,7033	1070,29	24,44	19,02

261,2734	387,7518	437,4489	1555,6	35,03	27,66
300,1045	313,2629	527,1453	2326,22	53,03	40,35
278,9756	295,4848	534,9751	2305,91	54,25	39,1
270,2438	315,1711	534,751	2311,91	53,67	39,49
242,2213	280,6312	545,0768	2313,29	53,56	39,01
214,6393	281,4462	532,5336	2317,77	53,21	38,91
222,8251	242,9642	519,9749	2305,89	54,45	38,19
261,4693	287,3378	520,3711	2314,45	55,14	38,05
288,279	376,969	521,5662	2320,37	54,28	39,74
288,5079	351,6024	519,0765	2336,21	54,07	40,25
258,003	268,7585	506,2682	2323,76	54,81	38,61
272,3155	334,1492	511,6129	2297,76	55,43	37,67
262,8905	276,243	520,9701	2301,78	55,7	38,16
285,9802	325,6145	525,1875	2316,98	55,66	38,77
257,7283	274,4582	515,1872	2332,05	54,76	38,49
282,8278	282,4443	516,628	2337,47	53,71	40,48
308,3612	228,2562	496,3304	2333,65	53,78	40,23
275,1888	281,9777	476,6954	2314,98	55,36	38,52
237,949	227,7368	520,6223	2350,43	53,58	38,89
237,0895	199,4748	504,2617	2310,16	53,19	39,21
271,7735	268,8544	501,8386	2311,75	53,63	39,67
306,8246	338,5118	516,7838	2312,85	52,25	40,19
330,7391	338,2463	517,3431	2277,71	53,34	40,58
308,8145	287,6757	515,6226	2260,81	52,13	42,2
306,1581	193,5606	486,3774	2290,38	54,11	40,11
323,7111	195,5848	456,5039	2318,77	54,3	39,94
346,2276	250,7509	453,8708	2316,83	54,13	39,79
318,2193	280,9874	468,6064	2291,92	54,9	38,54
266,1861	130,7971	463,6378	2280,03	57,18	36,7

265,8176	152,6627	453,406	2306,66	57,6	37,34
268,6012	127,242	449,8904	2314,73	56,97	36,9
265,8725	165,5405	447,6384	2305,9	57,05	36,89
70,60007	60,28466	110,6853	673,78	16,44	10,96
95,55847	44,1483	151,7101	387,37	9,04	6,55
204,9037	77,68374	298,1017	1153,59	27,55	18,8
283,0161	150,4752	381,801	1846,84	43,53	31,44
334,7445	199,1042	462,9733	2317,77	53,21	38,91
316,7763	209,384	462,7103	2317,23	54,34	40,04
317,3055	188,5064	462,7021	2315,54	55,05	39,83
374,0017	298,613	466,0995	2325,3	55,15	39,72
381,8611	302,973	472,27	2303,97	56,05	39,08
366,2876	282,6355	479,2417	2294,18	56,46	39,13
288,3983	252,0295	480,7954	2294,09	55,79	39,83
349,7889	295,9132	494,1216	2328,6	55,69	38,95
298,8045	338,3173	516,4622	2319,59	55,72	38,75
284,0079	324,6122	495,9353	2295,32	56,7	38,04
232,6514	221,3303	509,8242	2323,76	55,44	39,28
286,1456	303,9517	468,6835	2318,97	56,91	37,7
316,0407	298,7625	408,4098	2116,04	51,8	34,76
396,6331	401,1686	478,2269	2341,88	56,31	39,47
327,6218	265,1322	444,0067	2364,6	55,31	40,65
349,4811	299,9807	462,7568	2354,78	54,49	40,92
343,0321	331,8404	464,0817	2314,44	56,03	39,48
358,0434	348,5935	461,8112	2341,07	55,05	39,3
400,8814	407,4267	486,1896	2281,3	55,47	39,46
408,3262	381,1237	482,8132	2269,26	55,43	39,76
347,2597	281,2015	467,8421	2283,33	56,3	38,96
144,4864	40,00706	198,0567	1260,95	30,01	21,45

143,5629	126,2218	259,1007	481,7	11,38	8,1
376,07	399,8782	512,3858	2322,25	54,5	39,02
309,7298	353,1349	504,2895	2300,59	54,76	39,52
294,278	334,181	463,6123	2316,18	56,53	37,99
295,484	299,3716	468,3949	2321,12	55,89	38,28
326,6441	334,1793	475,0982	2350,63	57,05	37,51
322,002	352,2715	496,0764	2344,18	57,41	37,02
316,7997	332,0441	485,7608	1662,72	38,91	27,85
127,2826	95,96957	283,4626	786,07	17,59	12,11
363,9005	216,6029	461,4004	2332,79	54,99	37,58
358,1228	220,1576	461,1954	2331,95	55,99	38,73
284,5314	197,4032	457,274	2348,38	56,27	38,19
290,195	302,8355	453,1773	2317,07	55,88	38,24

Lampiran 1. Data Indeks Batubara Beserta Variabel Prediktor
(lanjutan)

X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}
2,7	30,01	30,16	0,15	37,14	4002,37
2,7	30,01	30,16	0,15	37,14	4002
2,7	30,01	30,16	0,15	37,14	4002
2,64	29,46	29,43	0,14	38,48	3961
2,81	31,4	31,03	0,15	34,76	4192
2,81	31,4	31,03	0,15	34,76	4192
2,81	31,4	31,03	0,15	34,76	4192
2,81	31,4	31,03	0,15	34,76	4192
3,9	30,35	30,74	0,15	35,01	4018
4,47	30,33	30,47	0,15	34,73	4058
4,47	30,33	30,47	0,15	34,73	4058
4,44	30,1	31,3	0,13	34,15	4141

4,44	30,1	31,3	0,13	34,15	4141
4,44	30,1	31,3	0,13	34,15	4141
3,04	31,49	31,38	0,14	34,09	4251
4,02	31,23	30,02	0,14	34,72	4091
4,02	31,23	30,02	0,14	34,72	4091
4,02	31,23	30,02	0,14	34,72	4091
3,54	31,34	31,23	0,13	33,89	4260
3,54	31,34	31,23	0,13	33,89	4260
3,54	31,34	31,23	0,13	33,89	4260
3,54	31,33699	31,22572	0,134262	33,89461	4260
3,49	29,88673	29,65774	0,128574	36,9626	3942
3,7	29,76025	28,49483	0,127224	38,04501	3928
3,7	29,76025	28,49483	0,127224	38,04501	3928
3,42	30,77428	30,91539	0,177941	34,88982	4162
3,99	30,74885	30,07949	0,16555	35,18341	4073
2,99	30,79	31,55	0,19	34,67	4230
2,99	30,79	31,55	0,19	34,67	4230
3	30,88	30,96	0,15	35,16	4127
3	30,88	30,96	0,15	35,16	4127
2,93	31,08169	30,74399	0,164274	35,24195	4146
2,99	30,90105	30,93761	0,14808	35,17012	4130
2,99	30,90105	30,93761	0,14808	35,17012	4130
2,99	30,90105	30,93761	0,14808	35,17012	4130
2,99	30,90105	30,93761	0,14808	35,17012	4130
2,99	30,90105	30,93761	0,14808	35,17012	4130
2,99	30,90105	30,93761	0,14808	35,17012	4130
2,99	30,90105	30,93761	0,14808	35,17012	4130
4,38	16,75	18,34	0,15	60,53	2121
3,31	27,32	28,58	0,15	40,79	3603
2,91	31,24	32,37	0,16	33,48	4152

3,14	28,96	30,16	0,15	37,74	3832
3,14	28,96	30,16	0,15	37,74	3832
2,01	30,22	30,53	0,14	37,23	4022
2,01	30,22	30,53	0,14	37,23	4022
2,01	30,22	30,53	0,14	37,23	4022
2,92	32,06	31,48	0,14	33,54	4247
2,91	31,7	31,82	0,15	33,56	4252
2,93	32,57	31	0,14	33,51	4240
2,9	30,9	32,58	0,17	33,62	4263
2,9	30,9	32,58	0,17	33,62	4263
2,9	30,9	32,58	0,17	33,62	4263
2,9	30,9	32,58	0,17	33,62	4263
2,51	30,89	29,14	0,17	37,46	4061
2,51	30,89	29,14	0,17	37,46	4061
2,51	30,89	29,14	0,17	37,46	4061
2,6	31,18	31,73	0,14	34,49	4256
2,6	31,18	31,73	0,14	34,49	4256
3,18	29,85	29,73	0,17	37,24	3990
3,26	31,01346	31,48588	0,14988	34,23686	4222
3,26	31,01346	31,48588	0,14988	34,23686	4222
3,26	31,01346	31,48588	0,14988	34,23686	4222
3,26	31,01346	31,48588	0,14988	34,23686	4222
2,82	38,27254	24,29758	0,1482	34,61394	4188
3,08	41,43445	22,35307	0,183432	33,13235	4234
3,08	41,43445	22,35307	0,183432	33,13235	4234
3,8	39,21	22,84	0,17	34,16	4081
3,8	39,21	22,84	0,17	34,16	4081
3,8	39,21	22,84	0,17	34,16	4081
3,8	39,21	22,84	0,17	34,16	4081

3,78	26,14	34,81	0,18	35,28	4066
3,78	26,14	34,81	0,18	35,28	4066
4,12	38,44	21,97	0,17	35,47	4040
4,12	38,44	21,97	0,17	35,47	4040
3,26	37,96	22,69	0,19	36,09	4156
3,11	39,4	21,73	0,18	35,75	4053
3,11	39,4	21,73	0,18	35,75	4053
3	31,82	31,48	0,11	33,7	4285
3	31,82	31,48	0,11	33,7	4285
3	31,82	31,48	0,11	33,7	4285
3	31,82	31,48	0,11	33,7	4285
2,4	30,48	31,05	0,16	36,07	4192
2,4	30,48	31,05	0,16	36,07	4192
2,4	30,48082	31,0541	0,162382	36,06577	4192
2,4	30,48082	31,0541	0,162382	36,06577	4192
2,73	32,42571	31,38576	0,1523	33,45488	4268
2,73	32,42571	31,38576	0,1523	33,45488	4268
2,73	32,42571	31,38576	0,1523	33,45488	4268
2,73	32,42571	31,38576	0,1523	33,45488	4268
2,73	32,42571	31,38576	0,1523	33,45488	4268
2,73	32,42571	31,38576	0,1523	33,45488	4268
3,65	29,07	30,72	0,14	36,55	3957
3,65	29,07	30,72	0,14	36,55	3957
3,65	29,07	30,72	0,14	36,55	3957
2,73	31,86	30,68	0,17	34,74	4212
2,87	29,59	31,11	0,12	36,43	4049
2,91	28,98	31,23	0,1	36,88	4005
2,8	30,82	30,88	0,15	35,51	4137
2,8	30,82	30,88	0,15	35,51	4137

2,8	30,82	30,88	0,15	35,51	4137
2,8	30,82	30,88	0,15	35,51	4137
2,94	31,17	31,14	0,15	34,76	4185
2,94	31,17	31,14	0,15	34,76	4185
3,59	30,17	31,2	0,1	35,05	4059
3,59	30,17	31,2	0,1	35,05	4059
4,04	29,58	29,6	0,21	36,78	3936
3,09	30,54	29,64	0,19	36,73	3994
3,09	30,54	29,64	0,19	36,73	3994
3,04	31,71	30,89	0,18	34,35	4231
3,04	31,71	30,89	0,18	34,35	4231
3,04	31,71	30,89	0,18	34,35	4231
3,04	31,71	30,89	0,18	34,35	4231
3,59	33,07	30,63	0,26	32,71	4254
3,45	30,82842	30,69376	0,203644	35,02	4092
3,45	30,82842	30,69376	0,203644	35,02	4092
4,05	31,03096	30,41228	0,199179	34,5	4100
4,05	31,03096	30,41228	0,199179	34,5	4100
4,05	31,03096	30,41228	0,199179	34,5	4100
4,05	31,03096	30,41228	0,199179	34,5	4100
3,27	30,70038	31,34815	0,2199	34,68	4222
4,18	30,93731	31,0921	0,2199	33,79	4069
4,18	30,93731	31,0921	0,2199	33,79	4069
3,25	31,75	31,86	0,12	33,14	4195
3,25	31,75	31,86	0,12	33,14	4195
3,25	31,75	31,86	0,12	33,14	4195
3,25	31,75	31,86	0,12	33,14	4195
4,27	31,42	31,01	0,18	33,29	4105
4,27	31,42	31,01	0,18	33,29	4105

3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
3,24	31,01	31,99	0,15	33,76	4195
4,62	30,51	31,69	0,21	33,17	4090
4,62	30,51	31,69	0,21	33,17	4090
3,47	31,13607	33,17218	0,173075	32,22	4237
3,47	31,13607	33,17218	0,173075	32,22	4237
3,47	31,13607	33,17218	0,173075	32,22	4237
3,47	31,13607	33,17218	0,173075	32,22	4237
3,11	32,33744	31,57728	0,214727	32,97	4250
2,63	32,87929	33,13073	0,167376	31,36	4262
2,63	32,87929	33,13073	0,167376	31,36	4262
2,63	32,87929	33,13073	0,167376	31,36	4262
2,63	32,87929	33,13073	0,167376	31,36	4262
2,63	32,87929	33,13073	0,167376	31,36	4262

Definisi Variabel

Y	Indeks Batubara
X₁	Speed Kiln
X₂	ILC Exit Temperature
X₃	SLC Exit Temperature
X₄	Temperature Stage 4 ILC
X₅	Temperature Stage 4 SLC
X₆	ILC Coal
X₇	SLC Coal
X₈	Kiln Coal

X₉	LSF
X₁₀	SIM
X₁₁	ALM
X₁₂	AC
X₁₃	VM
X₁₄	FC
X₁₅	TS
X₁₆	TM
X₁₇	GHV

Lampiran 2. *Syntax* Matlab Menghitung MSE dan SSE Regresi Ridge

```

>> clear all
>> format long
>> x=xlsread('D:TA_X1')
>> y=xlsread('D:TA_Y')
>> X=input('masukan matriks xx : ')
>> XX=X'*X
>> XY=X'*y
>> X=input('masukan matriks xx : ')
>> Y=input('masukan matriks xy : ')
>> disp('          nilai_c          betaR
VIF');
>> for c=0:0.001:0.009
C=c*eye(18,18);
bR=((X+C)^(-1))*X*((X+C)^(-1));
VIF=diag(bR);
VIFridge=transpose(VIF);
betaR=transpose(((X+C)^(-1))*Y);
nilai_c=[c c c c c c c c c c c c c c c c];
disp(transpose([nilai_c;betaR;VIFridge]));
end
>> for c=0.01:0.01:0.09
C=c*eye(18,18);
bR=((X+C)^(-1))*X*((X+C)^(-1));
VIF=diag(bR);
VIFridge=transpose(VIF);
betaR=transpose(((X+C)^(-1))*Y);
nilai_c=[c c c c c c c c c c c c c c c c];
disp(transpose([nilai_c;betaR;VIFridge]));
>> for c=0.1:0.1:1
C=c*eye(18,18);
bR=((X+C)^(-1))*X*((X+C)^(-1));
VIF=diag(bR);
VIFridge=transpose(VIF);
betaR=transpose(((X+C)^(-1))*Y);
nilai_c=[c c c c c c c c c c c c c c c c];
disp(transpose([nilai_c;betaR;VIFridge]));

```

```

end
>> disp('***Nilai c yang dipilih adalah nilai c yang semua
nilai VIF lebih kecil dari 10');
>> CRT=input('masukan nilai c yang semua VIFnya kurang dari
10 : ');
>> XX=input('masukan matriks XX : ');
>> XY=input('masukan matriks XY : ');
>> MSE_ols=input('masukan matriks MSE untuk hasil OLS : ');
>> BOLS=(XX^(-1))*XY;
>> e=eig(XX);
>> r=rank(XX);
>> p=r-1;
>> disp('----->Ridge Trace<-----')
>> bRT=[-0.07725948;937.185;0.675867;0.454152];
>> bRT=[-0.07726;-0.552735272;0.13383458;0.114180863;-
0.162305603;-
0.26859324;0.427955082;0.463370322;0.301139869;-
0.074132266;-0.099050177;0.107632283;-
0.027868537;0.014143408;0.066859024;-0.009763757;-
0.059912361;-0.040777006]
>> SSR_RT=bRT'*XY;
>> j = 0;
>> m = length(e);
>> for i = 1:m
eBaru = e(i)/((e(i)+CRT)^2);
jLama = j + eBaru;
j = jLama;
end
>> jumlah = j;
>> sisa=CRT^2*bOLS'*((XX+CRT*eye(18,18)).^(-1))*bOLS;
>> MSE_RT=jumlah+sisa;
>> format short eng
>> disp('          CRT          SSR_RT          MSE_RT');
>> disp([CRT SSR_RT MSE_RT]);
>> format long

```

Lampiran 3. Nilai VIF dengan Berbagai Nilai c

Nilai c	VIF β^*_1 (c)	VIF β^*_2 (c)	VIF β^*_3 (c)	VIF β^*_4 (c)	VIF β^*_5 (c)	VIF β^*_6 (c)
0	6,4422	2,3203	2,7316	10,213	10,275	2,7923
0,001	5,749	2,0516	2,6495	9,5	9,6934	2,5557
0,002	5,6485	2,0216	2,6233	9,1966	9,4096	2,5234
0,003	5,5514	1,9929	2,5979	8,9093	9,1394	2,4927
0,004	5,4576	1,9655	2,5732	8,6371	8,8817	2,4633
0,005	5,3668	1,9393	2,5492	8,3788	8,6358	2,435
0,006	5,279	1,9143	2,5259	8,1334	8,401	2,4079
0,007	5,1938	1,8902	2,5032	7,9	8,1766	2,3818

0,008	5,1113	1,8671	2,481	7,6779	7,9619	2,3566
0,009	5,0312	1,8449	2,4593	7,4662	7,7565	2,3323
0,01	4,9534	1,8236	2,4381	7,2643	7,5597	2,3087
0,02	4,2838	1,6459	2,2487	5,6673	5,9738	2,1074
0,03	3,7615	1,5132	2,0896	4,5901	4,8748	1,948
0,04	3,3414	1,4082	1,9521	3,8226	4,077	1,8153
0,05	2,9958	1,3219	1,8312	3,2524	3,4764	1,7012
0,06	2,7066	1,249	1,7236	2,8148	3,0108	1,6013
0,07	2,4611	1,1861	1,627	2,47	2,6412	1,5126
0,08	2,2504	1,1309	1,5398	2,1923	2,3418	1,433
0,09	2,0678	1,0819	1,4605	1,9646	2,0952	1,3611
0,1	1,9083	1,038	1,3882	1,7749	1,8892	1,2956
0,2	1,0074	0,7525	0,9078	0,8477	0,8768	0,8619
0,3	0,6365	0,5953	0,656	0,5245	0,5268	0,6327
0,4	0,4449	0,4914	0,5041	0,367	0,3594	0,4928
0,5	0,3319	0,4164	0,404	0,2763	0,2649	0,3997
0,6	0,2593	0,3593	0,3338	0,2183	0,2058	0,3338
0,7	0,2096	0,3143	0,2823	0,1786	0,166	0,2848
0,8	0,174	0,2779	0,243	0,15	0,1378	0,2472
0,9	0,1475	0,2479	0,2122	0,1286	0,1169	0,2175
1	0,1271	0,2227	0,1876	0,112	0,101	0,1935

Lampiran 3. Nilai VIF dengan Berbagai Nilai c (Lanjutan)

Nilai c	VIF β^*_7 (c)	VIF β^*_8 (c)	VIF β^*_9 (c)	VIF β^*_{10} (c)	VIF β^*_{11} (c)	VIF β^*_{12} (c)
0	2,8638	3,1194	21,027	9,6571	13,089	11560,4
0,001	2,7807	2,978	18,897	9,2688	12,101	3,1254
0,002	2,7409	2,9447	17,787	8,9128	11,527	2,9077
0,003	2,7031	2,9126	16,776	8,5834	11,001	2,7999
0,004	2,6671	2,8817	15,852	8,2778	10,516	2,7126
0,005	2,6327	2,8519	15,005	7,9935	10,069	2,6342
0,006	2,5999	2,823	14,227	7,7283	9,6554	2,5617
0,007	2,5684	2,795	13,511	7,4803	9,2717	2,494
0,008	2,5381	2,7679	12,851	7,2479	8,9149	2,4302
0,009	2,5091	2,7415	12,239	7,0296	8,5825	2,37
0,01	2,4811	2,7159	11,673	6,8242	8,2721	2,3131
0,02	2,2465	2,4907	7,7257	5,2815	6,027	1,8759

0,03	2,0667	2,3061	5,5555	4,3023	4,6964	1,5939
0,04	1,9201	2,1488	4,229	3,621	3,8227	1,399
0,05	1,7961	2,0119	3,3554	3,1175	3,2075	1,2572
0,06	1,6885	1,8909	2,7472	2,7292	2,752	1,1496
0,07	1,5935	1,7828	2,305	2,4204	2,4017	1,0653
0,08	1,5086	1,6854	1,9722	2,1689	2,1242	0,9972
0,09	1,432	1,5972	1,7147	1,9601	1,8994	0,941
0,1	1,3625	1,5168	1,5107	1,7841	1,7136	0,8936
0,2	0,9013	0,983	0,6454	0,8856	0,8169	0,6365
0,3	0,656	0,7015	0,3911	0,5537	0,5052	0,5138
0,4	0,506	0,5311	0,2737	0,3883	0,353	0,4332
0,5	0,4063	0,4188	0,2072	0,292	0,2653	0,3736
0,6	0,3359	0,3405	0,1649	0,2302	0,2092	0,3272
0,7	0,2841	0,2834	0,1358	0,1878	0,1709	0,2897
0,8	0,2446	0,2403	0,1148	0,1572	0,1432	0,2587
0,9	0,2135	0,2069	0,0989	0,1343	0,1226	0,2328
1	0,1887	0,1805	0,0866	0,1166	0,1066	0,2107

Lampiran 3. Nilai VIF dengan Berbagai Nilai c (Lanjutan)

Nilai c	VIF β^*_{13} (c)	VIF β^*_{14} (c)	VIF β^*_{15} (c)	VIF β^*_{16} (c)	VIF β^*_{17} (c)
0	89353,17	79331,7	1,6621	78295,06	10,8151
0,001	3,2174	3,3371	1,6093	5,9282	10,2925
0,002	2,1613	2,3795	1,5934	4,8644	9,8714
0,003	1,918	2,145	1,5782	4,5222	9,4773
0,004	1,7982	2,0216	1,5637	4,2969	9,108
0,005	1,7165	1,933	1,5498	4,1128	8,7614
0,006	1,6515	1,8604	1,5365	3,9505	8,4356
0,007	1,5958	1,797	1,5236	3,8027	8,1289
0,008	1,546	1,7398	1,5113	3,6662	7,8398
0,009	1,5007	1,6872	1,4994	3,539	7,567
0,01	1,4587	1,6385	1,4879	3,4199	7,3092
0,02	1,1522	1,2811	1,39	2,5338	5,3565
0,03	0,9609	1,0596	1,3132	1,9848	4,1318
0,04	0,8305	0,9099	1,2496	1,617	3,3096
0,05	0,7365	0,803	1,1949	1,3569	2,7289
0,06	0,666	0,7232	1,1469	1,1652	2,3023

0,07	0,6112	0,6617	1,1039	1,0194	1,9789
0,08	0,5675	0,6129	1,0649	0,9054	1,7273
0,09	0,5319	0,5732	1,0293	0,8143	1,5272
0,1	0,5022	0,5404	0,9964	0,7401	1,3653
0,2	0,3508	0,3754	0,7605	0,3998	0,6425
0,3	0,2865	0,3065	0,6134	0,2867	0,4201
0,4	0,2465	0,2638	0,5101	0,2293	0,3156
0,5	0,2175	0,2327	0,4332	0,1935	0,2549
0,6	0,1947	0,2084	0,3737	0,1684	0,2149
0,7	0,1761	0,1884	0,3265	0,1495	0,1862
0,8	0,1605	0,1716	0,2882	0,1345	0,1644
0,9	0,1471	0,1572	0,2566	0,1223	0,1472
1	0,1355	0,1447	0,2301	0,112	0,1332

Lampiran 4. Koefisien Regresi *Ridge* dengan Berbagai Nilai c

Nilai c	β^*_1	β^*_2	β^*_3	β^*_4	β^*_5	β^*_6
0	-0,4041	0,0791	0,0763	-0,0394	-0,1444	0,5619
0,001	-0,394	0,0846	0,0778	-0,0466	-0,1379	0,5657
0,002	-0,392	0,085	0,0769	-0,0473	-0,1371	0,5648
0,003	-0,39	0,0853	0,076	-0,0479	-0,1363	0,5639
0,004	-0,3881	0,0857	0,0751	-0,0485	-0,1355	0,563
0,005	-0,3862	0,086	0,0741	-0,0491	-0,1347	0,5622
0,006	-0,3843	0,0863	0,0732	-0,0496	-0,1339	0,5613
0,007	-0,3824	0,0866	0,0723	-0,0501	-0,1331	0,5604
0,008	-0,3806	0,0869	0,0714	-0,0506	-0,1324	0,5596
0,009	-0,3788	0,0872	0,0705	-0,0511	-0,1317	0,5587
0,01	-0,377	0,0875	0,0696	-0,0516	-0,131	0,5579
0,02	-0,3604	0,0897	0,0606	-0,0553	-0,1244	0,5498
0,03	-0,3456	0,0913	0,052	-0,0578	-0,1187	0,5424
0,04	-0,3323	0,0925	0,0438	-0,0597	-0,1136	0,5354
0,05	-0,3202	0,0933	0,036	-0,061	-0,1091	0,5288
0,06	-0,3093	0,0939	0,0287	-0,0619	-0,1051	0,5225
0,07	-0,2992	0,0942	0,0217	-0,0627	-0,1014	0,5166
0,08	-0,2899	0,0944	0,0152	-0,0632	-0,0981	0,5109
0,09	-0,2812	0,0944	0,009	-0,0635	-0,0951	0,5054
0,1	-0,2732	0,0943	0,0031	-0,0638	-0,0922	0,5002
0,2	-0,215	0,0899	-0,0408	-0,0639	-0,0728	0,4568
0,3	-0,1792	0,0834	-0,0672	-0,0629	-0,0617	0,4237

0,4	-0,1545	0,0768	-0,0839	-0,0618	-0,0546	0,3968
0,5	-0,1363	0,0706	-0,0945	-0,0607	-0,0496	0,3742
0,6	-0,1221	0,065	-0,1014	-0,0596	-0,0458	0,3546
0,7	-0,1109	0,06	-0,1058	-0,0585	-0,0428	0,3374
0,8	-0,1016	0,0555	-0,1085	-0,0574	-0,0404	0,3221
0,9	-0,0938	0,0516	-0,11	-0,0563	-0,0384	0,3083
1	-0,0872	0,048	-0,1106	-0,0552	-0,0366	0,2958

Lampiran 4. Koefisien Regresi *Ridge* dengan Berbagai Nilai c (Lanjutan)

Nilai c	β^*_7	β^*_8	β^*_9	β^*_{10}	β^*_{11}	β^*_{12}
0	0,5591	0,2251	-0,0211	-0,0521	0,0517	1,1002
0,001	0,5562	0,2211	-0,0321	-0,053	0,0588	-0,0293
0,002	0,5553	0,2206	-0,033	-0,0534	0,0596	-0,0315
0,003	0,5545	0,22	-0,0339	-0,0538	0,0603	-0,0322
0,004	0,5537	0,2195	-0,0347	-0,0542	0,061	-0,0326
0,005	0,5529	0,2189	-0,0354	-0,0546	0,0617	-0,0329
0,006	0,5521	0,2184	-0,0362	-0,055	0,0623	-0,0331
0,007	0,5512	0,2178	-0,0369	-0,0554	0,0629	-0,0332
0,008	0,5504	0,2173	-0,0375	-0,0558	0,0635	-0,0334
0,009	0,5496	0,2168	-0,0382	-0,0563	0,064	-0,0335
0,01	0,5489	0,2162	-0,0388	-0,0567	0,0645	-0,0336
0,02	0,5411	0,2113	-0,0436	-0,0606	0,0687	-0,0347
0,03	0,5336	0,2068	-0,0471	-0,0641	0,0715	-0,0358
0,04	0,5263	0,2028	-0,0497	-0,0671	0,0734	-0,0371
0,05	0,5194	0,1991	-0,0517	-0,0696	0,0748	-0,0383
0,06	0,5126	0,1957	-0,0534	-0,0717	0,0758	-0,0395
0,07	0,5062	0,1925	-0,0547	-0,0735	0,0765	-0,0407
0,08	0,4999	0,1896	-0,0557	-0,075	0,077	-0,0418
0,09	0,4938	0,1869	-0,0566	-0,0763	0,0774	-0,0429
0,1	0,488	0,1843	-0,0573	-0,0773	0,0776	-0,0439
0,2	0,4387	0,1648	-0,0594	-0,0807	0,0763	-0,0518
0,3	0,4013	0,1513	-0,0579	-0,0788	0,0735	-0,0563
0,4	0,3716	0,1407	-0,0554	-0,0755	0,0705	-0,0589
0,5	0,3472	0,1321	-0,0527	-0,072	0,0678	-0,0602
0,6	0,3265	0,1248	-0,0501	-0,0685	0,0652	-0,0607

0,7	0,3087	0,1184	-0,0477	-0,0652	0,0628	-0,0607
0,8	0,293	0,1127	-0,0454	-0,0622	0,0606	-0,0603
0,9	0,2792	0,1077	-0,0434	-0,0594	0,0586	-0,0596
1	0,2668	0,1031	-0,0415	-0,0569	0,0566	-0,0588

Lampiran 4. Koefisien Regresi *Ridge* dengan Berbagai Nilai c (Lanjutan)

Nilai c	β^*_7	β^*_8	β^*_9	β^*_{10}	β^*_{11}	β^*_{12}
0	0,5591	0,2251	-0,0211	-0,0521	0,0517	1,1002
0,001	0,5562	0,2211	-0,0321	-0,053	0,0588	-0,0293
0,002	0,5553	0,2206	-0,033	-0,0534	0,0596	-0,0315
0,003	0,5545	0,22	-0,0339	-0,0538	0,0603	-0,0322
0,004	0,5537	0,2195	-0,0347	-0,0542	0,061	-0,0326
0,005	0,5529	0,2189	-0,0354	-0,0546	0,0617	-0,0329
0,006	0,5521	0,2184	-0,0362	-0,055	0,0623	-0,0331
0,007	0,5512	0,2178	-0,0369	-0,0554	0,0629	-0,0332
0,008	0,5504	0,2173	-0,0375	-0,0558	0,0635	-0,0334
0,009	0,5496	0,2168	-0,0382	-0,0563	0,064	-0,0335
0,01	0,5489	0,2162	-0,0388	-0,0567	0,0645	-0,0336
0,02	0,5411	0,2113	-0,0436	-0,0606	0,0687	-0,0347
0,03	0,5336	0,2068	-0,0471	-0,0641	0,0715	-0,0358
0,04	0,5263	0,2028	-0,0497	-0,0671	0,0734	-0,0371
0,05	0,5194	0,1991	-0,0517	-0,0696	0,0748	-0,0383
0,06	0,5126	0,1957	-0,0534	-0,0717	0,0758	-0,0395
0,07	0,5062	0,1925	-0,0547	-0,0735	0,0765	-0,0407
0,08	0,4999	0,1896	-0,0557	-0,075	0,077	-0,0418
0,09	0,4938	0,1869	-0,0566	-0,0763	0,0774	-0,0429
0,1	0,488	0,1843	-0,0573	-0,0773	0,0776	-0,0439
0,2	0,4387	0,1648	-0,0594	-0,0807	0,0763	-0,0518
0,3	0,4013	0,1513	-0,0579	-0,0788	0,0735	-0,0563
0,4	0,3716	0,1407	-0,0554	-0,0755	0,0705	-0,0589
0,5	0,3472	0,1321	-0,0527	-0,072	0,0678	-0,0602
0,6	0,3265	0,1248	-0,0501	-0,0685	0,0652	-0,0607
0,7	0,3087	0,1184	-0,0477	-0,0652	0,0628	-0,0607
0,8	0,293	0,1127	-0,0454	-0,0622	0,0606	-0,0603
0,9	0,2792	0,1077	-0,0434	-0,0594	0,0586	-0,0596

1	0,2668	0,1031	-0,0415	-0,0569	0,0566	-0,0588
---	--------	--------	---------	---------	--------	---------

Lampiran 4. Koefisien Regresi *Ridge* dengan Berbagai Nilai *c* (Lanjutan)

Nilai <i>c</i>	β^*_{13}	β^*_{14}	β^*_{15}	β^*_{16}	β^*_{17}
0	3,1686	3,017	-0,0133	2,9	-0,0419
0,001	0,0276	0,0577	-0,0151	-0,0393	-0,0384
0,002	0,0212	0,0519	-0,015	-0,0445	-0,0378
0,003	0,0188	0,05	-0,0148	-0,046	-0,0372
0,004	0,0174	0,049	-0,0146	-0,0465	-0,0366
0,005	0,0165	0,0484	-0,0145	-0,0466	-0,0361
0,006	0,0158	0,048	-0,0143	-0,0466	-0,0355
0,007	0,0151	0,0477	-0,0142	-0,0465	-0,035
0,008	0,0146	0,0475	-0,014	-0,0464	-0,0346
0,009	0,0141	0,0474	-0,0138	-0,0462	-0,0341
0,01	0,0137	0,0473	-0,0137	-0,046	-0,0337
0,02	0,0105	0,0471	-0,0121	-0,0438	-0,0302
0,03	0,0081	0,0477	-0,0107	-0,0419	-0,0279
0,04	0,0061	0,0484	-0,0093	-0,0403	-0,0261
0,05	0,0044	0,0492	-0,0082	-0,0389	-0,0248
0,06	0,0028	0,05	-0,0071	-0,0376	-0,0238
0,07	0,0013	0,0508	-0,0061	-0,0365	-0,0229
0,08	0	0,0515	-0,0052	-0,0354	-0,0222
0,09	-0,0012	0,0522	-0,0043	-0,0344	-0,0216
0,1	-0,0024	0,0529	-0,0035	-0,0334	-0,0211
0,2	-0,0112	0,0575	0,0021	-0,0258	-0,0181
0,3	-0,0166	0,0596	0,0055	-0,0203	-0,0166
0,4	-0,0203	0,0603	0,0079	-0,0161	-0,0157
0,5	-0,0227	0,0603	0,0096	-0,013	-0,015
0,6	-0,0244	0,0597	0,0109	-0,0104	-0,0145
0,7	-0,0256	0,0589	0,0119	-0,0084	-0,014
0,8	-0,0263	0,0579	0,0126	-0,0067	-0,0136
0,9	-0,0268	0,0568	0,0132	-0,0054	-0,0133
1	-0,0271	0,0557	0,0136	-0,0042	-0,0129

Lampiran 5. Perhitungan Transformasi ke Data Awal

$$\begin{aligned}\beta_1 &= \left(\frac{S_y}{S_1} \right) \beta_1^* \\ &= \left(\frac{0,7857}{0,5123} \right) x (-0,3604) = -0,5527\end{aligned}$$

$$\begin{aligned}\beta_2 &= \left(\frac{S_y}{S_2} \right) \beta_2^* \\ &= \left(\frac{0,7857}{0,5266} \right) x (0,0897) = 0,1338\end{aligned}$$

$$\begin{aligned}\beta_3 &= \left(\frac{S_y}{S_3} \right) \beta_3^* \\ &= \left(\frac{0,7857}{0,4170} \right) x (0,0606) = 0,1142\end{aligned}$$

$$\begin{aligned}\beta_4 &= \left(\frac{S_y}{S_4} \right) \beta_4^* \\ &= \left(\frac{0,7857}{0,2677} \right) x (-0,0553) = -0,1623\end{aligned}$$

$$\begin{aligned}\beta_5 &= \left(\frac{S_y}{S_5} \right) \beta_5^* \\ &= \left(\frac{0,7857}{0,3639} \right) x (-0,1244) = -0,2686\end{aligned}$$

$$\begin{aligned}\beta_6 &= \left(\frac{S_y}{S_6} \right) \beta_6^* \\ &= \left(\frac{0,7857}{1,0094} \right) x (0,5498) = 0,4279\end{aligned}$$

$$\begin{aligned}\beta_7 &= \left(\frac{S_y}{S_7} \right) \beta_7^* \\ &= \left(\frac{0,7857}{0,9175} \right) x (0,5411) = 0,4634\end{aligned}$$

$$\begin{aligned}\beta_8 &= \left(\frac{S_y}{S_8} \right) \beta_8^* \\ &= \left(\frac{0,7857}{0,5513} \right) x (0,2113) = 0,3011\end{aligned}$$

$$\begin{aligned}\beta_9 &= \left(\frac{S_y}{S_9} \right) \beta_9^* \\ &= \left(\frac{0,7857}{0,4621} \right) x (-0,0436) = -0,0741\end{aligned}$$

$$\begin{aligned}\beta_{10} &= \left(\frac{S_y}{S_{10}} \right) \beta_{10}^* \\ &= \left(\frac{0,7857}{0,4807} \right) x (-0,0606) = -0,0990\end{aligned}$$

$$\begin{aligned}\beta_{11} &= \left(\frac{S_y}{S_{11}} \right) \beta_{11}^* \\ &= \left(\frac{0,7857}{0,5015} \right) x (0,0687) = 0,1076\end{aligned}$$

$$\begin{aligned}\beta_{12} &= \left(\frac{S_y}{S_{12}} \right) \beta_{12}^* \\ &= \left(\frac{0,7857}{0,9783} \right) x (-0,0347) = -0,0279\end{aligned}$$

$$\begin{aligned}\beta_{13} &= \left(\frac{S_y}{S_{13}} \right) \beta_{13}^* \\ &= \left(\frac{0,7857}{0,5833} \right) x (0,0105) = 0,0141\end{aligned}$$

$$\begin{aligned}\beta_{14} &= \left(\frac{S_y}{S_{14}} \right) \beta_{14}^* \\ &= \left(\frac{0,7857}{0,5535} \right) x (0,0471) = 0,0668\end{aligned}$$

$$\begin{aligned}\beta_{15} &= \left(\frac{S_y}{S_{15}} \right) \beta_{15}^* \\ &= \left(\frac{0,7857}{0,9737} \right) x (-0,0121) = -0,0098\end{aligned}$$

$$\begin{aligned}\beta_{16} &= \left(\frac{S_y}{S_{16}} \right) \beta_{16}^* \\ &= \left(\frac{0,7857}{0,5744} \right) x (-0,0438) = -0,0599\end{aligned}$$

$$\begin{aligned}\beta_{17} &= \left(\frac{S_y}{S_{17}} \right) \beta_{17}^* \\ &= \left(\frac{0,7857}{0,5819} \right) x (-0,0302) = -0,0408\end{aligned}$$

Lampiran 6. Syntax R Support Vector Regression (SVR)

```

library(caret)
library(miscTools)
library(kernlab)

set.seed(12345)
datafull=read.csv("D:/traData.csv", sep=";")

modellrb<-train(IndeksBB~., data=datafull,
               method="svmRadialSigma")
print(modellrb)
prediction1rb<-predict(modellrb, datafull)
pred1rb=data.frame(prediction1rb)
rmse1=RMSE(datafull$IndeksBB, prediction1rb); rmse1
mae1=MAE(datafull$IndeksBB, prediction1rb); mae1
y1rb=as.vector(datafull[,1])
resid1rb=y1rb-prediction1rb
rsq1rb=rSquared(y1rb, resid1rb); rsq1rb

modell1lin<-train(IndeksBB~., data=datafull,
                 method="svmLinear")
print(modell1lin)
prediction1lin<-predict(modell1lin, datafull)
pred1lin=data.frame(prediction1lin)
rmse1lin=RMSE(datafull$IndeksBB, prediction1lin); rmse1lin
mae1lin=MAE(datafull$IndeksBB, prediction1lin); mae1lin
y1lin=as.vector(datafull[,1])
resid1lin=y1lin-prediction1lin
rsq1lin=rSquared(y1lin, resid1lin); rsq1lin

modell1pol<-train(IndeksBB~., data=datafull,
                 method="svmPoly")
print(modell1pol)
prediction1pol<-predict(modell1pol, datafull)
pred1pol=data.frame(prediction1pol)
rmse1pol=RMSE(datafull$IndeksBB, prediction1pol); rmse1pol
mae1pol=MAE(datafull$IndeksBB, prediction1pol); mae1pol
y1pol=as.vector(datafull[,1])
resid1pol=y1pol-prediction1pol
rsq1pol=rSquared(y1pol, resid1pol); rsq1pol

```

Lampiran 7. Syntax R Recursive Feature Elimination (RFE)

```

library(caret)
library(randomForest)

datamodelclass=read.csv("D:/traData.csv",
                        sep=";", header=TRUE)

set.seed(12345)

```

```
control<-rfeControl(functions=rffuncs,method="cv",number=10)
results<-
rfe(datamodelclass[,2:18],datamodelclass[,1],sizes=c(1:18),
rfeControl=control)

print(results)
```

Lampiran 8. Prediksi SVR Menggunakan Metode Terbaik

```
library(caret)
library(miscTools)
library(e1071)
set.seed(12345)
train1=read.csv("D:/traData.csv",sep=";")
test1=read.csv("D:/tesData.csv",sep=";")

model1pol<-train(IndeksBB~,data=train1,
method="svmPoly",gamma=0.1,cost=1)
print(model1pol)
prediction1pol<-predict(model1pol,test1[-1])
pred1pol=data.frame(prediction1pol)
rmse1pol=RMSE(test1$IndeksBB,prediction1pol);rmse1pol
mae1pol=MAE(test1$IndeksBB,prediction1pol);mae1pol
y1pol=as.vector(test1[,1])
resid1pol=y1pol-prediction1pol
rsq1pol=rsquared(y1pol, resid1pol);rsq1pol
```

Lampiran 9. Rumus Statistika Deskriptif

A. Statistika Deskriptif

a. Rata-rata (*mean*)

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

dimana

x_i = data ke- i

n = jumlah unit penelitian

b. Varians

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

s = standar deviasi (akar dari varians)

B. Korelasi *Pearson*

$$r_{y_1, y_2} = \frac{\sum_{i=1}^n (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2)}{\sqrt{\sum_{i=1}^n (y_{1i} - \bar{y}_1)^2} \sqrt{\sum_{i=1}^n (y_{2i} - \bar{y}_2)^2}}$$

dimana

r_{y_1, y_2} = koefisien korelasi antara y_1 dan y_2

C. *Variance Inflation* Factor (VIF)

$$VIF_j = \frac{1}{1 - R_j^2}$$

dimana R_j^2 adalah koefisien determinasi pada regresi x_j dengan variabel prediktor lainnya.

Lampiran 10. Surat Permohonan Ijin Memperoleh Data



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
 FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
 DEPARTEMEN STATISTIKA

Kampus ITS Sukolilo - Surabaya 60111
 Telp. : 031-594 3352, 031-599 4251 Fax. : 031-592 2940 PABX: 1213, 1214
<http://www.statistics.its.ac.id> email : statistika@its.ac.id

Nomor : B/104152/IT2.V1.9.2 /TU.00.09/2019

20 Desember 2019

Perihal : Permohonan ijin memperoleh data untuk Tugas Akhir

Yth : Pimpinan Diklat
 PT. XYZ
 Desa Sumberarum, Kec. Kerek
 Tuban

Dalam rangka menyelesaikan studi di Departemen Statistika FMKSD – ITS, mahasiswa diwajibkan untuk melakukan Tugas Akhir. Sehubungan dengan hal tersebut, kami mohon bantuan kepada **Bapak/Ibu**, agar mahasiswa berikut:

Nama : RIZKY AMALIA PUTRI
 NRP : 1316100040 / 06211640000040
 Dosen pembimbing : Dra. Wiwiek Setya Winahju, MS
 Program Studi : Sarjana (S1)

diperkenankan memperoleh data, untuk keperluan pelaksanaan Tugas Akhir di Instansi yang Bapak/Ibu pimpin. Adapun judul Tugas Akhir mahasiswa tersebut adalah "**PREDIKSI INDEX BATUBARA DI PT. XYZ MENGGUNAKAN RIDGE REGRESSION DAN SUPPORT VECTOR REGRESSION (SVR)**". Pengambilan data tersebut dilaksanakan pada bulan 1 Agustus 2019 s.d 31 Agustus 2019.

Adapun data yang dibutuhkan adalah :

- Data index Batubara (bahan baku, kualitas dan operasional)

Demikian atas ijin yang diberikan kami sampaikan terima kasih.

Kepala Departemen,

Drs. Suhartono
 NIP. 19710929 199512 1 001

Lampiran 11. Surat Keterangan Permintaan Data

SURAT KETERANGAN

Saya yang bertanda tangan di bawah ini menerangkan bahwa :

1. Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

Nama : Rizky Amalia Putri

NRP : 0621164000040

Telah mengambil data di instansi/perusahaan kami :

Nama Instansi : PT XYZ

Divisi/ bagian : Production Planning and Eval

sejak tanggal 1 Agustus 2019 sampai dengan 31 Agustus 2019 untuk keperluan Tugas Akhir/ Thesis Semester Gasal/Genap* 2019 / 2020.

2. ~~Tidak Keberatan~~/Keberatan* nama perusahaan dicantumkan dalam Tugas Akhir/ Thesis mahasiswa Statistika yang akan di simpan di Perpustakaan ITS dan dibaca di lingkungan ITS.
3. Tidak Keberatan/~~Keberatan~~* bahwa hasil analisis data dari perusahaan dipublikasikan dalam E journal ITS yaitu Jurnal Sains dan Seni ITS.

Tubean, 26 Desember 2019
Pimpinan Perusahaan
Unit of Production
Planning & Eval



(Imron Gozali, ST.)

*(coret yang tidak perlu)

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Penulis dilahirkan di Gresik, 11 September 1997 dengan nama lengkap Rizky Amalia Putri namun biasa dipanggil Riris. Penulis menempuh pendidikan formal di SD Negeri Latsari Tuban, SMP Negeri 3 Tuban, dan SMA Negeri 1 Tuban. Kemudian penulis diterima sebagai mahasiswa Departemen Statistika ITS pada tahun 2016. Selama masa perkuliahan, penulis aktif di Himpunan Mahasiswa Statistika ITS (HIMASTA-ITS) sebagai staf KESMA (Kesejahteraan Mahasiswa) periode 2017-2018 dan Ketua Departemen KESMA (Kesejahteraan Mahasiswa) pada periode 2018-2019. Selain itu penulis juga aktif di forum daerah yaitu Forum Mahasiswa Ronggolawe Tuban sebagai staff KWU pada periode 2017-2018 dan Ketua Divisi KWU periode 2018-2019. Selain itu, penulis juga aktif dalam UKM UKTK-ITS di divisi *modern dance* sebagai staf Departemen Produksi 2017-2018 dan Ketua Departemen HUMAS 2018-2019. Penulis juga merupakan Pemandu Integralistik (Mentor) GERIGI ITS dan pernah bergabung dikepanitian ITS EXPO 2017. Selain aktif dalam organisasi, penulis juga pernah magang di BMKG Stasiun Klimatologi Malang dan PT Semen Indonesia (Persero) Tbk pada tahun 2019. Bagi pembaca yang ingin berdiskusi, memberikan saran, dan kritik mengenai Tugas Akhir ini dapat disampaikan melalui riskyfemale@gmail.com atau melalui nomor 081333303785.