



TUGAS AKHIR - KS184822

**PREDIKSI KUAT TEKAN SEMEN UNTUK PRODUK
PORTLAND COMPOSITE CEMENT (PCC) di PT.
SEMEN INDONESIA (Persero) Tbk. MENGGUNAKAN
SUPPORT VECTOR REGRESSION (SVR) DENGAN
*FEATURE SELECTION***

**RIZKI FEBRIASTO
NRP 062117 4500 0020**

**Dosen Pembimbing
Ni Luh Putu Satyaning P. Paramita, S.Si, M.Sc.
Wibawati, S.Si, M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



TUGAS AKHIR - KS184822

**PREDIKSI KUAT TEKAN SEMEN UNTUK PRODUK
PORTLAND COMPOSITE CEMENT (PCC) di PT.
SEMEN INDONESIA (Persero) Tbk. MENGGUNAKAN
SUPPORT VECTOR REGRESSION (SVR) DENGAN
*FEATURE SELECTION***

**RIZKI FEBRIASTO
NRP 062117 4500 0020**

**Dosen Pembimbing
Ni Luh Putu Satyaning P. Paramita, S.Si, M.Sc.
Wibawati, S.Si, M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



FINAL PROJECT - KS184822

**PREDICTION OF CEMENT COMPRESSIVE STRENGTH
FOR *PORTLAND COMPOSITE CEMENT* (PCC)
PRODUCT at SEMEN INDONESIA (Persero) Tbk.
USING *SUPPORT VECTOR REGRESSION* (SVR) WITH
*FEATURES SELECTION***

**RIZKI FEBRIASTO
SN 062117 4500 0020**

Supervisors

**Ni Luh Putu Satyaning P. Paramita, S.Si, M.Sc.
Wibawati, S.Si, M.Si.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**

LEMBAR PENGESAHAN

**PREDIKSI KUAT TEKAN SEMEN UNTUK PRODUK
PORTLAND COMPOSITE CEMENT (PCC) di PT. SEMEN
INDONESIA (Persero) Tbk. MENGGUNAKAN SUPPORT
VECTOR REGRESSION (SVR) DENGAN FEATURE
SELECTION**

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

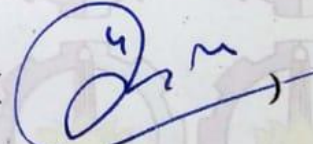
Oleh :

RIZKI FEBRIASTO
NRP. 062117 4500 0020

Disetujui oleh Pembimbing:
Ni Luh Putu Satyaning P.P, S.Si, M.Sc
NIP. 1991201712036

()

Wibawati, S.Si, M.Si
NIP. 19741213 199802 2 001

()

Mengetahui,
Kepala Departemen



Dr. Suhartono *st*

NIP. 19710929 199512 1 001
SURABAYA, JULI 2019

**PREDIKSI KUAT TEKAN SEMEN UNTUK PRODUK
PORTLAND COMPOSITE CEMENT (PCC) di PT. SEMEN
INDONESIA (Persero) Tbk. MENGGUNAKAN SUPPORT
VECTOR REGRESSION (SVR) DENGAN FEATURE
SELECTION**

Nama Mahasiswa : Rizki Febriasto
NRP : 062117 4500 0020
Departemen : Statistika-FMKSD ITS
**Dosen Pembimbing : N.L.P. Satyaning P.P, S.Si, M.Sc.
Wibawati, S.Si, M.Si.**

Abstrak

Indonesia sebagai negara berkembang terus-menerus melakukan pembangunan dalam segala bidang aspek kehidupan. Salah satu aspek yang selalu berkembang adalah sektor infrastruktur. Dengan menghadapi fenomena pertumbuhan infrastruktur ini, khususnya di sektor pembangunan perlu adanya hal yang menunjang dalam setiap pembangunan yang terjadi, salah satunya adalah material semen. Salah satu perusahaan yang bergerak di bidang produksi semen adalah PT. Semen Indonesia (Persero) Tbk. Terdapat beberapa produk yang dihasilkan oleh PT. Semen Indonesia (Persero) Tbk., salah satunya adalah Portland Composite Cement (PCC). Dalam penelitian ini akan dilakukan prediksi terhadap kuat tekan semen PCC di hari ke-28. Data yang digunakan yaitu bulan Juli 2018 hingga Maret 2019. Pada hasil analisis, hasil regresi linier telah terindikasi mengalami multikolinearitas, sehingga ditangani menggunakan PCR. Feature selection yang digunakan adalah Recursive Feature Elimination (RFE) berbasis Random Forest (RF-RFE) yang telah menghilangkan lima variabel. Dengan teknik machine learning yaitu SVR didapatkan model terbaik yaitu SVR dengan fungsi kernel Radial Basis Function (RBF) dengan complete feature. Feature selection tidak menghasilkan hasil yang lebih baik dibanding complete feature. Model yang terbaik digunakan untuk memprediksi kuat tekan pada bulan April 2019 yang menghasilkan kriteria terbaik RMSE sebesar 8,78.

Kata Kunci : Kuat Tekan, Multikolinearitas, PCC, PCR, RBF, RFE, RMSE, SVR.

**PREDICTION OF CEMENT COMPRESSIVE STRENGTH
FOR PORTLAND COMPOSITE CEMENT (PCC) PRODUCT
at PT. SEMEN INDONESIA (Persero) Tbk. USING
SUPPORT VECTOR REGRESSION (SVR) WITH
FEATURES SELECTION**

Name : Rizki Febriasto
SN : 062117 4500 0020
Department : Statistics-FMKSD ITS
Supervisors : N.L.P. Satyaning P.P, S.Si, M.Sc.
Wibawati, S.Si, M.Si.

Abstract

Indonesia as a developing country has continued to carry out development in all sectors of life. One of the sectors that are always developing is the infrastructure sector. By dealing with the phenomenon of infrastructure growth, especially in the development sector, it is necessary to have things to support in every development matter, one of them is cement material. One of the companies which engaged in cement production is PT. Semen Indonesia (Persero) Tbk. There are several products manufactured by PT. Semen Indonesia (Persero) Tbk. One of them is Portland Composite Cement (PCC). In this observation, predictions will be made on the compressive strength of PCC cement on the 28th day. The data used in this observation is in July 2018 until March 2019. The result of linear regression analysis is, it has been indicated to multicollinearity, then it is solved by PCR. Feature selection based on Random Forest (RF-RFE) is applied and has eliminated five variables. By machine learning techniques namely SVR, the best model is SVR with kernel Radial Basis Function (RBF) with complete feature. The best model is used to predict compressive strength in April 2019 which has the best RMSE criteria of 8.78.

Keywords : Compressive Strength, Multicollinearity, PCC, PCR, RBF, RFE, RMSE, SVR.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat, taufiq, serta hidayah-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “**PREDIKSI KUAT TEKAN SEMEN UNTUK PRODUK *PORTLAND COMPOSITE CEMENT (PCC)* di PT. SEMEN INDONESIA (Persero) Tbk. MENGGUNAKAN *SUPPORT VECTOR REGRESSION (SVR)* DENGAN *FEATURE SELECTION*”.**

Penyusunan Tugas Akhir ini dapat terselesaikan dengan baik dan lancar karena tidak lepas dari dukungan berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Ibu Ni Luh Putu Satyaning P.P dan ibu Wibawati selaku dosen pembimbing serta seluruh dosen yang telah membimbing dan berbagi ilmu kepada penulis semasa perkuliahan.
2. Seluruh civitas akademika Departemen Statistika ITS yang telah membantu kelancaran dan kemudahan dalam pelaksanaan kegiatan perkuliahan.
3. Pihak PT. Semen Indonesia (Persero) Tbk. khususnya pabrik Semen Tuban yang telah membantu dan membimbing penulis untuk menyelesaikan tugas akhir ini.
4. Ibu dan Bapak yang sangat penulis cintai dan hormati, Kustono dan Farida, tanpa doa dan seluruh dukungan mereka, penulis bukanlah apa-apa.
5. Prima Esti Retnani dan Arifianto sebagai kakak yang peduli dengan adanya Tugas Akhir ini dan tidak pernah lelah mendorong penulis untuk menyelesaikan Tugas Akhir.
6. Affanda, Rahmat, Ilma dan Rossy serta teman-teman PIONEER 2014 dan angkatan Lintas Jalur Statistika ITS 2017 yang telah membantu penulis dalam masa perkuliahan

7. Teman-teman diluar departemen Statistika ITS yang ikut turut menghibur dan menyemangati pembuatan tugas akhir ini.
8. Semua pihak yang telah memberikan dukungan yang tidak dapat disebutkan satu persatu oleh penulis.

Penulis menyadari bahwa laporan Tugas Akhir ini masih jauh dari kata sempurna, oleh karena itu penulis sangat mengharapkan kritik dan saran yang membangun agar dapat menjadi bahan untuk perbaikan berikutnya.

Surabaya, Juli 2019

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA	
2.1 <i>Pre-Processing</i>	7
2.2 <i>Recursive Feature Elimination</i>	7
2.3 <i>K-Fold Cross Validation</i>	8
2.4 Analisis Regresi Linier.....	9
2.4.1 Asumsi IIDN.....	9
2.5 <i>Principle Component Regression</i>	10
2.6 <i>Support Vector Regression</i>	12
2.6.1 Fungsi <i>Kernel</i>	14
2.7 Pemilihan Model Terbaik.....	14
2.7.1 <i>Root Mean Square Error</i>	15
2.8 Kuat Tekan Semen.....	15
2.9 Profil PT. Semen Indonesia (Persero) Tbk.....	15
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data.....	19
3.2 Variabel Penelitian.....	19
3.3 Langkah Analisis.....	20
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Eksplorasi Data Semen PCC.....	23

4.2	<i>K-Fold Cross Validation</i> Data Semen PCC	24
4.3	Analisis Regresi Linier Data Semen PCC	25
4.4	<i>Pinciple Component Regresion</i> Data Semen PCC	26
	4.4.1 <i>Modelling</i> PCR	26
	4.4.2 Validasi Model PCR.....	30
4.5	<i>Support Vector Regression</i> Data Semen PCC	31
	4.5.1 SVR Dengan <i>Complete Feature</i>	32
	4.5.1.1 <i>Modelling</i> SVR.....	32
	4.5.1.2 Validasi Model SVR	35
	4.5.2 SVR Dengan <i>Feature Selection</i>	36
	4.5.2.1 RF- <i>Recursive Feature Elimination</i>	37
	4.5.2.2 <i>Modelling</i> SVR Dengan RF-RFE.....	37
4.6	Pemilihan Model Terbaik Data Semen PCC	40
4.7	Prediksi Kuat Tekan Semen PCC.....	42
BAB V KESIMPULAN DAN SARAN		
5.1	Kesimpulan.....	43
5.2	Saran.....	43
DAFTAR PUSTAKA		
LAMPIRAN.....		
BIOGRAFI PENULIS		

DAFTAR TABEL

Tabel 3.1	Variabel Penelitian	19
Tabel 4.1	Statistika Deskriptif	24
Tabel 4.2	<i>Variance Inflation Factor</i>	25
Tabel 4.3	<i>Eigenvalue</i>	27
Tabel 4.4	Komponen Utama.....	27
Tabel 4.5	Skor Komponen Utama	28
Tabel 4.6	Uji Parsial	29
Tabel 4.7	Validasi Model PCR.....	31
Tabel 4.8	<i>Tuning Parameter Kernel-Linear</i>	32
Tabel 4.9	<i>Tuning Parameter Kernel-Polynomial</i>	33
Tabel 4.10	<i>Tuning Parameter Kernel-RBF</i>	34
Tabel 4.11	<i>Parameter Kernel</i>	35
Tabel 4.12	Validasi Model SVR.....	35
Tabel 4.13	Rata-Rata Kriteria Setiap Model <i>Kernel</i>	36
Tabel 4.14	<i>Ranking RF-RFE</i>	37
Tabel 4.15	<i>Tuning Parameter Kernel-Linear RF-RFE</i>	38
Tabel 4.16	<i>Tuning Parameter Kernel-Polynomial RF-RFE</i>	38
Tabel 4.17	<i>Tuning Parameter Kernel-RBF RF-RFE</i>	39
Tabel 4.18	Pemilihan Model Terbaik	40
Tabel 4.19	Koefisien Parameter Model SVR	41

DAFTAR GAMBAR

Gambar 2.1	<i>Insensitive Zone</i>	12
Gambar 2.2	Ilustrasi Proses Pembuatan Semen.....	16
Gambar 3.1	Diagram Alir	21
Gambar 4.1	Pola Data Kuat Tekan Semen PCC Hari ke-28...	23
Gambar 4.2	Asumsi IIDN Data Semen PCC	30
Gambar 4.3	Prediksi Kuat Tekan Semen PCC April 2019	42

DAFTAR LAMPIRAN

Lampiran 1	Data Kuat Tekan Semen PCC Mill #8	47
Lampiran 2	10-Fold Cross Validation.....	48
Lampiran 3	Hubungan Antar Variabel (<i>Pearson Corr.</i>)	48
Lampiran 4	<i>Output</i> Analisis Regresi Linier.....	49
Lampiran 5	<i>Output Eigenvalue</i>	50
Lampiran 6	<i>Output</i> PCA.....	50
Lampiran 7	<i>Output</i> Regresi PCA vs Kuat Tekan Semen	51
Lampiran 8	<i>Syntax</i> PCR dengan <i>Cross Validation</i>	52
Lampiran 9	<i>Syntax Support Vector Regression</i>	53
Lampiran 10	SVR-Complete Feature dengan Parameter	54
Lampiran 11	SVR-Complete Feature Setiap Fold	55
Lampiran 12	<i>Syntax</i> RF-Recursive Feature Elimination.....	56
Lampiran 13	<i>Output</i> RF-RFE	56
Lampiran 14	SVR-Feature Selection (RF-RFE)	57
Lampiran 15	Prediksi Kuat Tekan Semen Bulan April 2019 Menggunakan Model <i>Kernel-Radial Basis</i> <i>Function Complete Feature</i>	58
Lampiran 16	Surat Permohonan Ijin Memperoleh Data.....	59
Lampiran 17	Surat Keaslian Data.....	60
Lampiran 18	Surat Penerimaan Penelitian	61

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia sebagai negara berkembang terus-menerus melakukan pembangunan dalam segala bidang aspek kehidupan. Salah satu sektor yang selalu berkembang adalah sektor infrastruktur. Hal ini ditunjukkan dengan jaranginya keberadaan lahan persawahan maupun tanah lapang yang berada di kota-kota besar seperti Jakarta dan Surabaya, peristiwa ini disebabkan karena banyaknya bangunan yang baru dan digunakan sebagai gedung perkantoran, perumahan, pertokoan dan sebagainya. Perkembangan bidang pembangunan juga selalu berkembang sejalan dengan semakin banyaknya penduduk yang berada di Indonesia. Berdasarkan proyeksi badan perencanaan pembangunan nasional (Bappenas) jumlah penduduk Indonesia pada tahun 2018 mencapai 265 juta jiwa (Bappenas, 2013). Dengan menghadapi fenomena pertumbuhan penduduk ini, khususnya di sektor pembangunan perlu adanya hal yang menunjang dalam setiap pembangunan yang terjadi, salah satunya adalah material-material pembangunan seperti beton, mortar dan semen. Semen digunakan sebagai komponen utama dalam pembangunan infrastruktur seperti gedung, jalan tol, pelabuhan, bandara dan berbagai infra-struktur lainnya, maka dari itu kebutuhan semen selalu meningkat, hal ini ditunjukkan dengan penjualan semen yang mengalami kenaikan setiap tahunnya. Penjualan semen di Indonesia mulai 2002 dari 27.527.968 ton kemudian tahun 2003 sebanyak 30.191.188 ton dan pada tahun 2016 sebesar 62.007.529 ton, dapat disimpulkan bahwa selama 14 tahun penjualan semen di Indonesia meningkat sebesar 125% dari titik sebelumnya (CEIC, 2017).

Salah satu perusahaan yang bergerak di bidang produksi semen adalah PT. Semen Indonesia (Persero) Tbk. Semen Indonesia merupakan BUMN *Multinomial Company* pertama yang ada di Indonesia. Semen Indonesia pertama diresmikan pada tanggal 7 Agustus 1957 yang kemudian menjadi perusahaan

ternama di Indonesia. Dengan bertambah pesatnya bisnis industri semen membuat semakin ketat pula persaingan antar industri dalam memperebutkan *customer* serta mempertahankan pasar yang ada. Pihak Semen Indonesia memantau berbagai aspek agar dapat mempertahankan kepercayaan pelanggan. Semen Indonesia memiliki berbagai cara agar menjaga kualitas produksinya, diantaranya adalah melakukan evaluasi pada setiap produksi dan selalu melihat dari keluhan pelanggan dan kemudian melakukan evaluasi setelahnya agar kualitas semen bisa terjaga. Sebenarnya banyak sekali yang dapat dilakukan untuk menjaga kualitas semen, tidak harus menunggu keluhan konsumen. Salah satu penilaian inti dari semen adalah kuat tekan semen tersebut, jika kuat tekan semakin bagus, maka pelanggan akan semakin puas. Selama ini pengujian kuat tekan hanya dilakukan dengan mesin, sedangkan dengan adanya revolusi industri 4.0 yaitu *IoT (Internet of Things)*, Semen Indonesia harus melakukan *improvement* disetiap sistem yang berjalan. Maka dari itu, Semen Indonesia harus melakukan beberapa strategi agar *customer* tidak beralih ke produk semen lainnya. Dengan tujuan tersebut maka diperlukan inovasi dan evaluasi pada setiap produk dan sistematis yang berjalan di PT. Semen Indonesia (Persero) Tbk.. Terdapat beberapa produk yang dihasilkan oleh PT. Semen Indonesia (Persero) Tbk., yaitu OPC (*Ordinary Portland Cement*), PPC (*Portland Pozzolan Cement*) dan PCC (*Portland Composite Cement*).

Portland Composite Cement (PCC) merupakan bahan pengikat hidraulis hasil penggilingan bersama-sama terak semen *portland* dan gips dengan satu atau lebih bahan anorganik, atau hasil pencampuran antara bubuk semen *portland* dengan bubuk bahan anorganik lainnya, bahan anorganik tersebut antara lain terak tanur tinggi, pozolan, senyawa silikat, batu kapur dengan kadar total bahan anorganik sekitar 6%-35% dari massa *Portland Composite Cement* (PCC) (SNI 15-7064, 2004). Semen akan mengeras jika dicampur dengan air dan berguna untuk merekat-kan bahan bangunan lainnya seperti batu dan kerikil, maka dari itu semen *portland* disebut semen hidraulis. Komposisi kimia semen

portland umumnya terdiri dari CaO , SiO_2 , Al_2O_3 dan Fe_2O_3 . Dari hasil produksi PCC ini PT. Semen Indonesia (Persero) Tbk. perlu untuk terus meningkatkan kualitas produknya, karena terlihat dari meningkatnya penjualan semen disetiap tahunnya di Indonesia, khususnya pada perusahaan Semen Indonesia.

Statistikan banyak memiliki peranan dalam bidang produksi, misalnya pengumpulan data, pengolahan data, interpretasi maupun pengambilan keputusan. Dalam bidang produksi, PT. Semen Indonesia (Persero) Tbk. membutuhkan evaluasi terus menerus agar menghasilkan suatu produk yang memiliki kualitas yang baik dan memenuhi standar kualitas yang telah ditetapkan oleh perusahaan. Seiring perkembangan jaman, evaluasi yang dilakukan perusahaan bersifat konvensional. Dengan adanya revolusi industri 4.0 yang berbasis teknologi digital, seharusnya PT. Semen Indonesia (Persero) Tbk. dapat meringankan beban pekerjaan dengan memanfaatkan fenomena ini. Hal ini didukung karena adanya *Big Data* yang disebabkan oleh meningkatnya penjualan semen disetiap tahunnya.

Dalam penelitian ini akan diambil salah satu produk dari PT. Semen Indonesia (Persero) Tbk. yaitu PCC (*Portland Composite Cement*). Variabel yang digunakan adalah variabel kuat tekan dengan beberapa komposisi senyawa dan beberapa proses, kemudian akan dilakukan estimasi terhadap kuat tekan semen. Kuat tekan semen diuji pada hari ke-7, ke-14 dan hari ke-28, namun pada penelitian ini hanya dilakukan estimasi pada kuat tekan semen hari ke-28, karena pada hari ke-28, semen tersebut telah menjadi acuan apakah produk telah memiliki komposisi yang tepat atau belum, jika kuat tekan mengalami penurunan kekuatan, akan dilakukan evaluasi dari proses awal pembuatan semen dan jika diukur pada hari ke-7 atau hari ke-14, semen masih beresiko mengalami perubahan kualitas dan tidak menja-min apakah dapat menghasilkan semen yang baik pada hari ke-28. Metode yang digunakan untuk mengestimasi kuat tekan semen PCC adalah metode *Support Vector Regression* (SVR) yang merupakan pengembangan metode dari *Support Vector Machine* (SVM). 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 nonlinier. 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 memper-masalahkan hal tersebut. Metode SVR juga dapat mengatasi *overfitting*. Variabel yang menunjang untuk memestimasi ini adalah senyawa kimia yang terkandung didalam semen, kehalusan semen, ekspansi dan beberapa variabel lainnya. Sebelum menggunakan SVR, akan dilakukan pemilihan variabel atau *Feature Selection* dengan metode *Recursive Feature Elimination* 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 akan dibandingkan dengan metode regresi linier yang akan dipilih model terbaiknya diantara kedua model 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*. Hasbi (2014) yang memprediksi kurs rupiah terhadap dollar amerika menggunakan SVR disimpulkan bahwa pada pengujian data *testing* kernel linier dan *polynomial* menghasilkan akurasi yang baik dan eror yang kecil. Hendra (2012) yang menggunakan metode *feature selection* RFE menyimpulkan bahwa dengan RFE memiliki hasil RMSE yang lebih kecil dan *Rquared* lebih besar dibandingkan *complete feature*.

1.2 Perumusan Masalah

Produksi semen yang ada di perusahaan PT. Semen Indonesia (Persero) Tbk. memiliki beberapa jenis seperti OPC,

PPC dan PCC. Namun, pada bulan Juli 2018 produksi *Portland Pozzolan Cement* (PPC) di perusahaan PT. Semen Indonesia (Persero) Tbk. pabrik Tuban telah dihentikan, yang kemudian produksi difokuskan ke produksi PCC. Dalam penjualan yang dilakukan oleh PT. Semen Indonesia (Persero) Tbk., OPC lebih dikenal dan sering digunakan di masyarakat luas, sementara itu jika dipaparkan, produk OPC dengan PPC tidak jauh berbeda, bahkan PPC terdapat kandungan *Pozzolan* yang mempengaruhi kualitas semen. Karena perusahaan menghentikan produk PPC dan beralih ke PCC, maka dari itu akan dilakukan prediksi terhadap kuat tekan semen PCC di hari ke-28, agar dapat mengetahui adanya kenaikan atau penurunan kuat tekan semen produk PCC di hari pertama semen telah selesai diproduksi.

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah melakukan prediksi kuat tekan semen di hari ke-28 dengan menggunakan produk *Portland Composite Cement* (PCC) di PT. Semen Indonesia (Persero) Tbk. pabrik Tuban.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini diharapkan, hasil prediksi dapat menjadi referensi bagi PT. Semen Indonesia (Persero) Tbk. untuk meningkatkan dan mengevaluasi kualitas *Portland Composite Cement* dengan cara melihat kuat tekan semen pada hari ke-28 yang telah diprediksi.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah menggunakan data kualitas pada produksi *Portland Composite Cement* (PCC) di PT. Semen Indonesia (Persero) Tbk. pada bulan Juli 2018 hingga April 2019, dengan variabel dependen adalah kuat tekan semen PCC hari ke-28 serta 11 variabel prediktor yang terbagi menjadi variabel komponen senyawa kimia dan variabel lain-lain. Metode konvensional yang digunakan dalam prediksi kuat tekan semen ini adalah regresi linier dan *Principle Component Regression* (PCR) untuk menanggulangi indikasi multikolinearitas. *Machine*

Learning yang digunakan adalah *Support Vector Regression* (SVR) dengan penggunaan tiga *kernel* yaitu *Linear*, *Polynomial* dan *RBF*. *Feature selection* yang digunakan adalah *Recursive Feature Elimination* (RFE) dengan berbasis *Random Forest* (RF).

BAB II

TINJAUAN PUSTAKA

2.1 *Pre-Processing*

Pre-processing data merupakan sebuah langkah penting dalam data *minning* untuk membuat data lebih berkualitas. Data perlu dilakukan *pre-processing* karena dalam data mentah terdapat data yang tidak lengkap, *noise* dan tidak konsisten. Terdapat beberapa cara untuk melakukan *pre-processing* salah satunya adalah *data cleaning* yang merupakan proses untuk membersihkan data salah satunya adalah membuang data yang bersifat *outlier*. Kemudian adalah *data integration*, yaitu merupa-kan *integrasi* dari data-data yang digunakan seperti korelasi. Selanjutnya adalah *data transformation* yang membuat perubahan data ketika terdapat perbedaan satuan antar variabel dan pola dapat dipahami. Kemudian, yang terakhir adalah *data reduction* yang bertujuan untuk mereduksi data ataupun variabel agar lebih mudah diolah tetapi tidak menghilangkan karakteristik data tersebut (Alfarisi, 2017).

2.2 *Recursive Feature Elimination*

Feature selection adalah upaya untuk memilih fitur subset dari fitur asli yang paling berguna. *Feature extraction* 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* (Nakariyakul & Casasent 2004). Salah satu dari banyak *feature selection* yang tersedia adalah RFE (*Recursive Feature Elimination*).

Salah satu metode yang dapat digunakan dengan RFE ini adalah *Random Forest*. RF menunjukkan kelebihanannya antara lain dapat menghasilkan eror 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 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, dimana seluruh rangkaian prediktor digunakan untuk menghitung peringkat fitur.

2.3. *K-Fold Cross Validation*

Sebelum dilakukan tahap estimasi, dilakukan pembagian data menjadi *training* dan *testing*. *Cross validation* merupakan salah satu metode pembagian data. Metode ini mempartisi data ke dalam dua subset data yang berukuran sama, salah satu sebagai *training* dan salah satu sebagai data *testing*, kemudian dilakukan pertukaran fungsi dari subset sedemikian sehingga subset sebelumnya sebagai *training set* dan *test set*. Metode *k-fold cross validation* menggeneralisasi pendekatan ini dengan mensegmentasi data ke dalam k partisi berukuran sama. Selama proses, prosedur pembagian data *training* dan *testing* diulang sebanyak k kali, sehingga setiap subset akan menjadi data uji dari model. Proses yang telah dilakukan sebanyak k kali akan mendapatkan k buah nilai dari proses pembelajaran. Semua nilai performa ini akan dicari rata-ratanya dan nilai dengan rata-rata tertinggi akan dipilih sebagai model. *k-fold cross validation* memiliki kelebihan dapat mengklasifikasi dataset lebih efisien, namun metode ini memiliki kelemahan dalam proses komputasi yang digunakan akan lebih besar karena akan melakukan proses sebanyak k kali. *Cross*

validation adalah bentuk sederhana dari statistik, jumlah *fold* standar untuk memprediksi tingkat eror dari data adalah dengan menggunakan 10-fold *cross validation* (Witten, dkk., 2011) Selain untuk pembagian data *training* dan *testing*, teknik ini untuk melakukan validasi pada dataset untuk menemukan akurasi yang baik.

2.4. Analisis Regresi Linier

Analisis regresi merupakan sebuah alat statistik yang berguna untuk mendapatkan hubungan fungsional antara dua variabel atau lebih yaitu variabel respon dan prediktor. Tujuan dari analisis regresi agar mendapatkan pengaruh antara variabel prediktor terhadap responnya. Bentuk persamaan umum regresi liner berganda adalah sebagai berikut :

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_i x_{ki} + \varepsilon_i \quad (2.1)$$

keterangan :

y_i : variabel respon ke-i dari model.

$\beta_0, \beta_1, \dots, \beta_k$: parameter dari model.

x_1, x_2, \dots, x_k : variabel prediktor ke-i dari model.

ε_i : galat model ke-i.

(Drapper, 1992).

Dari model yang terbentuk, terdapat beberapa asumsi pada residual yang harus dipenuhi, yaitu asumsi identik, independen dan berdistribusi normal.

2.4.1 Asumsi IIDN

Terdapat beberapa asumsi yang perlu dipenuhi dalam penggunaan metode regresi linear yaitu asumsi identik, independen dan berdistribusi normal. Dalam asumsi ini dapat dilakukan dalam dua hal yaitu secara visual dan secara inferensia. Penelitian ini akan melakukan pendeteksian melalui visual

Asumsi identik merupakan salah satu asumsi residual yang penting dari model regresi. Varians residual harus bersifat homoskedastisitas atau varians residual bersifat identik (Drapper,

1992). Suatu data dikatakan identik apabila plot residualnya menyebar secara acak dan tidak membentuk suatu pola tertentu (Sudjana, 1996).

Asumsi saling bebas (*independent*) atau autokorelasi residual, yang dilakukan untuk mengetahui apakah ada korelasi antar residual. Suatu data dikatakan independen apabila plot residualnya menyebar secara acak dan tidak membentuk suatu pola tertentu (Sudjana, 1996).

Asumsi Residual berdistribusi normal dilakukan untuk melihat apakah residual memenuhi asumsi berdistribusi normal atau tidak. Kenormalan suatu data dapat dilihat dari plotnya. Apabila plot sudah mendekati garis linier, dapat dikatakan bahwa data tersebut memenuhi asumsi yaitu berdistribusi normal (Sudjana, 1996). Jika asumsi kenormalan tidak terpenuhi, estimasi OLS tidak dapat digunakan, sehingga perlu penanganan terhadap pelanggaran asumsi ini. Beberapa pengujian yang dapat dilakukan untuk asumsi distribusi normal adalah *Anderson Darling*, *Kolmogorov-Smirnov*, *Jarque-Bera test*, dan *Skewness-Kurtosis* (Daniel, 1989).

2.5. Principle Component Regression

Principle Component Regression (PCR) merupakan suatu teknik analisis yang mengkombinasikan antara analisis regresi dengan *Principal Component Analysis* (PCA). Analisis Regresi digunakan untuk mengetahui ada tidaknya hubungan antara variabel *dependent* dan *independent*, sedangkan PCA pada dasarnya bertujuan untuk menyederhanakan variabel yang diamati dengan cara menyusutkan (mereduksi) dimensinya. Hal ini dilakukan dengan jalan menghilangkan korelasi di antara variabel independen melalui transformasi variabel asal ke variabel baru (merupakan kombinasi linear dari variabel-variabel asal) yang tidak saling berkorelasi. Dalam hal ini, PCR bisa menanggulangi kasus multikolinearitas. Dari p buah variabel asal dapat dibentuk p buah komponen utama, dipilih k buah komponen utama saja ($k < p$) yang telah mampu menerangkan keragaman data cukup tinggi

(antara 80% sampai dengan 90%) (Johnson & Wichern, 2010, hal. 356). Komponen utama yang dipilih tersebut (k buah) dapat mengganti p buah variabel asal tanpa mengurangi informasi.

Analisis regresi komponen utama (PCR) merupakan analisis regresi variabel *dependent* terhadap komponen-komponen utama yang tidak saling berkorelasi yang telah diolah dengan analisis PCA sebelumnya, regresi komponen utama dapat dinyatakan sebagai berikut :

$$Y = w_o + w_1 K_1 + w_2 K_2 + \dots + w_m K_m + \varepsilon \quad (2.2)$$

$K_1, K_2, K_3, \dots, K_m$ menunjukkan komponen utama yang dilibatkan dalam analisis regresi komponen utama, dengan besaran m lebih kecil daripada banyaknya variabel *independent* yaitu sejumlah p , serta Y sebagai variabel *dependent*. Komponen utama merupakan kombinasi linear dari variabel baku Z , sehingga:

$$\begin{aligned} K_1 &= a_{11}Z_1 + a_{21}Z_2 + \dots + a_{p1}Z_p \\ K_2 &= a_{12}Z_1 + a_{22}Z_2 + \dots + a_{p2}Z_p \\ &\vdots \\ K_m &= a_{1m}Z_1 + a_{2m}Z_2 + \dots + a_{pm}Z_p \end{aligned} \quad (2.3)$$

Apabila K_1, K_2, \dots, K_m dalam persamaan (2.2) didistribusikan kembali ke dalam persamaan regresi komponen utama, yaitu persamaan (2.4) maka diperoleh:

$$\begin{aligned} Y &= w_o + w_1 (a_{11}Z_1 + a_{21}Z_2 + \dots + a_{p1}Z_p) + \\ &w_2 (a_{12}Z_1 + a_{22}Z_2 + \dots + a_{p2}Z_p) + \dots + \\ &w_m (a_{1m}Z_1 + a_{2m}Z_2 + \dots + a_{pm}Z_p) + \varepsilon \end{aligned} \quad (2.4)$$

Persamaan regresi linear dugaan komponen utama sebagai berikut:

$$Y = b_0 + b_1 Z_1 + b_2 Z_2 + \dots + b_p Z_p \quad (2.5)$$

2.6. Support Vector Regression

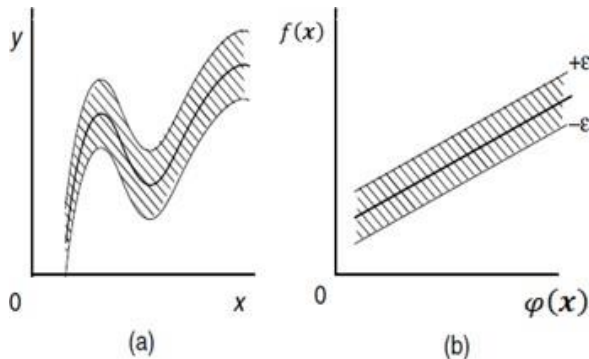
Support Vector Regression (SVR) merupakan suatu metode *Machine Learning* yang dikembangkan dari metode *Support Vektor Machine (SVM)*. Yang berbeda dari kedua metode ini adalah jika SVM satu kumpulan teknik klasifikasi, sedangkan SVR digunakan pada kasus regresi. Tujuan dari SVR adalah untuk menemukan sebuah fungsi sebagai suatu *hyperplane* (garis pemisah) berupa fungsi regresi yang mana sesuai dengan semua input data dengan sebuah error dan membuat error sekecil mungkin (Scholkopf dan Smola, 2002).

Tujuan lain dari SVR ini adalah untuk memetakan *vector* input ke dalam dimensi yang lebih tinggi. Misalkan sebuah fungsi berikut adalah garis regresi sebagai *optimal hyperplane* :

$$f(x) = w^T \varphi(x) + b \quad (2.6)$$

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

$$E(r) = \begin{cases} 0, & \text{untuk } |r| \leq \varepsilon \\ |r| - \varepsilon, & \text{untuk yang lain} \end{cases} \quad (2.7)$$



Gambar 2.1 *Insensitive zone* (a) *original input space*, dan (b) *feature space*

(Sumber : ISSN 0973-1768 Volume 12, Number 4, 2016).

$D(x, y)$ adalah jarak terjauh *support vector* dari *hyperplane*, kemudian disebut margin. Memaksimalkan margin akan meningkatkan probabilitas data ke dalam radius $\pm \varepsilon$. Jarak dari *hyperplane* $D(x, y) = 0$ ke data adalah (x, y) adalah $|D(x, y)| / \|W^*\|$, dengan:

$$W^* = (1 - W^T)^T \quad (2.8)$$

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

$$\begin{aligned} \frac{|D(x, y)|}{\|W^*\|} &\leq \delta \\ |D(x, y)| &\leq \delta \|W^*\| \\ \delta \|W^*\| &= \varepsilon \end{aligned} \quad (2.9)$$

Oleh karena itu untuk memaksimalkan margin φ , diperlukan $\|W^*\|$ yang minimum, Optimasi penyelesaian masalah dengan bentuk *Quadratic Programming* adalah sebagai berikut :

$$\min \frac{1}{2} \|W^*\|^2 \quad (2.10)$$

dengan syarat :

$$y_i - W^T \varphi(x_i) - b \leq \varepsilon \text{ untuk } i = 1, \dots, l \quad (2.11)$$

$$W^T \varphi(x_i) - y_i + b \leq \varepsilon \text{ untuk } i = 1, \dots, l \quad (2.12)$$

Faktor $\|W^*\|$ dinamakan regulas. Meminimalkan $\|W^*\|$ akan membuat suatu fungsi setipis (*flat*) mungkin, sehingga bias mengontrol kapasitas fungsi (*function capacity*) (Abe, 2005).

2.6.1 Fungsi Kernel

Menurut Santosa (2007) banyak teknik *Machine Learning* yang dikembangkan dengan asumsi kelinieran, sehingga algoritma yang dihasilkan terbatas untuk kasus-kasus yang linier. Dengan metode kernel suatu data x di *input space* dipetakan ke *feature space* dengan dimensi yang lebih tinggi melalui φ . Fungsi Kernel yang digunakan adalah :

1. Kerner Linier

$$\varphi(x) = K(x, x') = x^T x \quad (2.13)$$

2. Kernel *Polynomial*

$$\varphi(x) = K(x, x') = (\gamma(x^T x) + 1)^d \quad (2.14)$$

3. *Radial Basis Function* (RBF)

$$\varphi(x) = K(x, x') = \exp(-\gamma \|x - x_i\|^2) \quad (2.15)$$

Nilai $K(x_i, x_j)$ merupakan fungsi kernel yang menunjukkan pemetaan linier pada *feature space*. Perlu dijelaskan bahwa 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 dan sulit dihitung. Sedangkan x dan x' adalah pasangan dua data *training*. Parameter $d > 0$ merupakan konstanta. Fungsi kernel mana yang harus digunakan untuk substitusi *dot product* di *feature space* sangat tergantung pada data karena fungsi kernel ini akan menentukan fitur baru dengan fungsi pemisah akan dicari.

2.7. Pemilihan Model Terbaik

Pemilihan model terbaik dapat dilakukan dengan berbagai metode. Hal ini dilakukan karena terciptanya beberapa model yang layak pakai. Metode yang digunakan dalam penelitian ini adalah *Root Mean Square Error* (RMSE) karena RMSE lebih stabil

terhadap varians data yang cukup besar dan baik dalam memberikan penalti atau hukuman terhadap nilai eror yang besar.

2.7.1 *Root Mean Square Error*

RMSE (*Root Mean Square Error*) merupakan kriteria pemilihan model terbaik berdasarkan pada data yang telah dibagi menjadi data *training* dan data *testing*. Model terbaik dipilih yang memiliki nilai kriteria error terkecil (Gooijer dan Hyndman, 2006).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.16)$$

2.8. Kuat Tekan Semen

Kuat tekan merupakan kekuatan tekan maksimum yang dapat dipikul benda persatuan luas. Semen yang dicampur dengan air dan menjadi semen hidraulis akan diuji kuat tekannya. Biasanya, kuat tekan diukur pada hari ke-3, hari ke-7, hari ke-14, hari ke-21 dan hari ke-28. Kuat tekan selalu dipantau di setiap minggunya karena akan melihat perubahan pengerasan semen yang terjadi. Kuat tekan pada hari ke-3 minimum 160 kg/cm², pada hari ke-7 minimum 230 kg/cm² dan hari ke-28 minimum 320 kg/cm². Nilai kuat tekan semen didapatkan melalui tata cara pengujian standar, menggunakan mesin uji dengan cara memberikan beban bertingkat dengan kecepatan peningkatan beban tertentu dengan benda uji (SNI-2049, 2004). Penentuan kuat tersebut mengacu kepada ASTM C 109/109M-02 (*Standard Test Method for Compressive Strength of Hydraulic Cement Mortar*). Metode uji ini melingkupi penentuan kuat tekan mortar semen hidraulis dengan menggunakan cetakan kubus berukuran sisi 50 mm.

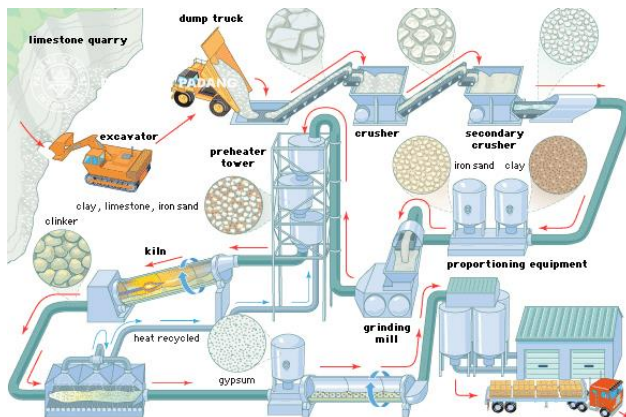
2.9. Profil PT. Semen Indonesia (Persero) Tbk.

PT Semen Indonesia (Persero) Tbk., yang sebelumnya bernama PT Semen Gresik (Persero) Tbk. merupakan Badan Usaha Milik Negara (BUMN) yang bergerak di bidang industri bahan bangunan. Perseroan berperan sebagai Strategic Holding Company dengan berbagai lini usaha yang menawarkan solusi lengkap dalam pembangunan.

Diresmikan pada tanggal 7 Agustus 1957 oleh Ir Soekarno Presiden pertama Republik Indonesia, Perseroan menjadi penopang pembangunan Indonesia pada masa awal kemerdekaan hingga saat ini. Pada tahun 1991, Perseroan mencatatkan diri sebagai perusahaan BUMN pertama yang Go Public di Bursa Efek Surabaya dan Bursa Efek Jakarta (sekarang menjadi Bursa Efek Indonesia), dengan kode emiten SMGR.

Dalam perkembangannya, Perseroan melakukan langkah strategis dengan mengakuisisi dua perusahaan BUMN lain, PT Semen Padang (Persero) dan PT Semen Tonasa (Persero) dan menjadi perusahaan persemenan terbesar di Indonesia. Seiring dengan visi perusahaan, tahun 2012 Perseroan melakukan langkah korporasi dengan mengakuisisi Thang Long Cement Company (TLCC) Vietnam.

Dalam industri persemenan, kapasitas produksi Perseroan terus mengalami pertumbuhan. Saat ini, Perseroan memiliki 14 Integrated Cement Plant yang tersebar di Indarung (Sumatera Barat), Tuban (Jawa Timur), Pangkep (Sulawesi Selatan), Rembang (Jawa Tengah) dan Quang Ninh (Vietnam) dengan total kapasitas terpasang sebesar 31,8 juta Ton semen per tahun.



Gambar 2.2 Ilustrasi Proses Pembuatan Semen
(Sumber: tentangteknikkimia.wordpress.com)

Pembuatan semen menggunakan bahan baku utama Batu Kapur dan Tanah Liat yang diambil dari proses penambangan di *Quarry* milik Perseroan. Penambangan Batu Kapur dilakukan dengan cara peledakan dan *Surface Minner*, sedangkan untuk mem-peroleh Tanah Liat dilakukan dengan cara pengerukan. Selanjutnya Batu Kapur dan Tanah Liat diangkut ke *crusher* dengan *dump truck*. Batu Kapur dan Tanah Liat dikecilkan ukurannya sampai 8 cm di *Crusher* untuk kemudian disimpan di *Stock Pile (storage)*.

Bahan baku yang didapat dari proses penambangan (Batu Kapur dan Tanah Liat) akan ditampung di dalam *storage* untuk selanjutnya dilakukan proses *prehomogenisasi* yang disebut *reclaimer*. Proses *prehomogenisasi* di *reclaimer* adalah proses yang sangat penting untuk menjamin kualitas dari produk yang dihasilkan baik dari *raw meal* hingga produk akhir, yaitu semen. Dari *Stock Pile* dimasukkan ke *Raw Mill* ditambahkan Pasir Besi dan Pasir Silika untuk digiling dan dikeringkan menjadi *Raw Meal*. *Raw Meal* atau tepung baku adalah bahan baku untuk pembuatan terak (*Clinker*). *Raw Meal* berbentuk seper-ti *powder* yang mempunyai kehalusan tertentu. *Raw Meal* mempunyai sifat fisika dan sifat kimia tertentu yang digunakan sebagai kontrol kualitas produk. Sifat kimia digunakan sebagai pengatur proporsi bahan-bahan yang akan diumpankan ke dalam pro-ses. *Raw Meal* dihasilkan dari sebuah sistem peralatan yaitu *Raw Mill Plant* yang terdiri dari alat-alat utama, sistem transport dan alat-alat separasi untuk kemudian disimpan di *Raw Meal Silo*.

Kemudian dilakukan pemanasan dan pembakaran menggunakan bahan bakar batu bara yang telah digiling dan dikeringkan melalui *Coal Mill*. Setelah itu, dilanjutkan dengan penggilingan *clinker* yang dengan proses ini akan menjadi semen yang dapat ditemui di pasaran, sebelum itu dilakukan penjualan, semen akan dikemas secara baik didalam bungkus yang berkualitas baik.

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh pada produksi bulan Juli 2018 hingga Januari 2019. Data diperoleh dari PT. Semen Indonesia (Persero) Tbk. pabrik Tuban dengan produk yang diamati adalah produk *Portland Composite Cement* (PCC).

3.2 Variabel Penelitian

Berikut merupakan variabel penelitian yang dimana merupakan karakteristik kualitas dari produk *Portland Composite Cement* yang sesuai dengan standar perusahaan.

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan	Satuan	Skala Data
Y	Kuat Tekan Hari ke-28	Kg/cm ²	Rasio
X_1	Oksida Silikon (SiO ₂)	%	Rasio
X_2	Oksida Alumunium (Al ₂ O ₃)	%	Rasio
X_3	Oksida Besi (Fe ₂ O ₃)	%	Rasio
X_4	Oksida Kalsium (CaO)	%	Rasio
X_5	Magensium Oksida (MgO)	%	Rasio
X_6	Sulfur Trioksida (SO ₃)	%	Rasio
X_7	<i>Freelime</i>	%	Rasio
X_8	<i>Insoluble Residu</i>	%	Rasio
X_9	<i>Loss of Ignition</i>	%	Rasio
X_{10}	<i>Blaine</i>	Luasan/massa	Rasio
X_{11}	<i>Residu</i>	%	Rasio

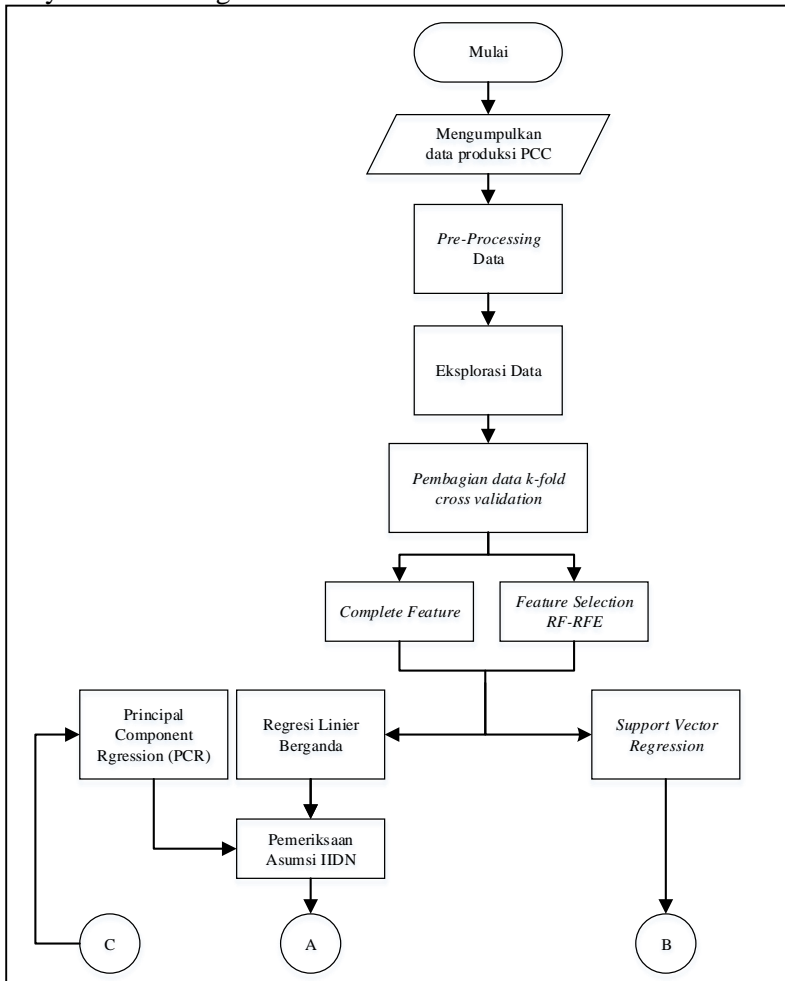
3.3 Langkah Analisis

Berikut ini adalah langkah analisis yang digunakan dalam prediksi kuat tekan produk *Portland Composite Cement* (PCC).

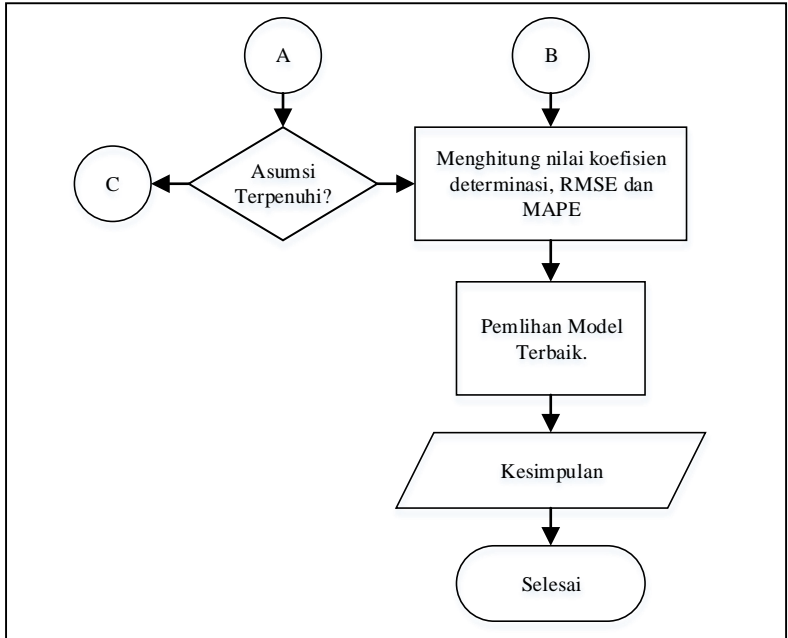
1. Mengumpulkan data sekunder hasil produksi pada produk PCC di PT. Semen Indonesia (Persero) Tbk.
2. Melakukan *pre-processing* pada data produksi PCC.
3. Melakukan eksplorasi data produksi PCC.
4. Pembagian data yaitu data *training* dan data *testing* menggunakan *k-fold cross validation*. Label *fold* pada data dengan variabel prediktor lengkap maupun yang terpilih adalah sama.
5. Melakukan pemilihan variabel dengan RFE dengan berbasis *Random Forest* (RF-RFE) untuk menentukan variabel yang akan dianalisis selanjutnya.
6. Pemodelan regresi linier dengan menggunakan Persamaan (2.1) dan pemodelan PCR menggunakan Persamaan (2.2).
 - a. Pemodelan dengan regresi linier didahului dengan pemeriksaan hubungan antara semua variabel, jika terdapat hubungan antara variabel prediktor dengan respon dan telah diindikasikan terdapat multikolinearitas maka pemodelan menggunakan PCR.
 - b. Melakukan pemodelan regresi linier berganda atau PCR
 - c. Pemeriksaan asumsi identik, independen, distribusi normal dan asumsi non-multikolinearitas. Jika terdapat pelanggaran asumsi, maka dilakukan penaggulangan.
 - d. Menghitung nilai RMSE dengan Persamaan (2.16).
7. Melakukan pemodelan SVR dengan menggunakan Persamaan (2.6) dengan fungsi *kernel* di Persamaan (2.13).
 - a. Melakukan pemodelan dengan SVR menggunakan semua variabel dan beberapa variabel yang terpilih dari langkah 4.
 - b. Melakukan *tuning* parameter agar mendapatkan parameter optimum dari ketiga *kernel*.
 - c. Mendapatkan model dan menghitung nilai RMSE dengan Persamaan (2.16).

8. Pemilihan model terbaik dengan menggunakan kriteria terbaik RMSE di Persamaan (2.16).
9. Menarik kesimpulan dan saran.

Berdasarkan langkah analisis yang telah dibuat, maka diagram alirnya adalah sebagai berikut:



Gambar 3.1 Diagram Alir



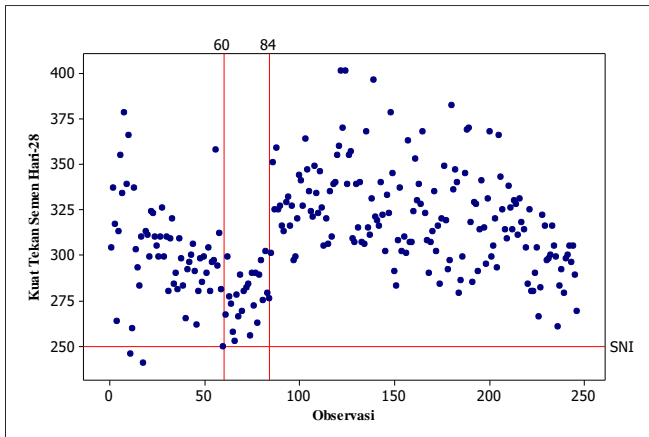
Gambar 3.1 Diagram Alir (Lanjutan)

BAB IV ANALISIS DAN PEMBAHASAN

Pada penelitian ini, analisis dan pembahasan mencakup beberapa tahap yaitu eksplorasi data, *feature selection* dengan RFE, pembagian data *training* dan *testing* dengan *k-fold cross validation*, analisis regresi linear serta identifikasi asumsinya, kemudian penggunaan metode *support vector regression* serta pemilihan model terbaik. Tahapan prediksi akan dilakukan setelah mendapatkan model terbaiknya.

4.1 Eksplorasi Data Semen PCC

Eksplorasi data bertujuan untuk mengetahui karakteristik data secara visual dan secara inferensia. Secara visual dapat ditunjukkan dengan beberapa *tools*. Berikut adalah pola data dari kuat tekan semen pada hari ke 28 dari bulan Juli 2018 hingga Maret 2019.



Gambar 4.1 Pola Data Kuat Tekan Semen PCC Hari ke-28

Gambar 4.1 memberikan visualisasi tentang pola data kuat tekan semen *PCC* pada hari ke-28 di *Mill 8* pabrik Tuban, bahwa hasil kuat tekan semen *PCC* hari ke-28, memiliki pola yang fluktuatif. Terlihat pula bahwa pada observasi ke-60 atau tepatnya

tanggal 5 September 2018 hingga observasi ke-84 yaitu 1 Oktober 2018 mengalami penurunan kuat tekan semen, setidaknya hampir 1 bulan penuh yaitu bulan Oktober 2018, kuat tekan semen PCC mengalami penurunan dibandingkan kuat tekan semen pada bulan lainnya. Berdasarkan acuan SNI 15-2049-2004 tentang semen portland komposit (PCC) bahwa persyaratan kuat tekan semen PCC pada hari ke-28 adalah minimal 250 kg/cm². Sedangkan, pada kuat tekan semen PCC pabrik semen Tuban, hanya 3 observasi atau sekitar 1% dari seluruh produksi dari bulan Juli 2018 hingga Maret 2019 yang memiliki kuat tekan semen dibawah persyaratan, sedangkan sisanya sudah memenuhi persyaratan SNI. Untuk eksplorasi lebih lanjut dapat dilihat hasil inferensia berikut ini:

Tabel 4.1 Statistika Deskriptif

Variabel	N	SNI	Rata-rata	Min.	Maks.	Median	Varians
Kuat Tekan (Y)	246	250	312,28	241	401	310	852,34

Tabel 4.1 memberikan informasi bahwa terdapat 246 data kuat tekan semen PCC yang digunakan, dan tidak terdapat data *missing* dalam penelitian ini. Rata-rata kuat tekan semen memiliki kekuatan tekan sekitar 312,28 kg/cm², dengan nilai varians yang besar dan berbanding lurus dengan gambar 4.1 yang mengindikasikan bahwa data kuat tekan semen PCC ini sangat beragam dengan nilai minimum tekan semen adalah 241 kg/cm² dan nilai maksimum sebesar 401 kg/cm². Nilai minimum kuat tekan semen berbeda seperti dugaan sebelumnya yaitu bulan September 2018, akan tetapi nilai minimum kuat tekan semen berada pada bulan Juli 2018. Bisa disimpulkan bahwa bulan Juli-September 2018 memiliki kuat tekan semen yang sangat berfluktuatif, karena pada bulan Juli 2018 sendiri adalah bulan pertama produksi PCC setelah produk PPC (*Portland Pozzolan Cement*) dihentikan di pabrik semen Tuban.

4.2 K-Fold Cross Validation Data Semen PCC

Dalam penelitian kali ini menggunakan 10 *fold cross validation*. Dengan jumlah data kuat tekan semen sebanyak 246

observasi, maka setiap *subset* akan memiliki 24 dan 25 observasi. Pengambilan data sebanyak 10 *fold* bersifat *random* dan tidak ada pengulangan data dalam metode ini. Penentuan *k* sebesar 10 *fold* dikarenakan jumlah data yang cukup banyak. Penentuan ini membuat persentase data menjadi 90% untuk *training* dan untuk *testing* 10% yang jika dilakukan prediksi akan menghasilkan hasil yang baik, karena prediksi yang baik tidak dilakukan dengan periode yang panjang.

4.3 Analisis Regresi Linear Data Semen PCC

Analisis regresi linear pada penelitian kali ini menggunakan variabel dependen kuat tekan semen PCC pada hari ke-28 dengan variabel independennya sebanyak 11 variabel. Sebelum melakukan pemodelan, perlu diketahui hubungan antara variabel dependen dengan masing-masing variabel independen.

Hasil hubungan antar variabel menggunakan *pearson correlation* yang terletak pada Lampiran 3. memperlihatkan bahwa hasil hubungan antara beberapa variabel independen dengan variabel dependen. Terlihat bahwa hanya 6 variabel independen yang berhubungan dengan variabel dependen, sedangkan variabel lainnya tidak berhubungan. Selain itu, banyak variabel independen yang memiliki korelasi yang kuat dengan variabel independen lain yang menyebabkan tidak terpenuhinya asumsi *non-multikolinearitas* pada model regresi linier. Untuk melihat apakah benar-benar terdapat multikolinearitas didalam penelitian ini, maka akan dilihat nilai *Variance Inflation Factor* (VIF).

Tabel 4.2 *Variance Inflation Factor*

Predictor	VIF
x1	14,656
x2	16,098
x3	2,397
x4	2,077
x5	1,448
x6	2,151
x7	1,101

Tabel 4.2 *Variance Inflation Factor* (Lanjutan)

Predictor	VIF
X ₈	2,573
X ₉	4,701
X ₁₀	3,057
X ₁₁	1,084

Tabel 4.2 memperlihatkan nilai VIF pada variabel SiO (x_1) dan Al (x_2) melebihi angka 10, yang menunjukkan bahwa terdapat multikolinearitas pada penelitian ini dan yang menyebabkan terjadinya adalah variabel SiO dan Al. Dikarenakan fenomena ini, sehingga penelitian tidak dapat dilanjutkan ke analisis regresi linier.

Terdapat beberapa cara untuk menanggulangi fenomena multikolinearitas, salah satu cara yang termudah adalah dengan mengeluarkan variabel independen yang mempunyai korelasi yang kuat dengan variabel independen lain dan tidak berkorelasi dengan variabel dependen, namun solusi ini mempunyai resiko terbuangnya informasi data. Alternatif lainnya dengan menggunakan *Principal Component Regression* (PCR), dengan tujuan mereduksi variabel tanpa membuang informasinya.

4.4 *Principle Component Regression Data Semen PCC*

PCR merupakan teknik analisis regresi yang dikombinasikan dengan teknik analisis komponen utama (PCA). Model PCR akan dibentuk menggunakan keseluruhan data dengan *complete feature*. Lalu, model PCR akan divalidasi menggunakan *10 fold cross validation* untuk mendapatkan nilai kriteria model terbaik dan RMSE.

4.4.1 *Modelling PCR*

Sebelum melakukan PCR, akan dicari nilai *eigenvalue* yang dalam analisis faktor terdapat beberapa komponen yang merupakan variabel. Setiap faktor mewakili variabel yang dianalisis. Kemampuan setiap faktor mewakili variabel yang dianalisis ditunjukkan oleh besarnya varians yang dijelaskan, yang disebut *eigenvalue*. Kriteria pemilihan faktor dengan melihat nilai *eigenvalue* yang bernilai lebih besar sama dengan satu. Dalam

analisis ini akan dilakukan pemodelan menggunakan data kuat tekan semen PCC Juli 2018 hingga Maret 2019. Berikut adalah hasil analisisnya.

Tabel 4.3 Eigenvalue

Faktor	<i>Eigenvalue</i>	Proporsi	Kumulatif
1	3,4639	0,315	0,315
2	1,8505	0,168	0,483
3	1,5367	0,140	0,623
4	1,0768	0,098	0,721
5	0,9046	0,082	0,803
6	0,8895	0,081	0,884
7	0,5644	0,051	0,935
8	0,2933	0,027	0,962
9	0,2446	0,022	0,984
10	0,1428	0,013	0,997
11	0,0329	0,003	1,000

Pada Tabel 4.3 terlihat bahwa nilai *eigenvalue* yang memiliki nilai lebih dari satu yaitu di faktor satu hingga faktor empat, dengan keempat faktor ini dapat menjelaskan data sebesar 72,1%. Langkah selanjutnya adalah mencari nilai skor komponen yang terbentuk untuk membentuk persamaan regresi komponen utama, yang nantinya persamaan ini akan membantu untuk mengembalikan ke persamaan awal, berikut adalah hasilnya.

Tabel 4.4 Komponen Utama

Variabel	PC1	PC2	PC3	PC4
Z ₁	-0,50068	0,02702	0,04558	0,03216
Z ₂	-0,50752	0,09249	0,08563	-0,02254
Z ₃	-0,38385	-0,09894	0,11251	0,07979
Z ₄	0,08497	-0,50530	0,07712	0,57451
Z ₅	0,04033	-0,09476	0,53391	-0,45016
Z ₆	0,21251	0,24892	-0,58534	-0,17937
Z ₇	-0,00965	0,27268	-0,04212	0,51350

Tabel 4.4 Komponen Utama (Lanjutan)

Variabel	PC1	PC2	PC3	PC4
Z ₈	-0,29523	0,52017	-0,02928	-0,04613
Z ₉	0,41803	0,26134	0,22594	-0,00228
Z ₁₀	0,16842	0,38505	0,52348	0,20694
Z ₁₁	0,00505	-0,30523	0,01775	-0,34435

Berdasarkan Tabel 4.4 yang memperlihatkan keempat faktor yang terbentuk dapat dibentuk persamaan regresi sebagai berikut.

$$K_1 = -0,50068Z_1 - 0,50752Z_2 - 0,38385Z_3 + 0,08497Z_4 + 0,04033Z_5 + \dots + 0,00505Z_{11}$$

$$K_2 = 0,02702Z_1 + 0,09249Z_2 - 0,09894Z_3 - 0,50530Z_4 - 0,09476Z_5 + \dots - 0,30523Z_{11}$$

$$K_3 = 0,04558Z_1 + 0,08563Z_2 + 0,11251Z_3 + 0,07712Z_4 + 0,53391Z_5 + \dots + 0,01775Z_{11}$$

$$K_4 = 0,03216Z_1 - 0,02254Z_2 + 0,07979Z_3 + 0,57451Z_4 - 0,45016Z_5 + \dots - 0,34435Z_{11}$$

Setelah komponen utama terbentuk, langkah selanjutnya adalah menentukan variabel-variabel baru untuk menggantikan variabel independen sebelumnya. Variabel pengganti juga berjumlah empat variabel karena mengikuti faktor yang terbentuk, berikut adalah variabel yang baru.

Tabel 4.5 Skor Komponen Utama

i	K _{1i}	K _{2i}	K _{3i}	K _{4i}
1	-4,69357	0,77447	3,72359	-0,95912
2	-5,24951	-3,44305	0,73046	-0,99116
3	-2,03896	-1,56704	0,46048	-0,31848
4	-0,53945	-0,23737	2,40233	2,44069
5	-2,75136	-0,72927	2,37485	0,80409
6	-5,14789	-2,06380	-0,12851	-0,82029
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
245	1,43106	0,90606	1,80560	-1,70655
246	0,38554	0,93918	1,65935	-1,64986

Pada Tabel 4.5 memberikan informasi bahwa terdapat empat variabel baru yaitu nilai skor komponen utama (W) yang didapatkan dari perkalian antara matrix nilai baku Z (data

standarisasi) dengan matrix *eigenvectors*, yang akan kemudian di regresikan dengan variabel respon yaitu kuat tekan semen pada PCC. Berikut adalah hasil regresi dengan variabel yang terbentuk.

Tabel 4.6 Uji Parsial

Prediktor	Koefisien	SE Koef	T	P	VIF
Konstan	312,276	1,445	216,18	0,000	
K ₁	-3,1877	0,777	-4,10	0,000	1,000
K ₂	-4,060	1,064	-3,82	0,000	1,000
K ₃	-13,295	1,168	-11,39	0,000	1,000
K ₄	-3,062	1,395	-2,20	0,029	1,000

Berdasarkan *output* yang disajikan di Tabel 4.6 terlihat bahwa nilai VIF disetiap variabel baru memiliki nilai yang kurang dari 10, sehingga fenomena multikolinearitas telah berhasil diatasi dan pada pengujian parsial dari variabel K₁ hingga K₄ telah berpengaruh signifikan terhadap model. Dari hasil regresi tersebut dapat dibentuk model regresi sebagai berikut.

$$\hat{Y} = 312,28 - 3,18K_1 - 4,06K_2 - 13,29K_3 - 3,06K_4$$

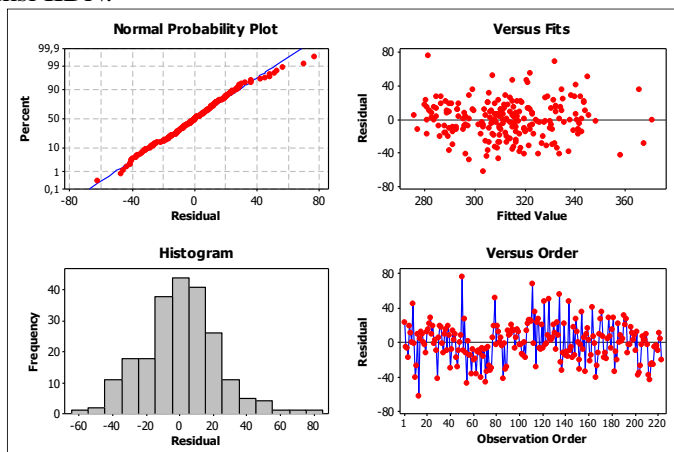
Dengan persamaan regresi linear diatas, dapat dikembalikan ke model awal dengan cara mendistribusikan persamaan (2.3) ke persamaan (2.2) dengan hasil sebagai berikut.

$$\begin{aligned} \hat{Y} = & 312,28 - 3,18(-0,50068Z_1 - 0,50752Z_2 - 0,38385Z_3 + 0,08497Z_4 + 0,04033Z_5 + \dots + 0,00505Z_{11}) \\ & - 4,06(0,02702Z_1 + 0,09249Z_2 - 0,09894Z_3 - 0,50530Z_4 - 0,09476Z_5 + \dots - 0,30523Z_{11}) \\ & - 13,29(0,04558Z_1 + 0,08563Z_2 + 0,11251Z_3 + 0,07712Z_4 + 0,53391Z_5 + \dots + 0,01775Z_{11}) \\ & - 3,06(0,03216Z_1 - 0,02254Z_2 + 0,07979Z_3 + 0,57451Z_4 - 0,45016Z_5 + \dots - 0,34435Z_{11}) \end{aligned}$$

$$\hat{Y} = 312,28 + 0,782x_1 + 0,173x_2 - 0,115x_3 - 1,004x_4 - 5,464x_5 + 6,643x_6 - 2,088x_7 - 0,640x_8 - 5,850x_9 - 9,694x_{10} + 2,042x_{11}$$

Hasil *output* juga memberikan keputusan bahwa setiap variabel independen berpengaruh signifikan terhadap model dan memberikan nilai *Rsquared* yaitu sebesar 40,8% yang berarti variabel independen dapat menjelaskan variabel dependen sebesar 40,8%, sisanya dijelaskan di luar variabel yang digunakan. Kemudian setelah didapatkan model dan hasil regresinya, akan dilakukan pendeteksian asumsi regresi linear, asumsinya adalah berdistribusi normal, identik dan independen. Pendeteksian dapat dilakukan secara visual dan inferensia, pada penelitian kali ini,

akan dilakukan secara visual. Berikut adalah hasil pendeteksian asumsi IIDN.



Gambar 4.2 Asumsi IIDN Data Semen PCC

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 (berwarna biru) 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.

4.4.2 Validasi Model PCR

Validasi model PCR kali ini menggunakan metode *10 fold cross validation*, yang berarti subset data akan terbagi menjadi *10 fold* yang nantinya *9 fold* menjadi subset data *training* dan *1 fold* menjadi subset data *testing*. Validasi menggunakan kriteria terbaik RMSE yang akan dicari rerata dari kesepuluh *fold*

Tabel 4.7 Validasi Model PCR

<i>Fold</i>	RMSE
<i>Fold 1</i>	22,58
<i>Fold 2</i>	22,17
<i>Fold 3</i>	20,06
<i>Fold 4</i>	22,09
<i>Fold 5</i>	24,38
<i>Fold 6</i>	16,62
<i>Fold 7</i>	22,71
<i>Fold 8</i>	23,16
<i>Fold 9</i>	26,78
<i>Fold 10</i>	23,39
Rata-Rata	22,39

Tabel 4.7 menunjukkan hasil perhitungan dari kriteria model terbaik yaitu RMSE menggunakan metode PCR. Untuk *fold* yang menghasilkan RMSE terendah adalah *fold 6* dengan nilai RMSE sebesar 16,62. Akan tetapi, hasil dari validasi model tidak diperbolehkan hanya melihat dari salah satu *fold*, oleh karena itu digunakan indikator rata-rata dari kesepuluh *fold* yang memiliki nilai RMSE sebesar 22,39, yang kemudian nilai ini akan digunakan sebagai pembandingan dengan metode lainnya.

Setelah melakukan pemodelan dan validasi menggunakan metode PCR, kemudian dilanjutkan dengan pemodelan menggunakan salah satu teknik *machine learning* yaitu SVR, dengan *kernel-linear*, *kernel-polynomial* dan *kernel-RBF*

4.5 Support Vector Regression Data Semen PCC

SVR merupakan salah satu teknik *machine learning* yang berbasis SVM yang diterapkan ke dalam kasus-kasus regresi. Model SVR akan dibentuk menggunakan keseluruhan data dengan *complete feature* dan *feature selection* dengan RFE, lalu validasi menggunakan 10 *fold cros validation* untuk mendapatkan nilai kriteria model terbaik yaitu RMSE masing-masing model. Metode

yang digunakan dalam SVR kali ini adalah *Kernel-Linear*, *Kernel-Polynomial* dan *Kernel-Radial Basis Function*.

4.5.1 SVR Dengan *Complete Feature*

Dalam penelitian ini dilakukan pemodelan menggunakan *Support Vector Regression* dengan *complete feature* yang terdiri dari 11 variabel prediktor dan 1 variabel respon yaitu kuat tekan semen PCC.

4.5.1.1 *Modelling SVR*

Pembentukan model SVR disertai parameter *tuning* menggunakan metode *grid seacrh* untuk mendapatkan prediksi parameter yang paling optimum. Untuk pemodelan kali ini menggunakan semua data kuat tekan semen PCC dari bulan Juli 2018 hingga Maret 2019 dengan 11 variabel, pemodelan meliputi *Kernel-Linear*, *Kernel-Polynomial* dan *Kernel-Radial Basis Function* (RBF).

a). *Kernel Linear*

Kernel-Linear merupakan sebuah metode yang ketika digunakan, data akan terpisah oleh sebuah garis linier yang disebut *hyperlane*. Berikut adalah hasil penggunaan *kernel-linear* dengan seluruh data kuat tekan semen PCC.

Tabel 4.8 *Tuning Parameter Kernel-Linear*

<i>Kernel</i>	C	RMSE
<i>Linear</i>	1	23,51

Tabel diatas menunjukkan hasil *tuning* parameter menggunakan *kernel-linear* yang hasilnya mendapatkan nilai *parameter cost* (C) yaitu senilai 1, dengan nilai kriteria RMSE yaitu sebesar 23,51. Hal ini menyatakan bahwa dengan metode *kernel-linear complete feature*, model memiliki nilai akurasi prediksi sebesar 23.51, yang kemudian nilai ini akan dibandingkan dengan metode lainnya.

b). *Kernel Polynomial*

Kernel-polynomial merupakan *kernel* yang bersifat *non-linear*. *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 4.9 *Tuning Parameter Kernel-Polynomial*

<i>Degree</i>	<i>Gamma</i>	<i>C</i>	RMSE
1	0,001	0,25	28,64327
1	0,001	0,5	27,8415
1	0,001	1	26,60177
1	0,01	0,25	24,69781
1	0,01	0,5	23,76333
1	0,01	1	23,30962
1	0,1	0,25	23,1993
1	0,1	0,5	23,29292
1	0,1	1	23,41949
2	0,001	0,25	27,83821
2	0,001	0,5	26,59642
2	0,001	1	25,08911
2	0,01	0,25	23,6739
2	0,01	0,5	23,1744
2	0,01	1	22,9691
2	0,1	0,25	23,50165
2	0,1	0,5	24,22072
2	0,1	1	25,44794
3	0,001	0,25	27,15214
3	0,001	0,5	25,66091
3	0,001	1	24,38279
3	0,01	0,25	23,28626
3	0,01	0,5	22,95288
3	0,01	1	22,91489
3	0,1	0,25	26,5429
3	0,1	0,5	29,68679
3	0,1	1	34,79056

Tabel 4.9 merupakan *tuning parameter* untuk *kernel polynomial*, didapatkan hasil bahwa terdapat dua *parameter* yang digunakan yaitu *gamma* dan *cost* (C). Dengan kriteria terbaik yaitu RMSE sebesar 22,914, didapatkan *parameter* yang paling optimal adalah 0,01 (*gamma*) dan 1 (*cost*) dengan nilai *degree* bernilai tiga. Dengan *kernel-polynomial*, akurasi prediksi menggunakan model dan parameternya memiliki nilai sebesar 22.914 yang kemudian

dilakukan perbandingan terhadap semua model yang telah didapatkan.

c). Kernel Radial Basis Function

Kernel-RBF (Radial Basis Function) merupakan salah satu dari beberapa *kernel* untuk kasus *non-linear*. *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 4.10 *Tuning Parameter Kernel-RBF*

<i>Gamma</i>	C	RMSE
0,030096	0,25	22,58855
0,030096	0,5	22,17093
0,030096	1	22,03268
0,098053	0,25	22,63819
0,098053	0,5	22,33979
0,098053	1	22,49458
0,16601	0,25	23,42606
0,16601	0,5	22,96372
0,16601	1	22,89459

Tabel diatas menunjukkan hasil *tuning parameter* untuk SVR dengan *kernel-RBF* dengan seluruh data kuat tekan semen. Terdapat dua *parameter* untuk *kernel-RBF* yaitu *gamma* dan *cost* (C). Dengan kriteria RMSE sebesar 22,03 maka *parameter* yang terpilih dan paling optimal untuk data kuat tekan semen PCC ini adalah *gamma* bernilai 0,030096 dan nilai *cost* sebesar 1. Dengan model *kernel-RBF* ini, maka akurasi prediksi menggunakan model dan parameter terpilihnya memiliki nilai sebesar 22.03, yang kemudian akan dibandingkan.

Jika dibandingkan dengan model *kernel-linear* dan *kernel-polynomial*, maka *kernel-RBF* yang memiliki model terbaik dengan kriteria RMSE yang paling rendah. Selanjutnya, dilakukan evaluasi model dengan *10 fold cross validation* yang membagi data menjadi data *training* dan *testing*.

4.5.1.2 Validasi Model SVR

Validasi dalam kali ini menggunakan metode *10 fold cross validation*. Sebanyak 10 *fold* yang terbagi menjadi data *training* dan data *testing* akan dilakukan pemodelan dengan masing-masing *kernel* yang digunakan dan *parameter* yang telah optimal, adapun *parameter*nya adalah sebagai berikut.

Tabel 4.11 *Parameter Kernel*

<i>Kernel</i>	<i>Parameter</i>		
	<i>Gamma</i>	<i>Degree</i>	<i>Cost</i>
<i>Linear</i>	-	-	1
<i>Polynomial</i>	0,01	3	1
RBF	0,030096	-	1

Dengan *parameter* yang berada di Tabel 4.11 akan dilakukan pemodelan di setiap *fold*, yang kemudian akan diprediksi menggunakan data *testing* dan akan dicari nilai kriteria RMSE, untuk model di setiap *fold*, yang kemudian hasil dari seluruh *fold* akan dirata-rata. Berikut adalah hasil pemodelannya.

Tabel 4.12 *Validasi Model SVR*

<i>Fold</i>	<i>Kernel</i>		
	RBF	<i>Linear</i>	<i>Polynomial</i>
<i>Fold 1</i>	22,75	23,74	23,32
<i>Fold 2</i>	19,83	20,79	21,06
<i>Fold 3</i>	20,29	21,71	18,98
<i>Fold 4</i>	22,03	21,85	22,09
<i>Fold 5</i>	25,29	25,1	25,40
<i>Fold 6</i>	16,58	16,69	17,06
<i>Fold 7</i>	22,20	22,95	22,36
<i>Fold 8</i>	20,83	22,72	21,72
<i>Fold 9</i>	29,14	29,68	29,35
<i>Fold 10</i>	23,05	23,79	23,5

Tabel 4.12 memberikan informasi bahwa pemodelan yang dilakukan dengan *parameter* yang sama di tiap *fold* menghasilkan hasil yang berbeda, yang menyebabkan perbedaan adalah pembagian data yang tidak boleh ada pengulangan. Terlihat dari *fold* 6 memberikan hasil yang terbaik karena dalam pemodelan *kernel-linear*, *kernel-polynomial* dan *kernel-RBF* memiliki nilai RMSE yang terendah diantara *fold* yang lain, sedangkan dari *fold* yang kurang baik hasilnya adalah pada *fold* 9 yang memiliki RMSE yang tinggi. Karena setiap *fold* mengandung informasi dari kuat tekan semen PCC, maka validasi model menggunakan rerata dari setiap hasil yang telah dilakukan pemodelan, berikut adalah hasilnya.

Tabel 4.13 Rata-Rata Kriteria Setiap Model *Kernel*

Indikator	Kriteria	<i>Kernel</i>		
		RBF	<i>Linear</i>	<i>Polynomial</i>
Rata-Rata	RMSE	22,19	22,90	22,48

Tabel 4.13 diatas memperlihatkan bahwa, jika digunakan indikator rata-rata dalam penentuan nilai kriteria terbaik akan menurunkan performa dari masing-masing fungsi *kernel*. Hal ini disebabkan variansi nilai kriteria dari tiap *fold*. Dapat diketahui bahwa model yang terbaik untuk digunakan adalah *kernel-RBF* kriteria RMSE yang terendah dibandingkan *kernel* lainnya yaitu sebesar 22,19.

Setelah dilakukan pemodelan serta validasi untuk SVR *Complete Feature*, kemudian dilanjutkan dengan pemodelan menggunakan SVR *Feature Selection*. Seleksi fitur yang digunakan adalah *Recursive Feature Elimination* (RFE).

4.5.2 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*. Variabel yang direduksi adalah variabel prediktor, yaitu variabel yang menunjang adanya kuat tekan semen produk PCC. *Feature selection* yang

digunakan adalah RFE yaitu *Recursive Feature Elimination* berbasis *Random Forest* (RF-RFE).

4.5.2.1 *Random Forest-Recursive Feature Elimination*

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 4.14 *Ranking* RF-RFE

<i>Ranking</i>	Variabel	RMSE
1	Blaine	28,00
2	SO₃	25,82
3	LOI	23,16
4	MgO	22,87
5	SiO₂	22,42
6	Fe ₂ O ₃	22,30
7	Al ₂ O ₃	21,90
8	Cao	21,72
9	Insol	21,81
10	Residu	21,78
11	FL	21,59

Berdasarkan tabel diatas, menunjukkan bahwa berdasarkan *ranking* RF-RFE, variabel Blaine, SO₃, LOI, MgO dan SiO₂ adalah merupakan variabel dengan *ranking* tertinggi yang menurut metode RF-RFE kurang relevan, sehingga kelima variabel tersebut akan dihilangkan (tidak digunakan).

4.5.2.2 *Modelling SVR Dengan Feature Selection*

Modelling merupakan suatu pembentukan model dengan cara *tuning* parameter untuk mendapatkan *parameter* paling optimal, dengan menggunakan tiga fungsi *kernel* yaitu *linear*, *polynomial* dan RBF. Berikut adalah hasil dari pemodelan SVR dengan *feature selection* (RF-RFE).

a). Kernel Linear

Kernel-Linear merupakan sebuah metode yang ketika digunakan, data akan terpisah oleh sebuah garis linier yang disebut *hyperlane*. Berikut adalah hasil penggunaan *kernel-linear* dengan *feature selection*.

Tabel 4.15 *Tuning Parameter Kernel-Linear* RF-RFE

<i>Kernel</i>	<i>C</i>	RMSE
<i>Linear</i>	1	29,881

Tabel 4.15 menunjukkan hasil *tuning* parameter menggunakan RF-RFE dengan model *kernel-linear*. Dengan nilai *parameter cost* (*C*) yaitu senilai 1 didapatkan kriteria RMSE yang menjelaskan bahwa akurasi prediksi untuk *kernel-Linear feature selection* sebesar 29,881. Kemudian kriteria terbaik RMSE ini dibandingkan dengan *kernel lainnya* yang telah terseleksi fitur.

b). Kernel Polynomial

Kernel-polynomial merupakan *kernel* yang bersifat *non-linear*. *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* dengan *feature selection*.

Tabel 4.16 *Tuning Parameter Kernel-Polynomial* RF-RFE

<i>Degree</i>	<i>Gamma</i>	<i>C</i>	RMSE
1	0,001	0,25	29,7996
1	0,001	0,5	29,82194
1	0,001	1	29,8526
1	0,01	0,25	29,9418
1	0,01	0,5	30,03972
1	0,01	1	30,13421
1	0,1	0,25	30,25217
1	0,1	0,5	30,32718
1	0,1	1	30,3524
2	0,001	0,25	29,82154
2	0,001	0,5	29,85231
2	0,001	1	29,92208
2	0,01	0,25	30,01467

Tabel 4.16 *Tuning Parameter Kernel-Polynomial RF-RFE(Lanjutan)*

<i>Degree</i>	<i>Gamma</i>	<i>C</i>	RMSE
2	0,01	0,5	30,08402
2	0,01	1	30,13817
2	0,1	0,25	31,59392
2	0,1	0,5	32,95569
2	0,1	1	34,5827
3	0,001	0,25	29,83633
3	0,001	0,5	29,89195
3	0,001	1	29,95689
3	0,01	0,25	30,04116
3	0,01	0,5	30,10939
3	0,01	1	30,1058
3	0,1	0,25	32,96752
3	0,1	0,5	34,39058
3	0,1	1	36,57519

Tabel 4.16 merupakan hasil dari *tuning parameter* untuk *kernel polynomial* dengan RF-RFE, terdapat dua *parameter* yang digunakan yaitu *gamma* dan *cost* (C). Dengan kriteria terbaik yaitu RMSE sebesar 29,79 didapatkan *parameter* yang paling optimal untuk *kernel-Polynomial feature selection* adalah 0,001 (*gamma*) dan 0,25 (*cost*) dengan nilai *degree* adalah 3.

c). Kernel Radial Basis Function

Kernel-RBF (Radial Basis Function) merupakan salah satu dari beberapa *kernel* yang bersifat *non-linear*. *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 RBF dengan *feature selection*.

Tabel 4.17 *Tuning Parameter Kernel-RBF RF-RFE*

<i>Gamma</i>	<i>C</i>	RMSE
0,046757	0,25	28,41493
0,046757	0,5	28,33156
0,046757	1	28,17513
0,235164	0,25	27,49736
0,235164	0,5	27,32093
0,235164	1	27,43508

Tabel 4.17 *Tuning Parameter Kernel-RBF RF-RFE* (Lanjutan)

<i>Gamma</i>	C	RMSE
0,423572	0,25	27,46448
0,423572	0,5	27,26652
0,423572	1	27,47271

Tabel 4.17 diatas menunjukkan hasil *tuning parameter* untuk *kernel-RBF* dengan RF-RFE. Hasil menunjukkan bahwa terdapat dua *parameter* untuk *kernel-RBF* yaitu *gamma* dan *cost* (C). Maka *parameter* yang terpilih adalah *gamma* bernilai 0,423572 dan nilai *cost* sebesar 0,25, dengan nilai akurasi prediksi RMSE adalah sebesar 27.2665

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

4.6 Pemilihan Model Terbaik Data Semen PCC

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

Tabel 4.18 Pemilihan Model Terbaik

No.	Metode		RMSE
1	PCR		22,39
2	SVR -	Linear	22,90
	<i>Complete Feature</i>	Polynomial	22,48
3		RBF	22,19
	SVR -	Linear	29,88
	<i>Feature Selection</i>	Polynomial	29,79
		RBF	27,27

Tabel 4.18 memberikan informasi bahwa, dari beberapa model yang terbentuk, kriteria model terbaik adalah pada metode *Support Vector Regression Kernel-Radial Basis Function* dengan *Complete Feature* yang memiliki nilai kriteria RMSE bernilai 22,19, dengan parameter γ bernilai 0,030096 dan nilai $cost$ sebesar 1. Berikut adalah model *Kernel Radial Basis Function*.

$$K(x, x') = \exp(-0,030096\|x - x_i\|^2) + \varepsilon$$

Dengan model *kernel-RBF*, kemudian dilakukan pemodelan ke *Support Vector Regression*.

$$f(x) = w^T K(x, x') + b + \varepsilon$$

$$\hat{f}(x) = \hat{w}_i^T \exp(-0,030096\|x - x_i\|^2) + \hat{b}$$

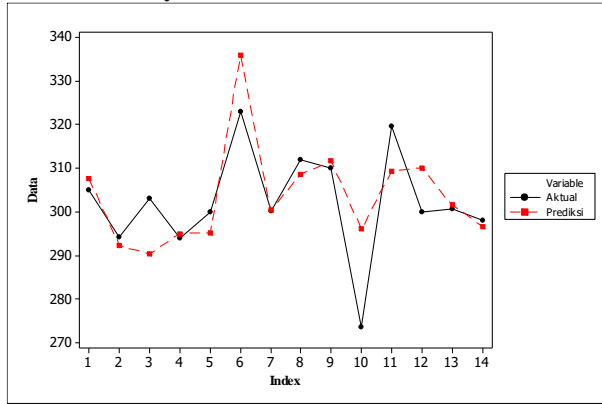
Tabel 4.19 Koefisien Parameter Model SVR

Variabel	\hat{w}_i^T
x ₁	3,427
x ₂	3,266
x ₃	-2,751
x ₄	0,173
x ₅	-6,875
x ₆	7,776
x ₇	-5,090
x ₈	-3,249
x ₉	-8,365
x ₁₀	-9,747
x ₁₁	1,073

Tabel 4.19 memberikan informasi bahwa pada pemodelan dengan SVR *kernel-radial basis function* memiliki parameter γ sedangkan untuk pemodelan SVR terdapat koefisien parameter \hat{w}_i^T yang setiap variabel memiliki koefisien masing-masing serta nilai \hat{b} yaitu bernilai -0,165, nilai \hat{b} merupakan nilai *intercept*. Setelah mendapatkan model terbaik, dilakukan prediksi kuat tekan semen produk PCC dengan *complete feature* metode SVR *kernel-radial basis function*.

4.7 Prediksi Kuat Tekan Semen PCC

Dengan model *kernel-Radial Basis Function Complete Feature* yang terpilih sebagai metode terbaik dalam penelitian ini disertai *parameter* pada tabel 4.11, akan dilakukan prediksi kuat tekan semen PCC pada hari ke-28 pada bulan April 2019 di *Mill 8*, berikut adalah hasilnya.



Gambar 4.3 Prediksi Kuat Tekan Semen PCC April 2019

Gambar 4.3 memberikan informasi bahwa, terdapat 14 data kuat tekan semen PCC di bulan April 2019 di *Mill 8*. Secara visual nilai prediksi dengan model SVR *kernel-RBF* telah mengikuti data aktual, namun pada observasi keenam dan kesepuluh yaitu tanggal 6 April 2019 dan 12 April 2019 cukup memiliki perbedaan yang jauh pada besaran kuat tekan semen PCC, namun untuk observasi yang lain, telah mengikuti data aktual, dengan prediksi ini didapatkan nilai kesalahan RMSE adalah 8,78. Nilai kesalahan RMSE ini lebih rendah dibandingkan dengan saat pemodelan PCR, SVR *Complete Feature* dan SVR *Feature Selection*.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Pada hasil analisis regresi linier terindikasi mengalami multikolinearitas, sehingga perlu ditangani menggunakan metode PCR yang mereduksi variabel menjadi empat komponen. Kemudian dilakukan pemodelan dengan metode 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 kuat tekan semen. SVR *complete feature* juga menghasilkan model dan *parameter* yang lebih baik dibandingkan PCR. Metode yang terpilih adalah *Kernel – Radial Basis Function* dengan *complete feature*.
2. Hasil prediksi menunjukkan bahwa dengan model *Kernel-RBF* telah memprediksi kuat tekan semen PCC di bulan April 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.

5.2 Saran

Berdasarkan kesimpulan, maka saran yang dapat diberikan untuk penelitian selanjutnya adalah penggunaan metode yang tepat dan sebanding menurut data maupun metode dapat memberikan hasil analisis yang lebih baik. Selain itu, peneliti dapat mengambil informasi lebih banyak dalam pengaruh kuat tekan semen, karena variabel-variabel yang digunakan dalam penelitian ini tidak sepenuhnya menggambarkan kuat tekan semen secara keseluruhan, diduga variabel-variabel seperti

manusia, *setting* mesin, cuaca dan sebagainya juga turut mempengaruhi kuat tekan semen dan diharapkan pengambilan data dengan periode yang lebih panjang, agar pemodelan dan prediksi lebih akurat.

DAFTAR PUSTAKA

- Abe, S. 2005. *Support Vector Machine for Pattern Classification*. Springer - Verlag. London Limited.
- Alfarisi. *Data Preprocessing - Konsep Pembelajaran Data Mining*. Steetmit.com. [Dikutip: 9 Desember 2018.] <https://steetmit.com/education/@alfarisi/data-reprocessing-konsep-pembelajaran-datamining?sort=new#comments>, 2017.
- Breiman, L. 2001. *Random Forests*. *Machine Learning*. Vol. 45, hal. 5-32.
- ASTM Standards. 2002. ASTM C 109/C 109M – 02. *Standard Test Method for Compressive Strength of Hydraulic Cement Mortars (Using 2-in. or 50- mm Cube Specimens)*. ASTM International, West Conshohocken, PA
- BAPPENAS, BPS dan UNFPA. 2013. *Proyeksi Penduduk Indonesia 2010-2035*. Jakarta: BPS
- Casasent, D. dan Nakariyakul, S. 2004. *Hyperspectral Ratio Feature Selection: Agricultural Product Inspection Example. Nondestructive Sensing for Food Safety, Quality, and Natural Resources 5587, Proceedings of SPIE; Philadelphia, 26 Oktober 2004*. Hal. 133-143.
- Daniel, W.W. 1989. *Statistika Nonparametrik Terapan*. Terjemahan oleh Alex Tri Kuncoro. Jakarta: PT. Gramedia Pustaka Utama.
- Draper, N. R. dan Smith, H. 1992. *Analisis Regresi Terapan Edisi Kedua*. Terjemahan oleh Bambang Sumantri. Jakarta: PT Gramedia Pustaka Utama.
- De Gooijer, J. G. dan Hyndman, R. J. 2006. *25 Years Of Time Series Forecasting. International Journal of Forecasting*. Vol. 22, hal. 443-473.

- Johnson, R. A dan Wichern, D. W. 2010. *Applied Multivariate Statistical Analysis. Sixth Edition*, Prentice Hall. New Jersey.
- Kohavi, R. 1995. *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. Di dalam: *Proceedings of the International Joint Conference on Artificial Intelligence*; 1995. San Mateo, California: Morgan Kaufmann. Hal. 1137-1143.
- Santosa, B. 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Graha Ilmu: Yogyakarta.
- Scholkopf, B. dan Smola, A. J. 2002. *Learning With Kernel*. MIT Press.
- Sudjana. 2005. *Metode Statistika*. Bandung : Tarsito.
- SNI 15-2049-2004 Semen Portland
- Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., dan Feuston, B. P. 2003. *Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling*. *Journal of Chemical Information and Computer Sciences*. Vol. 7, No.3.
- Witten, I. H., Frank, E., dan Hall, M. A. 2011. *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition*. Burlington: Morgan Kaufmann Publishers.

LAMPIRAN

Lampiran 1. Data Kuat Tekan Semen PCC Mill #8.

No. Prod.	ELEMEN MAYOR						LAIN - LAIN					Kuat Tekan
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	Y
182	24.72	7.65	2.72	55.72	2.94	1.38	0.76	5.32	11.78	395	7.13	304
183												
184	23.64	6.65	2.61	60.59	2.78	1.38	0.82	7.06	5.45	310	12.92	337
185	20.87	5.78	2.40	59.17	2.39	1.49	0.76	6.42	7.82	338	8.39	317
186	20.37	5.47	2.38	61.58	2.31	1.33	1.54	7.26	11.69	386	8.57	264
187	22.47	6.25	2.65	60.32	2.20	1.38	0.80	7.03	11.32	380	8.99	313
188	23.06	6.48	2.68	59.26	2.19	1.48	0.66	8.27	5.89	315	10.69	355
189*	23.27	6.45	2.79	59.60	2.28	1.45	0.66	9.39	7.14	330	12.68	334
...
...
...
85*	17.92	4.84	2.34	59.48	1.95	1.41	0.91	5.07	11.57	387	10.64	305
86	18.26	4.99	2.36	58.78	2.68	1.38	0.93	5.38	11.53	384	9.26	299
87	18.33	4.99	2.40	58.39	2.89	1.45	0.89	5.60	11.48	375	9.85	310

*data produk nomor 182-189 adalah bulan Juli 2018 dan seterusnya, data produk nomor 85 adalah data bulan Maret 2019 dan seterusnya.

Y	Kuat Tekan Hari ke-28	X_9	<i>Loss of Ignition</i>
X_1	Oksida Silikon (SiO ₂)	X_{10}	<i>Blaine</i>
X_2	Oksida Aluminium (Al ₂ O ₃)	X_{11}	<i>Residu</i>
X_3	Oksida Besi (Fe ₂ O ₃)		
X_4	Oksida Kalsium (CaO)		
X_5	Magensium Oksida (MgO)		
X_6	Sulfur Trioksida (SO ₃)		
X_7	<i>Freelime</i>		
X_8	<i>Insoluble Residu</i>		

Lampiran 2. 10-Fold Cross Validation.

```
data = read.csv("D:/skripsii/mill8.csv" , sep=";")
N = nrow(data)
folds = 10
Holdout = split(sample(1:N), 1:folds)
```

Lampiran 3. Hubungan Antar Variabel (*Pearson Correlation*).

	y	x1	x2	x3	x4
x1	0,152 0,017				
x2	0,134 0,036	0,940 0,000			
x3	0,084 0,189	0,513 0,000	0,627 0,000		
x4	-0,028 0,659	-0,111 0,081	-0,219 0,001	0,047 0,460	
x5	-0,285 0,000	-0,088 0,171	-0,063 0,329	0,042 0,515	-0,125 0,050
x6	0,304 0,000	-0,391 0,000	-0,376 0,000	-0,311 0,000	-0,264 0,000
x7	-0,122 0,055	0,005 0,939	-0,006 0,919	-0,037 0,562	-0,059 0,355
x8	-0,046 0,476	0,513 0,000	0,566 0,000	0,277 0,000	-0,496 0,000
x9	-0,413 0,000	-0,672 0,000	-0,597 0,000	-0,474 0,000	-0,032 0,614
x10	-0,563 0,000	-0,187 0,003	-0,126 0,048	-0,181 0,004	-0,074 0,250
x11	0,003 0,960	0,004 0,947	-0,015 0,821	-0,029 0,650	0,121 0,057

Lampiran 3. Hubungan Antar Variabel (Lanjutan).

	x6	x7	x8	x9	x10
x7	0,046				
	0,476				
x8	0,050	0,200			
	0,438	0,002			
x9	0,231	-0,001	-0,156		
	0,000	0,988	0,014		
x10	-0,166	0,114	0,175	0,664	
	0,009	0,074	0,006	0,000	
x11	-0,087	-0,109	-0,142	-0,037	-0,153
	0,173	0,089	0,026	0,562	0,016

Lampiran 4. Output Analisis Regresi Linier.

Regression Analysis: y versus x1; x2; ...

The regression equation is

$$y = 463 - 3,24 x_1 + 19,2 x_2 - 13,4 x_3 - 0,11 x_4 - 8,00 x_5 + 48,1 x_6 - 8,21 x_7 - 2,99 x_8 - 3,30 x_9 - 0,404 x_{10} - 0,745 x_{11}$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	463,46	97,61	4,75	0,000	
x1	-3,241	4,001	-0,81	0,419	14,656
x2	19,21	10,79	1,78	0,076	16,098
x3	-13,39	13,49	-0,99	0,322	2,397
x4	-0,107	1,250	-0,09	0,932	2,077
x5	-7,999	3,477	-2,30	0,022	1,448
x6	48,06	11,26	4,27	0,000	2,151
x7	-8,206	5,028	-1,63	0,104	1,101
x8	-2,992	1,556	-1,92	0,056	2,573
x9	-3,304	1,709	-1,93	0,054	4,701
x10	-0,4042	0,1064	-3,80	0,000	3,057
x11	-0,7450	0,8440	-0,88	0,378	1,084

S = 22,4448 R-Sq = 43,5% R-Sq(adj) = 40,9%

PRESS = 130548 R-Sq(pred) = 37,48%

Lampiran 4. Output Analisis Regresi Linier (Lanjutan).

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	11	90941,0	8267,4	16,41	0,000
Residual Error	234	117882,2	503,8		
Total	245	208823,2			

Lampiran 5. Output Eigenvalue.

Principal Component Analysis: x1; x2; x3; x4; x5; x6; x7; x8; x9; x10; x11					
Eigenanalysis of the Correlation Matrix					
Eigenvalue	3,4639	1,8505	1,5367	1,0768	0,9046
	0,8895	0,5644	0,2933		
Proportion	0,315	0,168	0,140	0,098	0,082
	0,081	0,051	0,027		
Cumulative	0,315	0,483	0,623	0,721	0,803
	0,884	0,935	0,962		
Eigenvalue	0,2446	0,1428	0,0329		
Proportion	0,022	0,013	0,003		
Cumulative	0,984	0,997	1,000		

Lampiran 6. Output PCA.

Variable	PC1	PC2	PC3	PC4	PC5	...
x1	-0,501	0,027	0,046	0,032	-0,055	...
x2	-0,508	0,092	0,086	-0,023	-0,021	...
x3	-0,384	-0,099	0,113	0,080	0,165	...
x4	0,085	-0,505	0,077	0,575	-0,002	...
x5	0,040	-0,095	0,534	-0,450	0,077	...
x6	0,213	0,249	-0,585	-0,179	0,040	...
x7	-0,010	0,273	-0,042	0,514	-0,489	...
x8	-0,295	0,520	-0,029	-0,046	-0,139	...
x9	0,418	0,261	0,259	0,002	-0,043	...
x10	0,168	0,385	0,523	0,207	-0,058	...
x11	0,005	-0,305	0,018	-0,344	-0,836	...

Lampiran 6. Output PCA (lanjutan).

Variable	PC9	PC10	PC11
x1	0,360	-0,093	0,669
x2	0,392	-0,241	-0,701
x3	0,009	0,032	0,168
x4	-0,084	-0,031	-0,113
x5	0,036	0,028	-0,033
x6	0,414	0,190	-0,024
x7	0,115	-0,068	-0,001
x8	-0,666	-0,044	-0,029
x9	0,155	-0,736	0,132
x10	0,235	0,583	-0,024
x11	0,011	0,075	0,003

Lampiran 7. Output Regresi PCA vs Kuat Tekan Semen PCC.

Regression Analysis: y versus W1; W2; W3; W4					
The regression equation is					
$y = 312 - 3,19 W1 - 4,06 W2 - 13,3 W3 - 3,06 W4$					
Predictor	Coef	SE Coef	T	P	VIF
Constant	312,276	1,445	216,18	0,000	
W1	-3,1877	0,7777	-4,10	0,000	1,000
W2	-4,060	1,064	-3,82	0,000	1,000
W3	-13,295	1,168	-11,39	0,000	1,000
W4	-3,062	1,395	-2,20	0,029	1,000
S = 22,6565 R-Sq = 40,8% R-Sq(adj) = 39,8%					
Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	4	85114	21278	41,45	0,000
Residual Error	241	123709	513		
Total	245	208823			

Lampiran 8. *Syntax* PCR dengan *Cross Validation*.

```
library(caret)
library(miscTools)
library(e1071)
library(pls)

datafull=read.csv("D:/skripsii/mill8.csv" ,
sep=";")

train1=read.csv("D:/skripsii/train1.csv" , sep=";")
test1=read.csv("D:/skripsii/test1.csv" , sep=";")

#----FOLD 1----#
train1 <- train1
y_test1 <- test1[, -1]
test1 <- test1
pcr_model1 <- pcr(y~., data = train1 ,scale =TRUE,
validation = "CV")
pcr_pred1 <- predict(pcr_model, test1, ncomp = 4)
rmse1=RMSE(test1$y, pcr_pred1);rmse1
mae1=MAE(test1$y, pcr_pred1);mae1
y1=as.vector(test1[,1])
resid1=as.matrix(y1-pcr_pred1)
rsq1=rSquared(y1, resid1);rsq1

...
...
...

#----FOLD 10----#
```


Lampiran 9. Syntax Support Vector Regression.

```
library(caret)
library(miscTools)

datafull=read.csv("D:/skripsii/mill8.csv" ,
sep=";")

modellrb <- train(y~., data=datafull,
method="svmRadialSigma")
print(modellrb)
predictionlrb <- predict (modellrb, datafull)
predlrb=data.frame(predictionlrb)
rmse1=RMSE(datafull$y, predictionlrb);rmse1
mael=MAE(datafull$y, predictionlrb);mael
ylrb=as.vector(datafull[,1])
residlrb=y1rb-predictionlrb
rsqlrb=rSquared(ylrb, residlrb);rsqlrb

modelllin <- train(y~., data=datafull,
method="svmLinear")
print(modelllin)
predictionllin <- predict (modelllin, datafull)
predllin=data.frame(predictionllin)
rmse1lin=RMSE(datafull$y, predictionllin);rmse1lin
mael1lin=MAE(datafull$y, predictionllin);mael1lin
yllin=as.vector(datafull[,1])
residllin=y1lin-predictionllin
rsq1lin=rSquared(y1lin, residllin);rsq1lin

modellpol <- train(y~., data=datafull,
method="svmPoly")
print(modellpol)
prediction1pol <- predict (modellpol, datafull)
pred1pol=data.frame(prediction1pol)
rmse1pol=RMSE(datafull$y, prediction1pol);rmse1pol
mael1pol=MAE(datafull$y, prediction1pol);mael1pol
y1pol=as.vector(datafull[,1])
resid1pol=y1pol-prediction1pol
rsq1pol=rSquared(y1pol, resid1pol);rsq1pol
```

Lampiran 10. SVR – Complete Feature dengan Parameter.

```
library(caret)
library(miscTools)
library(e1071)

datafull=read.csv("D:/skripsii/mill8.csv" ,
sep=";")

train1=read.csv("D:/skripsii/train1.csv" , sep=";")
test1=read.csv("D:/skripsii/test1.csv" , sep=";")

modellrb <- svm(y~., data=train1,
method="svmRadialSigma", gamma=0.03009583 , cost=1)
print(modellrb)
prediction1rb <- predict (modellrb, test1[-1])
pred1rb=data.frame(prediction1rb)
rmse1=RMSE(test1$y, prediction1rb);rmse1
mae1=MAE(test1$y, prediction1rb);mae1
y1rb=as.vector(test1[,1])
resid1rb=y1rb-prediction1rb
rsq1rb=rSquared(y1rb, resid1rb);rsq1rb

modelllin <- train(y~., data=train1,
method="svmLinear")
print(modelllin)
prediction1lin <- predict (modelllin, test1[-1])
pred1lin=data.frame(prediction1lin)
rmse1lin=RMSE(test1$y, prediction1lin);rmse1lin
mae1lin=MAE(test1$y, prediction1lin);mae1lin
y1lin=as.vector(test1[,1])
resid1lin=y1lin-prediction1lin
rsq1lin=rSquared(y1lin, resid1lin);rsq1lin

modellpol <- train(y~., data=train1,
method="svmPoly")
print(modellpol)
prediction1pol <- predict (modellpol, test1[-1])
pred1pol=data.frame(prediction1pol)
rmse1pol=RMSE(test1$y, prediction1pol);rmse1pol
mae1pol=MAE(test1$y, prediction1pol);mae1pol
y1pol=as.vector(test1[,1])
```

Lampiran 11. SVR – Complete Feature dengan Parameter Setiap Fold.

```
library(caret)
library(miscTools)
library(e1071)

datafull=read.csv("D:/skripsii/mill8.csv" ,
sep=";")

train1=read.csv("D:/skripsii/train1.csv" , sep=";")
test1=read.csv("D:/skripsii/test1.csv" , sep=";")

modellrb <- svm(y~., data=train1,
method="svmRadialSigma", gamma=0.03009583 , cost=1)
print(modellrb)
predictionlrb <- predict (modellrb, test1[-1])
predlrb=data.frame(predictionlrb)
rmse1=RMSE(test1$y, predictionlrb);rmse1
mae1=MAE(test1$y, predictionlrb);mae1
ylrb=as.vector(test1[,1])
residlrb=y1rb-predictionlrb
rsqlrb=rSquared(y1rb, residlrb);rsqlrb

modelllin <- train(y~., data=train1,
method="svmLinear")
print(modelllin)
predictionllin <- predict (modelllin, test1[-1])
predllin=data.frame(predictionllin)
rmse1lin=RMSE(test1$y, predictionllin);rmse1lin
mae1lin=MAE(test1$y, predictionllin);mae1lin
yllin=as.vector(test1[,1])
residllin=y1lin-predictionllin
rsqllin=rSquared(y1lin, residllin);rsqllin

modellpol <- train(y~., data=train1,
method="svmPoly")
print(modellpol)
predictionlpol <- predict (modellpol, test1[-1])
predlpol=data.frame(predictionlpol)
rmse1pol=RMSE(test1$y, predictionlpol);rmse1pol
mae1pol=MAE(test1$y, predictionlpol);mae1pol
ylpol=as.vector(test1[,1])
```

Lampiran 12. Syntax RF-Recursive Feature Elimination.

```
library(caret)

datamodelclass=read.csv("D:/skripsii/mill8.csv" ,
sep=";" , header=TRUE)

set.seed(99)

control <- rfeControl(functions = rfFuncs, method
="cv",number=10)

results <-
rfe(datamodelclass[,2:12],datamodelclass[,1],sizes
= c(1:12),rfeControl = control)

print(results)
```

Lampiran 13. Output RF-Recursive Feature Elimination.

```
Recursive feature selection

Outer resampling method: Cross-Validated (10 fold)

Resampling performance over subset size:
```

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD	Selected
1	27.68	0.2108	21.83	2.504	0.07812	2.585	
2	26.27	0.2361	20.72	2.802	0.14240	3.130	
3	23.12	0.3920	18.26	3.008	0.14763	2.595	
4	22.61	0.4195	17.84	2.746	0.15809	2.272	
5	22.40	0.4324	17.68	2.491	0.15089	2.137	
6	22.41	0.4317	17.60	2.305	0.14177	2.051	
7	22.21	0.4475	17.48	2.146	0.14515	1.904	
8	21.71	0.4749	17.06	2.191	0.14701	1.890	
9	21.90	0.4602	17.05	2.406	0.15467	2.136	
10	21.92	0.4627	17.16	2.498	0.15486	1.888	
11	21.60	0.4901	16.89	2.639	0.16932	2.060	*

```
The top 5 variables (out of 11):
  blaine, so3, loi, mgo, sio2
```

Lampiran 14. SVR – Feature Selection (RF-RFE).

```
library(caret)
library(miscTools)
library(e1071)

data=read.csv("D:/rfe.csv" , sep=";")

modellrb <- train(y~., data=data,
method="svmRadialSigma")
print(modellrb)
predictionlrb <- predict (modellrb, test1[-1])
predlrb=data.frame(predictionlrb)
rmse1=RMSE(data$y, predictionlrb);rmse1
mael1=MAE(data$y, predictionlrb);mael1
ylrb=as.vector(data[,1])
residlrb=y1rb-predictionlrb
rsqlrb=rSquared(y1rb, residlrb);rsqlrb

modelllin <- train(y~., data=data,
method="svmLinear")
print(modelllin)
predictionllin <- predict (modelllin, data[-1])
predllin=data.frame(predictionllin)
rmse1lin=RMSE(data$y, predictionllin);rmse1lin
mael1lin=MAE(data$y, predictionllin);mael1lin
yllin=as.vector(data[,1])
residllin=y1lin-predictionllin
rsqllin=rSquared(y1lin, residllin);rsqllin

modellpol <- train(y~., data=data, method="svmPoly")
print(modellpol)
prediction1pol <- predict (modellpol, data[-1])
pred1pol=data.frame(prediction1pol)
rmse1pol=RMSE(data$y, prediction1pol);rmse1pol
mael1pol=MAE(data$y, prediction1pol);mael1pol
y1pol=as.vector(data[,1])
resid1pol=y1pol-prediction1pol
rsq1pol=rSquared(y1pol, resid1pol);rsq1pol
```



Lampiran 15. Prediksi Kuat Tekan Semen Bulan April 2019
menggunakan model *Kernel-Radial Basis Function*
Complete Feature.

```
library(caret)
library(miscTools)
library(e1071)

train1=read.csv("D:/skripsii/mill8.csv" , sep=";")
test1=read.csv("D:/skripsii/april.csv" , sep=";")

modellrb <- svm(y~., data=train1,
method="svmRadialSigma", gamma=0.03009583 , cost=1)
print(modellrb)
prediction1rb <- predict (modellrb, test1[-1])
pred1rb=data.frame(prediction1rb)
rmse1=RMSE(test1$y, prediction1rb);rmse1
mae1=MAE(test1$y, prediction1rb);mae1
y1rb=as.vector(test1[,1])
resid1rb=y1rb-prediction1rb
rsq1rb=rSquared(y1rb, resid1rb);rsq1rb
```

Lampiran 16. Surat Permohonan Ijin Memperoleh Data.

	<p>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</p>
Nomor : 14700/IT2.V1.9.2 /TU.00.09/2019	19 Februari 2019
Perihal : Permohonan ijin memperoleh data untuk Tugas Akhir	
Yth : Senior Manager of Training & Development PT. Semen Indonesia (Persero) ,Tbk. Jl. Veteran, Sidokumpul. Gresik.	
<p>Dalam rangka menyelesaikan studi di Departemen Statistika FMKSD – ITS, mahasiswa diwajibkan untuk melakukan Tugas Akhir. Sehubungan dengan hal tersebut, kami mohon bantuan kepada Senior Manager of Training & Development PT. Semen Indonesia (Persero) ,Tbk, agar mahasiswa berikut:</p> <p>Nama : RIZKI FEBRIASTO NRP : 06211745000020 (1317 105 020) Program Studi : Sarjana (S1-Lintas Jalur)</p> <p>diperkenankan memperoleh data, untuk keperluan pelaksanaan Tugas Akhir di Instansi yang Bapak/Ibu pimpin. Dengan judul Tugas Akhir mahasiswa tersebut adalah : "Analisis Pengendalian Statistik Pada Produksi Portland Composite Cement". Pelaksanaan dari kegiatan pengambilan data tersebut dilaksanakan sekitar Tanggal 25 Februari 2019 s.d 22 Maret 2019.</p> <p>Adapun data yang dibutuhkan adalah data produksi Portland Composite Cement (PCC), Portland Pozzolan Cement (PPC) dan Ordinary Portland Cement (OPC) periode Januari 2018 s.d Januari 2019</p> <p>Demikian atas ijin yang diberikan kami sampaikan terima kasih.</p>	
<p style="text-align: right;"> Kepala Departemen, Sekretaris Departemen Kartika Fithriyani, M.Si. NIP. 19691212 199303 2 002</p>	

Lampiran 17. Surat Keaslian Data.

SURAT KETERANGAN

Saya yang bertanda tangan di bawah ini menerangkan bahwa :

1. Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

Nama : RIZKI FEBRIASTO

NRP : 062117 4500 0020

Telah mengambil data di instansi/perusahaan kami :

Nama Instansi : PT. SEMEN INDONESIA (Persero) Tbk.

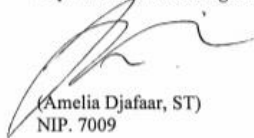
Pabrik Tuban.

Divisi/ bagian : *Quality Assurance*

sejak tanggal 18 Maret 2019 sampai dengan 29 Maret 2019 untuk keperluan Tugas Akhir/ Thesis Semester Gasal/Genap*2018/ 2019.

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.

Tuban, Juni 2019
Dept. Production Management



(Amelia Djafaar, ST)
NIP. 7009

*(coret yang tidak perlu)

Lampiran 18. Surat Penerimaan Penelitian.



F/26203200/004-3

Nomor : 0002439/SM.15/SUP/50032582/2000/03.2019
Lamp. : -
Perihal : **Panggilan Kerja Praktek**

Kepada Yth.

Dr. Kartika Fithriasari, M.Si.

Sekretaris Dept. Statistika - Fak. Matematika, Komputasi dan Sains Data
Institut Teknologi Sepuluh Nopember Surabaya

Menunjuk Surat Saudara No : 14700/IT2.VI.9.2/TU.00.9/2019 tanggal 19 Februari 2019,
Perihal : Permohonan Ijin Kerja Praktek, dengan ini kami beritahukan bahwa kami
tidak dapat menerima mahasiswa/siswa Saudara :

No.	NAMA	NIM	JURUSAN
1.	Rizki Febriasto	06211745000020	Matematika, Komputasi dan Sains Data

Untuk melakukan Kerja Praktek di PT Semen Indonesia (Persero) Tbk.,
di Unit of Tuban & Gresik Quality Assurance, Pabrik Tuban dengan ketentuan sbb :

1. Setiap mahasiswa/siswa yang melakukan Kerja Praktek harus diikutsertakan dalam Asuransi Kecelakaan oleh Institusi ybs.
2. Kerja Praktek dilaksanakan mulai tanggal 18 Maret 2019 s.d. 29 Maret 2019
3. Perusahaan tidak menyediakan sarana akomodasi (penginapan) & transportasi.
4. Mahasiswa/siswa tersebut di atas diharapkan kehadirannya pada :
 - Hari/Tanggal : Senin, 18 Maret 2019
 - Pukul : 07.30 WIB sd. Selesai
 - Tempat : Gedung Auditorium Lt.2 Kantor Pusat PT Semen Gresik Tuban Desa Sumberarum, Kecamatan Kerek, Kabupaten Tuban
 - Acara : Pengarahan dari Perusahaan & Penyerahan Perleng. Administrasi
 - Membawa :
 1. Foto Copy Kartu Tanda Pelajar/Mahasiswa (KTP) sebanyak 1 (satu) lembar.
 2. Foto Copy Polis Asuransi Kecelakaan Kerja/Kesehatan sebanyak 1 (satu) lembar.
 3. Pas foto berwarna ukuran 2x3 sebanyak 2 (dua) lembar.
 4. Printout Surat Panggilan dan Dokumen Pendukung.

Demikian atas perhatian Saudara kami sampaikan terima kasih.


Gresik, 08 Maret 2019
PT Semen Indonesia (Persero) Tbk
An. Direksi,
SM of Training & Development

SEMIEN
INDONESIA
TONY GUNAWAN, ST., MM.

Office :

(1) Gedung Utama PTs/II. Veteran Gresik 61122, Indonesia/Telp: +62 31 3981 - 3/Fax + 62 31 3972264, 3983209
(2) The East Tower Lantai 18 / Jl. Dr. Ide Anak Agung Gde Agung Kav. E.3.2 No. 1 / Jakarta 12950, Indonesia / Telp: +62 21 526 1176

Lampiran 18. Surat Penerimaan Penelitian (Lanjutan).

 PT Semen Indonesia (Persero) Tbk.	
Kepada Yth : Unit of Tuban & Gresik Quality Assurance	F/26203200/004-2
Perihal : Permohonan Kerja Praktek	
Terlampir kami sampaikan data mahasiswa permohonan Kerja Praktek dari : Institut Teknologi Sepuluh Nopember Surabaya	
Nama mahasiswa : Rizki Febriasto	
Jumlah mahasiswa : 1 (satu) orang	
Dalam rangka : Kerja Praktek	
Jurusan : Matematika, Komputasi dan Sains Data	
Tanggal pengajuan : 18 Maret 2019 s.d. 29 Maret 2019	
Lama Kerja Praktek : 2 (dua) minggu	
Materi Proposal Mhs. : Analisis Pengendalian Statistik pada Produksi Portland Composite Cement	
Gresik, 06 Maret 2019 Hormat Kami Training & Development Officer Ttd. Moch. Soebchan, SE	
Mohon <i>konfirmasi</i> atas permohonan kami,	
Mahasiswa tersebut : (<input checked="" type="checkbox"/>) dapat dibantu (<input type="checkbox"/>) tidak dapat dibantu	
Tanggal disetujui Kerja Praktek : 18 Maret 2019 s.d 29 Maret 2019	
Pembimbing yang ditunjuk Nopeq : 1775 Nama pegawai : DJOHAR ARIFIN Unit Kerja : Unit of Tuban & Gresik Quality Assurance Jabatan :	
Tuban, 08 Maret 2019 (DJOHAR ARIFIN)	
<small>Office : [1] Gedung Utama 30/9, Veteran Gresik 61122, Indonesia/Telp: +62 31 3881 - 5000 / Fax + 62 31 3872268, 3884209 [2] The East Tower Lantai 18 / 8. St. Jde Arah Agung Side Agung Km. 6.2.2 No. 1 / Jakarta 12050, Indonesia / Telp: +62 21 5261176</small>	

BIODATA PENULIS



Penulis memiliki nama lengkap RIZKI FEBRIASTO. Lahir di Balikpapan tanggal 6 Februari 1996 dengan predikat anak bungsu. Bertempat tinggal di Perumahan Royal Ketintang Regency Blok E-03 Surabaya. Penulis telah menempuh pendidikan formal hingga Sekolah Menengah Atas di Surabaya. Mulai dari SD Laboratorium UNESA, SMP Negeri 22 Surabaya, dan SMA Negeri 16 Surabaya. Lulus dari SMA, penulis melanjutkan pendidikannya di Departemen STATISTIKA BISNIS Fakultas Vokasi ITS tahun 2014 dan dilanjutkan Lintas Jalur tahun 2017 di Departemen STATISTIKA-FMKSD ITS. Selama perkuliahan penulis aktif organisasi di HIMADATA-ITS (Himpunan Mahasiswa Diploma Statistika ITS) serta mengikuti kegiatan kepanitiaan dan organisasi di lingkup Kampus ITS. Penulis pernah menjadi Staff PSDM (Pengembangan Sumber Daya Mahasiswa) HIMADATA-ITS pada tahun kedua dan menjadi Ketua Tim Ahli HIMADATA-ITS pada tahun ketiga. Penulis juga mengikuti serangkaian *big event* yang diselenggarakan oleh Departemen Statistika dan Departemen Statistika Bisnis yaitu PRS 2016, penulis menjadi koordinator Akomodasi dan Transportasi. Penulis pernah mengikuti dan mengambil pengalaman di salah satu lembaga survey. Kritik, saran, dan masukan dapat dikirim melalui email penulis rizkifebriato11@gmail.com.