

**PROGNOSIS KERUSAKAN BANTALAN GELINDING DENGAN  
MENGUNAKAN METODE *SUPPORT VECTOR REGRESSION* (SVR)**

**Tesis**  
**untuk memenuhi sebagian persyaratan**  
**mencapai derajat Sarjana S-2 Program Studi**  
**Magister Sistem Informasi**



**Hevi Herlina Ullu**  
**24010411400025**

**PROGRAM PASCASARJANA**  
**UNIVERSITAS DIPONEGORO**  
**SEMARANG**  
**2013**

**HALAMAN PENGESAHAN**

**[HALAMAN INI SENGAJA DIKOSONGKAN]**

**PERNYATAAN PERSETUJUAN**  
**PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS**

**[HALAMAN INI SENGAJA DIKOSONGKAN]**

**PERNYATAAN**

**[HALAMAN INI SENGAJA DIKOSONGKAN]**

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yesus Kristus yang sudah menolong, memberikan hikmat dan pengetahuan sehingga penulis dapat menyelesaikan tesis yang berjudul “*Prognosis Kerusakan Bantalan Gelinding dengan menggunakan Metode Support Vector Regression (SVR)*” dengan baik. Tesis ini disusun sebagai salah satu syarat untuk mencapai derajat Sarjana S-2 pada program studi Magister Sistem Informasi Universitas Diponegoro Semarang. Pada kesempatan ini penulis ingin mengucapkan terimakasih kepada :

1. Ketua program studi Magister Sistem Informasi Bapak Drs. Bayu Surarso, M.Sc., Ph.D.
2. Bapak Ir. Toni Prahasto, M.Asc., Ph.D., selaku dosen pembimbing I yang sudah memberikan penulis kesempatan untuk bergabung dalam penelitian ini serta secara tidak langsung mengajarkan arti pentingnya kerja keras, tanggungjawab dan konsekuensi dalam segala hal.
3. Bapak Dr. Achmad Widodo, ST., MT., selaku dosen pembimbing II yang juga sudah memberikan waktu, ilmu, arahan dan bimbingan yang baik dalam melakukan penelitian ini.
4. Semua pihak yang secara langsung maupun tidak langsung membantu penulis dalam menyelesaikan penelitian tesis ini.

Akhir kata, penulis menyadari bahwa tesis ini masih memiliki banyak kekurangan keterbatasan, oleh karena itu penulis mengharapkan saran dan kritik yang membangun demi pengembangan kesempurnaan tesis ini.

Semarang, 10 Juni 2013

Penulis.

## DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS .....	iii
HALAMAN PERNYATAAN .....	iv
KATA PENGANTAR .....	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR .....	x
DAFTAR TABEL.....	xiv
DAFTAR ARTI LAMBANG DAN SINGKATAN .....	xv
ABSTRAK .....	xvii
ABSTRACT.....	xviii
 <b>BAB I PENDAHULUAN</b>	
1.1. Latar Belakang .....	1
1.2. Tujuan Penelitian .....	3
1.3. Manfaat Penelitian .....	4
 <b>BAB II TINJAUAN PUSTAKA DAN DASAR TEORI</b>	
2.1 Tinjauan Pustaka .....	5
2.1.1. Pemantauan Kondisi Mesin .....	5
2.1.2. Format Presentasi Data .....	6
2.1.3. <i>Data Preprocessing</i> .....	6
2.1.4. Pengertian Statistik Ddeskriptif dan Ekstraksi Fitur Statistik .....	7
2.1.5. Performa Statistik.....	8
2.2. Dasar Teori.....	9
2.2.1. Perkembangan Penelitian Menggunakan Metode SVR .....	9

2.2.2. Data Mining Getaran dengan menggunakan SVR .....	10
2.2.3. Pengertian Prognosis dan Bantalan Gelinding.....	19
<b>BAB III METODE PENELITIAN</b>	
3.1. Metode Penelitian.....	21
3.1.1. Objek dan Data Penelitian.....	21
3.2. Prosedur Penelitian.....	22
3.2.1. Studi Pustaka.....	24
3.2.2. Perumusan Masalah dan Pengambilan Data .....	24
3.2.3. Proses Prognosis menggunakan Metode SVR .....	25
3.2.4. Analisa dan Pembahasan.....	29
3.2.5. Kesimpulan dan Saran.....	29
<b>BAB IV HASIL PENELITIAN DAN PEMBAHASAN</b>	
4.1. Hasil Penelitian .....	30
4.1.1. Data <i>Benchmarking</i> .....	31
4.1.1.1. <i>Loading Data</i> .....	31
4.1.1.2. <i>Training dan Validation Data</i> .....	31
4.1.1.3. <i>Testing Data</i> .....	33
4.1.2. Data <i>Real</i> .....	33
4.1.2.1. Data Pengukuran Tahap I.....	33
4.1.2.2. Data Pengukuran Tahap II .....	35
4.1.2.3. Data Pengukuran Tahap III.....	36
4.2. Pembahasan.....	37
4.2.1. Prediksi Berdasarkan 3 Fitur pada NASA Data Pengukuran Tahap I ...	37
4.2.1.1. Fitur RMS .....	37
4.2.1.1.1. <i>Loading Data</i> .....	37
4.2.1.1.2. <i>Training dan Validation Data</i> .....	38
4.2.1.1.3. <i>Testing Data</i> .....	39
4.2.1.2. Fitur Skewness .....	40
4.2.1.2.1. <i>Loading Data</i> .....	40
4.2.1.2.2. <i>Training dan Validation Data</i> .....	41

4.2.1.2.3. <i>Testing Data</i> .....	42
4.2.1.3. Fitur Kurtosis .....	43
4.2.1.3.1. <i>Loading Data</i> .....	43
4.2.1.3.2. <i>Training dan Validation Data</i> .....	44
4.2.1.3.3. <i>Testing Data</i> .....	45
4.2.1.4. Fitur Mean .....	45
4.2.1.4.1. <i>Loading Data</i> .....	46
4.2.1.4.2. <i>Training dan Validation Data</i> .....	47
4.2.1.4.3. <i>Testing Data</i> .....	47
4.2.1.5. Fitur Betakurtosis .....	48
4.2.1.5.1. <i>Loading Data</i> .....	48
4.2.1.5.2. <i>Training dan Validation Data</i> .....	49
4.2.1.5.3. <i>Testing Data</i> .....	49
4.2.2. Prediksi Berdasarkan 3 Fitur pada NASA Data Pengukuran Tahap II ..	50
4.2.2.1. Fitur RMS .....	50
4.2.2.1.1. <i>Loading Data</i> .....	50
4.2.2.1.2. <i>Training dan Validation Data</i> .....	51
4.2.2.1.3. <i>Testing Data</i> .....	52
4.2.2.2. Fitur Skewness .....	53
4.2.2.2.1. <i>Loading Data</i> .....	53
4.2.2.2.2. <i>Training dan Validation Data</i> .....	54
4.2.2.2.3. <i>Testing Data</i> .....	55
4.2.2.3. Fitur Kurtosis .....	55
4.2.2.3.1. <i>Loading Data</i> .....	55
4.2.2.3.2. <i>Training dan Validation Data</i> .....	56
4.2.2.3.3. <i>Testing Data</i> .....	57
4.2.2.4. Fitur Mean .....	58
4.2.2.4.1. <i>Loading Data</i> .....	58
4.2.2.4.2. <i>Training dan Validation Data</i> .....	59
4.2.2.4.3. <i>Testing Data</i> .....	60
4.2.2.5. Fitur Betakurtosis .....	60



4.2.2.5.1. <i>Loading Data</i> .....	60
4.2.2.5.2. <i>Training dan Validation Data</i> .....	61
4.2.2.5.3. <i>Testing Data</i> .....	62
4.2.3. Prediksi Berdasarkan 3 Fitur pada NASA Data Pengukuran Tahap III.	62
4.2.3.1. Fitur Mean .....	62
4.2.3.1.1. <i>Loading Data</i> .....	62
4.2.3.1.2. <i>Training dan Validation Data</i> .....	63
4.2.3.1.3. <i>Testing Data</i> .....	64
4.2.3.2. Fitur Skewness .....	65
4.2.3.2.1. <i>Loading Data</i> .....	65
4.2.3.2.2. <i>Training dan Validation Data</i> .....	66
4.2.3.2.3. <i>Testing Data</i> .....	67
4.2.3.3. Fitur Kurtosis .....	67
4.2.3.3.1. <i>Loading Data</i> .....	67
4.2.3.3.2. <i>Training dan Validation Data</i> .....	68
4.2.3.3.3. <i>Testing Data</i> .....	69
4.2.3.4. Fitur RMS .....	70
4.2.3.4.1. <i>Loading Data</i> .....	70
4.2.3.4.2. <i>Training dan Validation Data</i> .....	71
4.2.3.4.3. <i>Testing Data</i> .....	71
4.2.3.5. Fitur Betakurtosis .....	72
4.2.3.5.1. <i>Loading Data</i> .....	72
4.2.3.5.2. <i>Training dan Validation Data</i> .....	73
4.2.3.5.3. <i>Testing Data</i> .....	73
4.2.4. Analisis Data Berdasarkan Nilai RMSE dan R.....	74
4.2.5. Prognosis Kondisi Mendatang Mesin .....	75
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1. Kesimpulan .....	78
5.2. Saran.....	78

## DAFTAR PUSTAKA

## DAFTAR GAMBAR

	Halaman
Gambar 2.1 Macam fungsi pemisah yang optimal secara linear .....	12
Gambar 2.2 Memperbesar Margin untuk meningkatkan probabilitas .....	13
Gambar 2.3 Kernel Map yang mengubah masalah tidak linear menjadi linear.....	14
Gambar 2.4 Batas <i>error</i> dalam <i>Support Vector Regression</i> .....	16
Gambar 2.5 Variabel <i>Slack</i> untuk mengatasi masalah ketidaklayakan pembatas.....	18
Gambar 2.6 Bantalan Gelinding.....	20
Gambar 3.1 Ilustrasi penempatan sinyal dan tes kondisi bantalan gelinding .	22
Gambar 3.2 Diagram alir penelitian.....	23
Gambar 3.3 Blok diagram sistem.....	24
Gambar 3.4 Flowchart prognosis kerusakan bantalan gelinding .....	25
Gambar 3.5 Plotting pengukuran data tahap I.....	26
Gambar 3.6 Proses ekstraksi fitur .....	27
Gambar 3.7 Hasil ekstraksi fitur statistik data pengukuran tahap I .....	27
Gambar 3.8 Plotting jumlah data training fitur kurtosis .....	28
Gambar 4.1 Data getaran amplitudo pada kompresor.....	31
Gambar 4.2 Jumlah data <i>training</i> dan <i>testing</i> .....	32
Gambar 4.3 hasil <i>training data</i> .....	32
Gambar 4.4. Hasil <i>Testing data</i> .....	33
Gambar 4.5 Plotting data NASA pengukuran Tahap I .....	34
Gambar 4.6 Plotting 5 fitur ((a) fitur mean, (b) fitur RMS, (c) fitur skewness (d) fitur kurtosis, (e) fitur betakurtosis) pada pengukuran tahap I .....	34
Gambar 4.7 Plotting NASA data pengukuran Tahap II.....	35
Gambar 4.8 Plotting 5 fitur ((a) fitur mean, (b) fitur RMS, (c) fitur skewness, (d) fitur kurtosis, (e) fitur betakurtosis) pada pengukuran	

tahap II.....	36
Gambar 4.9 Plotting Data NASA Pengukuran tahap III .....	36
Gambar 4.10 Plotting 5 fitur ((a) fitur mean, (b) fitur RMS, (c) fitur skewness, (d) fitur kurtosis, (e) fitur betakurtosis) pada pengukuran tahap III.....	37
Gambar 4.11 Plotting fitur <i>RMS</i> tahap I .....	38
Gambar 4.12 Plotting jumlah data <i>training</i> dan <i>testing</i> fitur RMS tahap I.....	38
Gambar 4.13 plotting <i>training data</i> dan <i>validasi model</i> Fitur RMS Tahap I .....	39
Gambar 4.14 Plotting Hasil <i>testing</i> data fitur RMS Tahap I.....	40
Gambar 4.15 Plotting data fitur <i>skewness</i> pada data pengukuran tahap I.....	40
Gambar 4.16 Jumlah data <i>training</i> dan <i>testing</i> fitur skewness Tahap I.....	41
Gambar 4.17 Plotting hasil <i>training</i> dan <i>validasi model</i> fitur skewness Tahap I.....	41
Gambar 4.18 Plotting hasil <i>testing</i> data fitur skewness tahap I .....	42
Gambar 4.19 Data fitur kurtosis pada data pengukuran tahap I.....	43
Gambar 4.20 Jumlah data <i>training</i> dan <i>testing</i> .....	44
Gambar 4.21 plotting <i>training</i> data fitur kurtosis tahap I .....	44
Gambar 4.22 Plotting hasil testing pada fitur kurtosis tahap I.....	45
Gambar 4.23 Plotting fitur mean tahap I.....	46
Gambar 4.24 Plotting Jumlah data <i>training</i> dan <i>testing</i> fitur mean Tahap I....	46
Gambar 4.25 Plotting <i>training</i> dan validasi model fitur mean Tahap I .....	47
Gambar 4.26 Plotting hasil testing pada fitur mean tahap I.....	48
Gambar 4.27 Plotting fitur betakurtosis tahap I.....	48
Gambar 4.28 Plotting <i>training</i> dan validasi model fitur betakurtosis Tahap I.....	49
Gambar 4.29 Plotting hasil testing pada fitur betakurtosis tahap I .....	50
Gambar 4.30 Plotting kondisi mesin berdasarkan fitur RMS Tahap II.....	51
Gambar 4.31 Plotting Jumlah data <i>training</i> dan <i>testing</i> fitur RMS tahap II....	51
Gambar 4.32 Plotting hasil training dan validasi model fitur <i>RMS</i> tahap I .....	52

Gambar 4.33 Plotting hasil <i>testing</i> fitur <i>RMS</i> tahap I .....	53
Gambar 4.34 Plotting data fitur <i>skewness</i> pengukuran tahap II.....	53
Gambar 4.35 Plotting Jumlah data <i>training</i> dan <i>testing</i> fitur <i>Skewness</i> Tahap II.....	54
Gambar 4.36 hasil training dan Validasi fitur <i>skewness</i> tahap II.....	54
Gambar 4.37 Plotting hasil testing fitur <i>skewness</i> tahap II.....	55
Gambar 4.38 Plotting hasil ekstraksi fitur kurtosis tahap II.....	56
Gambar 4.39 Plotting Jumlah data <i>training</i> dan <i>testing</i> fitur kurtosis Tahap II.....	56
Gambar 4.40 Plotting hasil <i>training</i> fitur kurtosis tahap II.....	57
Gambar 4.41 Plotting hasil testing fitur kurtosis tahap II .....	58
Gambar 4.42 Plotting hasil ekstraksi fitur mean tahap II.....	58
Gambar 4.43 Plotting Jumlah data <i>training</i> dan <i>testing</i> fitur mean Tahap II.....	59
Gambar 4.44 Plotting hasil <i>training</i> fitur mean tahap II.....	59
Gambar 4.45 Plotting hasil testing fitur mean tahap II .....	60
Gambar 4.46 Plotting hasil ekstraksi fitur betakurtosis pengukuran tahap II..	61
Gambar 4.47 Plotting hasil <i>training</i> fitur betakurtosis tahap II .....	62
Gambar 4.48 Plotting hasil testing fitur betakurtosis tahap II .....	62
Gambar 4.49 Plotting hasil ekstraksi fitur Mean tahap III.....	63
Gambar 4.50 Plotting jumlah data training dan testing fitur mean tahap III.....	63
Gambar 4.51 Plotting hasil training dan Validasi fitur mean tahap III.....	64
Gambar 4.52 Plotting hasil testing fitur mean tahap III.....	65
Gambar 4.53 Plotting Fitur <i>skewness</i> .....	65
Gambar 4.54 Plotting Jumlah data <i>training</i> dan <i>testing</i> fitur <i>skewness</i> tahap III .....	66
Gambar 4.55 Hasil proses <i>Training</i> dan validasi model fitur <i>Skewness</i> tahap III .....	66
Gambar 4.56 Plotting hasil proses testing fitur <i>skewness</i> tahap III .....	67
Gambar 4.57 Plotting trend data fitur kurtosis tahap III .....	68

Gambar 4.58 Plotting Point Data Training dan Testing	
Fitur Mean Tahap III .....	68
Gambar 4.59 Plotting hasil training fitur kurtosis tahap III .....	69
Gambar 4.60 Plotting hasil testing fitur kurtosis tahap III .....	70
Gambar 4.61 Plotting hasil ekstrasi fitur RMS tahap III .....	70
Gambar 4.62 Plotting hasil training fitur RMS tahap III .....	71
Gambar 4.63 Plotting hasil testing fitur RMS tahap III .....	72
Gambar 4.64 Plotting hasil ekstrasi fitur betakurtosis tahap III .....	72
Gambar 4.65 Plotting hasil training fitur betakurtosis tahap III .....	73
Gambar 4.66 Plotting hasil testing fitur betakurtosis tahap III .....	74
Gambar 4.67 Plotting prediksi <i>step ahead</i> fitur RMS .....	76

## DAFTAR TABEL

4.1. Spesifikasi perekaman Data NASA .....	30
4.2. Pemilihan model prognosis berdasarkan RMSE dan R .....	75

## DAFTAR ARTI LAMBANG DAN SINGKATAN

### DAFTAR ARTI LAMBANG

<b>Lambang</b>	<b>Arti Lambang</b>
$-\varepsilon / +\varepsilon$	<i>Derajat toleransi terhadap error</i>
$w$	<i>Weight vector</i>
$x$	<i>Input vector</i>
$\varphi(x)$	Hasil pemetaan $x$ didalam <i>input space</i>
$\xi$	Fungsi Slack
$\ w\ $	Normalisasi $W$ / regularisasi
<b><math>b</math></b>	Bias
<b><math>f(x)</math></b>	Fungsi nilai aktual sebelum adanya <i>tresholding</i>
$y_i$	Nilai yang diamati
$\hat{y}$	Nilai prediksi
$C$	Nilai pinalti jika deviasi $>$ batas error
$L$	Fungsi resiko
$L_\varepsilon$	<i>Loss function</i> yang bertipe $\varepsilon$ -insensitive
$N$	<i>Jumlah data</i>
$\sigma$	<i>Varian</i>
$\mu$	<i>Rata-rata</i>
$E$	<i>Expected value</i>
$m_4$	<i>Momen ke empat</i>

## DAFTAR SINGKATAN

<b>Singkatan</b>	<b>Kepanjangan Singkatan</b>
ANFIS	<i>Adaptive Neuro-Fuzzy Integrated System</i>
NN	<i>Neural Network</i>
NARX	<i>Nonlinear AutoRegressive eXogenous</i>
NF	<i>Neuro Fuzzy</i>
SVR	<i>Support Vector Regression</i>
SVM	<i>Support Vector Machines</i>
RMS	<i>Root Mean Square</i>
RMSE	<i>Root Mean Square Error</i>
R	<i>Correlation Statistic Coefficient</i>
COV	<i>Covariance</i>



# **PROGNOSIS KERUSAKAN BANTALAN GELINDING DENGAN MENGGUNAKAN METODE SUPPORT VECTOR REGRESSION (SVR)**

## **ABSTRAK**

Pada penelitian ini telah dilakukan prognosis kerusakan bantalan gelinding dengan menggunakan metode *Support Vector Regression* (SVR). Penelitian ini dilakukan dengan menggunakan objek bantalan gelinding karena bantalan gelinding merupakan komponen yang mampu membuat sebuah mesin terus berputar atau bekerja. Dengan melakukan prognosis terhadap kerusakan bantalan gelinding dapat mengoptimalkan biaya perawatan mesin karena bisa mengetahui sisa umur fungsi bantalan gelinding sebelum bantalan gelinding tersebut rusak.

Data masukan berasal dari hasil ekstraksi beberapa fitur statistik dari data trend sinyal getaran bantalan gelinding. Data fitur yang dihasilkan digunakan dalam proses pembelajaran dan proses pengujian dengan metode SVR, setelah itu akan menghasilkan prognosis kerusakan bantalan gelinding yang mendekati nilai ideal dari *Root Mean Square Error* (RMSE) dan *Coefisien Corelation* (R).

Hasil dari penelitian ini adalah fitur RMS merupakan fitur yang bagus dipakai untuk melakukan prognosis. Nilai RMSE dan R untuk kedua fitur tersebut mendekati nilai ideal yaitu nilai nilai RMSE untuk fitur RMS adalah 0.0129 sedangkan nilai R untuk fitur RMS adalah 0.9709.

Kata Kunci : Prognosis, sisa umur fungsi, bantalan gelinding, ekstraksi fitur, SVR

# **PROGNOSIS OF BEARING DAMAGE USING SUPPORT VECTOR REGRESSION METHOD (SVR)**

## **ABSTRACT**

This research has been performed on the prognosis of rolling bearing damage using by Support Vector Regression (SVR). This research was performed using rolling bearings object as a rolling bearing component capable of making a machine continues to spin or work. Prognosis for rolling bearing damage can optimize maintenance costs because it can determine the remaining life of rolling bearings function before rolling bearing is damaged.

Data derived from the extraction of some statistical features of the data trend rolling bearing vibration signal. The resulting feature data used in the process of learning and testing processes with SVR method, as it will result in rolling bearing damage prognosis approaching the ideal value of Root Mean Square Error (RMSE) and Coefisien Correlation (R).

The results of this research is RMS feature is a good feature to perform the prognosis. RMSE and R values for RMS feature is close to the ideal value. RMSE values for RMS feature is 0.0129, while the value of R for RMS feature is 0.9709.

Keywords: Prognosis, residual life function, rolling bearings, feature extraction, SVR

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

*Artificial intelligent* atau yang sering disebut dengan kecerdasan buatan merupakan bagian ilmu komputer yang membuat agar mesin komputer dapat melakukan pekerjaan seperti dan sebaik yang dilakukan oleh manusia. *Machine learning* merupakan salah satu area dalam kecerdasan buatan yang berhubungan dengan cara memrogram mesin atau komputer dengan contoh data dan pengalaman dari data masa lalu atau yang sudah pernah ada. *Machine learning* juga menjadi alat analisis dalam data mining. Algoritma *machine learning* dapat diatur dalam taksonomi berdasarkan hasil yang diinginkan dari algoritma, antara lain menemukan hubungan antar data (*learning associations*), proses pembelajaran dengan pemetaan data dari input ke output (*supervised learning*), proses pembelajaran yang hanya memiliki input data untuk menemukan regularitas dalam data tanpa output (*unsupervised learning*) dan proses pembelajaran berdasarkan bagaimana pengalaman memecahkan suatu permasalahan (*reinforcement learning*) (Santosa, 2007).

*Data Mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis (Hermawati, 2009). *Data mining* juga merupakan proses menganalisa set data observasi untuk menemukan hubungan yang tidak terduga dan meringkas data tersebut dengan cara baru yang dimengerti dan berguna bagi pemilik data (Hand dkk., 2001). Ada beberapa teknik dan sifat *data mining* antara lain yang bersifat prediktif yaitu klasifikasi, regresi, sedangkan yang bersifat deskriptif yaitu klusterisasi, kaidah asosiasi, pencarian pola sekuensial.

Regresi merupakan salah satu teknik dalam *data mining* yang digunakan untuk memprediksi nilai dari suatu variabel kontinu yang diberikan berdasarkan nilai dari variabel yang lain, dengan mengasumsikan sebuah model ketergantungan linier atau nonlinier. Proses pembelajaran yang dilakukan dalam

teknik regresi adalah *supervised learning* atau proses pembelajaran yang menggunakan data pelatihan dan data pengujian. Ada beberapa metode yang sering digunakan untuk melakukan regresi dengan proses pembelajaran *supervised learning*, antara lain *Adaptive Neuro-Fuzzy Integrated System* (ANFIS), *Neural Network* (NN), *Nonlinear AutoRegressive eXogeneous* (NARX), *Neuro Fuzzy* (NF), *Support Vector Regression* (SVR).

Pada penelitian ini, akan digunakan metode SVR untuk melakukan proses prognosis yaitu kemampuan untuk menilai baik dan tidaknya suatu bagian mesin dan untuk memprediksi waktu kerusakan dari bagian mesin tersebut. Pada awalnya Vapnik menerapkan metode *Support Vector Machines* (SVM) untuk kasus klasifikasi tetapi dengan menggunakan konsep  *$\epsilon$ -insensitive loss function*, SVM bisa digeneralisasi untuk melakukan pendekatan fungsi (*function approximation*) atau regresi (Scholkopf dan Smola, 2002). Prognosis menggunakan NN dibangun secara otomatis melalui pelatihan tanpa perlu untuk identifikasi struktur model dan parameter. NN memiliki dua arsitektur jaringan yang khas yaitu *feedforward* dan *recurrent network* dan keduanya telah digunakan dalam peramalan perilaku sistem. Sedangkan NF memiliki kelebihan dalam mengidentifikasi masalah berdasarkan informasi fuzzy dan pengenalan pola, pembelajaran dan *training* data yang baik tetapi kelemahannya diperlukan *training* data yang besar dan lambatnya konvergensi. SVR merupakan metode yang mampu mengatasi *overfitting* yaitu keadaan dimana sebuah model klasifikasi dapat melakukan klasifikasi data dengan sangat baik pada proses *training set*, tetapi sangat buruk dalam melakukan klasifikasi data *output* yang baru dan belum pernah ada, sehingga akan menghasilkan performa yang bagus. SVR digunakan untuk mendapatkan suatu fungsi dengan tingkat kesalahan paling kecil sehingga menghasilkan sebuah prediksi yang bagus. Ide dasar dari SVR adalah dengan menentukan set data yang dibagi menjadi *set training* dan *set validasi*. Kemudian dari *set training* tersebut ditentukan suatu fungsi regresi dengan batasan deviasi tertentu sehingga dapat menghasilkan prediksi yang mendekati target aktual. Metode SVR diaplikasikan dalam berbagai bidang antara lain untuk peramalan *time series* keuangan (Lu dkk., 2009), prediksi tingkat sirkulasi vertikal dalam

tabung *thermosipher reboiler* (Zaidi, 2012), peramalan berbasis ekstraksi aturan hibrida (Farquad dkk., 2010), sedangkan untuk penelitian ini, metode SVR akan digunakan untuk memprognosis kerusakan bantalan gelinding berdasarkan data sinyal getaran.

*Bearing* atau yang sering disebut sebagai bantalan gelinding merupakan sebuah elemen mesin yang berfungsi untuk membatasi gerak relatif antara dua atau lebih komponen mesin agar selalu bergerak pada arah yang diinginkan. Bantalan gelinding juga menjaga poros (*shaft*) agar selalu berputar terhadap sumbu porosnya, atau juga menjaga komponen yang bergerak linear agar selalu berada pada jalurnya. Bantalan gelinding mempunyai elemen yang berputar dan bagian yang diam saat bekerja yang terletak antara poros dan rumah bantalan gelinding. Kegagalan perputaran bantalan gelinding merupakan salah satu penyebab utama kerusakan dalam sebuah mesin (Niu, 2005). Kerusakan bantalan gelinding harus diprediksi dengan akurat karena akan menentukan berapa lama sisa umur fungsi dari sebuah mesin.

Dengan kelebihan dari metode SVR, maka penelitian ini akan menerapkan metode SVR untuk memprognosis kerusakan bantalan gelinding berdasarkan data sinyal getaran bantalan gelinding karena dengan melakukan prognosis akan membantu mengetahui kondisi mendatang dari sebuah mesin sehingga bisa meminimalisir biaya kerusakan yang akan terjadi.

## **1.2. Tujuan Penelitian**

Tujuan dari penelitian ini adalah :

- a. Menentukan fitur statistik yang bagus untuk dijadikan model prediksi pada metode SVR.
- b. Menguji tingkat validasi output dari metode SVR untuk menjamin kemampuan dari metode SVR.
- c. Menerapkan metode SVR untuk memprediksi kondisi mendatang sebuah mesin berdasarkan hasil ekstraksi fitur data getaran bantalan gelinding.

### **1.3. Manfaat Penelitian**

Manfaat dari penelitian ini yaitu :

- a. Dapat membantu mengatasi masalah sulitnya memprediksi kondisi sebuah mesin jika sering terjadi kerusakan pada bantalan gelinding dan membantu mengurangi biaya perawatan karena kerusakan sebuah mesin sudah dapat diprediksi sebelum mesin mengalami kerusakan yang lebih parah.
- b. Dapat membantu pengambil keputusan agar tidak bersifat subjektif dalam menentukan kondisi kerusakan sebuah mesin.

## **BAB II**

### **TINJAUAN PUSTAKA DAN DASAR TEORI**

#### **2.1. Tinjauan Pustaka**

##### **2.1.1. Pemantauan Kondisi Mesin**

Kondisi getaran sebuah mesin dapat dilihat dari pemantauan getaran bantalan gelinding atau rumah bantalan gelinding dan atau dari pemantauan getaran perputaran elemen pada mesin. Pemantauan kondisi getaran tersebut bisa dilakukan terus menerus atau hanya pada beberapa waktu tertentu saja. Pemantauan getaran dilakukan untuk memantau “sehat atau tidaknya” sebuah mesin dalam jangka waktu pemakaian mesin tersebut, tergantung dari tipe mesin dan komponen yang ingin dipantau. Dari waktu pemantauan yang dilakukan, kondisi “tidak sehat” dari sebuah mesin dapat lebih awal diketahui sehingga dapat diambil tindakan perbaikan pada mesin tersebut.

Pemantauan kondisi getaran sebuah mesin dapat dilakukan dengan berbagai cara antara lain :

- a. Pemantauan secara umum yaitu dengan memantau biaya operasional, biaya mesin, kekritisan operasi dari mesin, biaya keamanan, dampak penggunaan mesin terhadap lingkungan dan lain sebagainya.
- b. Pemantauan melalui sistem yang dipasang pada mesin yaitu *transducers*, kondisi sinyal, pengolahan data dan penyimpanan data terpasang secara permanen pada mesin. Data dapat dikumpulkan terus menerus atau secara berkala.
- c. Pemantauan yang ketiga dapat dilakukan dengan sistem semi permanen yaitu penggabungan antara sistem permanen dan *portable* dimana jenis sistem ini *transducers* dipasang secara permanen sedangkan data akuisis komponen dapat dihubungkan sesuai kebutuhan (ISO 13373-1:2002(E)).

##### **2.1.2. Format Presentasi Data**

Format penyajian analisis data pemantauan kondisi getaran sebuah mesin dapat disajikan secara umum sejak data awal diambil. Data awal ini terdiri dari

semua parameter getaran yang biasa digunakan untuk mendefinisikan kondisi getaran. Format penyajian data tersebut antara lain penyajian berdasarkan nilai dari *broadband*, waktu dan sinyal gelombang, amplitudo per frekuensi, getaran vektor, dan lainnya. Format penyajian data yang lengkap dari awal data diambil akan membantu menentukan penyebab dari perubahan getaran (ISO 13373-1:2002 (E)).

### **2.1.3. Data Preprocessing**

Data yang ada dalam dunia nyata adalah data yang “kotor” yang disebabkan oleh beberapa faktor antara lain:

- a. Data tidak lengkap karena ketika dikumpulkan data tersebut sudah tidak berlaku, adanya perbedaan pemikiran ketika data tersebut dikumpulkan dengan data ketika dianalisa, dan kesalahan dari perangkat lunak, perangkat keras, dan pengguna yang memakai data tersebut.
- b. Nilai data yang salah karena berasal dari instrumen pengumpulan datayang salah, kesalahan manusia atau komputer ketika memasukkan data serta terjadinya kesalahan transmisi data.
- c. Data yang tidak tetap karena perbedaan sumber data.

Tugas utama dari *data preprocessing* antara lain membersihkan data, mengintegrasikan data, mentransformasi data, mereduksi data. Data yang sudah dibuat format presentasinya harus dilakukan *data preprocessing* karena jika tidak ada data yang berkualitas maka tidak ada pertimbangan yang berkualitas. Kesimpulan yang berkualitas harus berdasarkan kualitas data (ISO 13373-1:2002(E)).

### **2.1.4. Pengertian Statistik Deskriptif dan Ekstraksi Fitur Statistik**

Statistik deskriptif adalah metode-metode yang berkaitan dengan pengumpulan data dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Statistik deskriptif hanya memberikan informasi mengenai data yang dipunyai dan sama sekali tidak menarik kesimpulan apapun



tentang populasi yang lebih besar. Contoh statistik deskriptif yang sering muncul adalah berupa tabel, diagram, grafik dan lain sebagainya (Wahyuni, 2011).

Ekstraksi fitur dilakukan karena merupakan bagian dari analisis data. Analisis data merupakan salah satu metode yang penting dalam memonitoring dan mendiagnosa kerusakan, dimana tujuannya adalah menemukan transformasi yang sederhana dan efektif dari data sinyal yang original.

Data yang direpresentasikan oleh fitur akan menjadi nilai yang memiliki karakteristik. Untuk suatu kondisi dimana diperlukan kegiatan pemantauan dan diagnosa kesalahan, fitur merupakan data yang dapat merepresentasikan sebuah nilai yang mengindikasikan kondisi – kondisi dari mesin. Selain itu fitur sangat penting karena dapat mentransformasi dan menyimpan permasalahan yang suatu saat dapat diselesaikan.

*Root Mean Square (RMS), variance, mean, skewness, kurtosis, betakurtosis* merupakan beberapa bentuk representasi fitur statistik pada domain waktu. Fitur yang digunakan dapat menentukan bagus dan tidaknya keluaran. Fitur yang bagus akan menghasilkan keluaran yang bagus (Medjaher dkk., 2012). Ada lima fitur yang digunakan untuk mengekstraksi data, yaitu *mean, RMS, skewness, kurtosis* dan *betakurtosis*. Lima fitur tersebut dapat dideskripsikan sebagai berikut:

1. Mean adalah rata-rata dari suatu array (deret angka). Persamaan dari mean adalah sebagai berikut:

$$\mu_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (2.1)$$

2. RMS adalah rata-rata kuadrat yang dirumuskan sebagai berikut:

$$RMS = \sqrt{\frac{1}{2} \sum_n x^2(t)} \quad (2.2)$$

3. Skewness adalah pengukuran data asimetri di sekitar rata-rata sampel. Jika bernilai positif maka data lebih tersebar ke kanan dan jika bernilai negatif maka data lebih tersebar ke kiri. Jika distribusi normal (simetris sempurna), maka skewness bernilai nol. Skewness dirumuskan dengan persamaan sebagai berikut:

$$y = \frac{E(x-\mu)^3}{\sigma^3} \quad (2.3)$$

4. Kurtosis adalah derajat keruncingan dari suatu distribusi. Kurtosis dapat dirumuskan dengan persamaan berikut :

$$K = \frac{E(x-\mu)^4}{\sigma^4} \quad (2.4)$$

5. Beta kurtosis adalah moment ke empat dari fungsi beta. Beta Kurtosis dirumuskan pada persamaan berikut :

$$BK = \frac{m_4}{(\sigma^2)^2} \quad (2.5)$$

### 2.1.5. Performa Statistik

Ada beberapa cara untuk mengukur atau menguji performa statistik antara lain dengan menggunakan *root mean square error* (RMSE) dan *correlation statistic coefficient* (R). RMSE memberi ilustrasi umum tentang baik dan tidaknya akurasi prediksi yang dilakukan dan RMSE merupakan ukuran dari kesalahan rata-rata menurut kuadrat kesalahan. Karena merupakan jumlah kuadrat, RMSE lebih kuat dipengaruhi oleh *error* yang besar daripada *error* yang kecil, sehingga memiliki jangkauan dari 0 hingga tak terbatas, dengan 0 menjadi nilai sempurna. RMSE dapat dihitung dengan persamaan :

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (2.6)$$

Dimana  $N$  merepresentasikan jumlah total dari data point pada data set *testing*,  $y$  merepresentasikan nilai yang akan diamati dan  $\hat{y}$  merepresentasikan nilai dari prediksi.

Sedangkan R mengukur korelasi linear antara nilai aktual dan nilai prediksi. R dapat dihitung dengan persamaan berikut

$$R_{y,\hat{y}} = \frac{Cov(y,\hat{y})}{\sigma_y \sigma_{\hat{y}}} \quad (2.7)$$

Dimana  $R$  adalah koefisien korelasi dan  $Cov(y,\hat{y})$  adalah kovarian antara nilai yang diamati dan nilai prediksi (Widodo, 2009).

## 2.2 . Dasar Teori

### 2.2.1. Perkembangan Penelitian Menggunakan Metode SVR

Penerapan metode SVR untuk peramalan memiliki keuntungan karena SVR memiliki kemampuan menggeneralisasi yang cukup bagus. Tetapi dalam kasus peramalan *time series* keuangan, banyak terdapat kebisingan data dan sering berubah-ubahnya *time series* tersebut. Untuk mengatasi hal tersebut SVR dikombinasikan dengan *independent component analysis* (ICA) karena ICA mampu mengurangi kebisingan data dengan menghasilkan komponen independen sehingga IC yang bising dapat dibuang dan IC yang tersisa digunakan sebagai inputan dalam SVR untuk melakukan peramalan. Hasil penelitian ini menunjukkan bahwa dengan mengkombinasikan SVR dan ICA maka dapat dilakukan peramalan dengan lebih baik lagi (Lu dkk., 2009)

Hidrodinamika dan transfer panas di *thermosiphon reboiler* berinteraksi satu sama lain membuat proses yang sangat kompleks. Prediksi tingkat perpindahan panas dan aliran termal yang disebabkan adalah persyaratan utama untuk desain *thermosiphon reboiler*. Tujuan dari penelitian ini adalah untuk mengembangkan, untuk pertama kalinya, sebuah model berbasis data terpadu, untuk prediksi laju sirkulasi dalam *thermosiphon reboiler* dengan menggunakan support vector regression (SVR) berbasis teknik pemodelan. Penelitian ini menggunakan 148 poin data eksperimen. Hasil dari penelitian ini adalah model prediksi dengan menggunakan metode SVR dapat menunjukkan prediksi yang bagus dengan nilai RMSE sebesar 0.0717 dan performa generalisasi yang cukup bagus (Zaidi, 2012)

Aplikasi SVM pada peramalan *time-series* keuangan (Cao dkk., 2001), pada penelitian ini SVM diuji kelayakannya pada peramalan waktu keuangan dengan dibandingkan oleh *multi-layer back propagation*. Hasil dari penelitian ini adalah SVM melebihi jaringan saraf *back propagation* berdasarkan kriteria kesalahan, oleh karena itu lebih menguntungkan jika SVM diterapkan dalam peramalan *time-series* keuangan.

Penerapan SVM untuk memprediksi konsumsi energi pada daerah Tropis. Hasil dari penelitian ini adalah fitur yang menonjol dari SVM diimplementasikan

untuk meminimalkan batas atas kesalahan generalisasi bukan pada kesalahan pelatihan seperti yang diterapkan dalam *Neural Network* oleh karena itu SVM layak digunakan untuk memprediksi tagihan bulanan dari pengguna energi di daerah tropis (Dong dkk., 2004).

Penerapan metode *adaptive predictor* dengan pendekatan *neuro fuzzy* untuk memprediksi perilaku sistem dinamis diusulkan untuk meningkatkan efisiensi peramalan. Metode *adaptive predictor* diimplementasikan untuk monitoring kondisi *gear* dan pengujian material yang lelah. Hasil pengujian dari penelitian ini menunjukkan bahwa metode *adaptive predictor* merupakan metode peramalan yang dapat diandalkan karena metode *adaptive predictor* dapat melacak karakteristik sistem dengan akurat, selain itu metode *adaptive predictor* dapat mengakomodasi kondisi sistem operasi yang berbeda dan variasi karakteristik dinamis sistem. Metode *adaptive predictor* efisien dalam meningkatkan kinerja peramalan dengan memodifikasi sifat dari batas-batas ruang keputusan (Wang, 2007).

Penerapan *Intelligent Predictive Decision Support System* (IPDSS) untuk memantau kondisi awal suatu mesin. Hasil dari penelitian ini, prediksi dengan menggunakan IPDSS dapat mengurangi biaya perawatan dengan mempersiapkan tindakan perbaikan yang tepat sebelum terjadi kerusakan fatal sebuah mesin (Yang dkk., 2001).

### **2.2.2. Data Mining Getaran dengan menggunakan SVR**

Metode SVM merupakan salah satu metode dalam *Machine Learning*. SVM adalah satu kumpulan teknik klasifikasi dan regresi, yang merupakan pengembangan algoritma non-linear dan diperkenalkan pertama kali oleh Vapnik pada tahun 1999. Dengan SVM, algoritmanya mencoba menemukan pemisah linear (*hyperplane*) antara titik data dari dua kelas dalam ruang multidimensi. SVM cocok digunakan untuk interaksi antara fitur dan fitur berlebihan.

Kasus klasifikasi secara linear bisa dipisahkan. Fungsi pemisah yang dicari adalah fungsi linear. Fungsi ini bisa didefinisikan sebagai berikut :

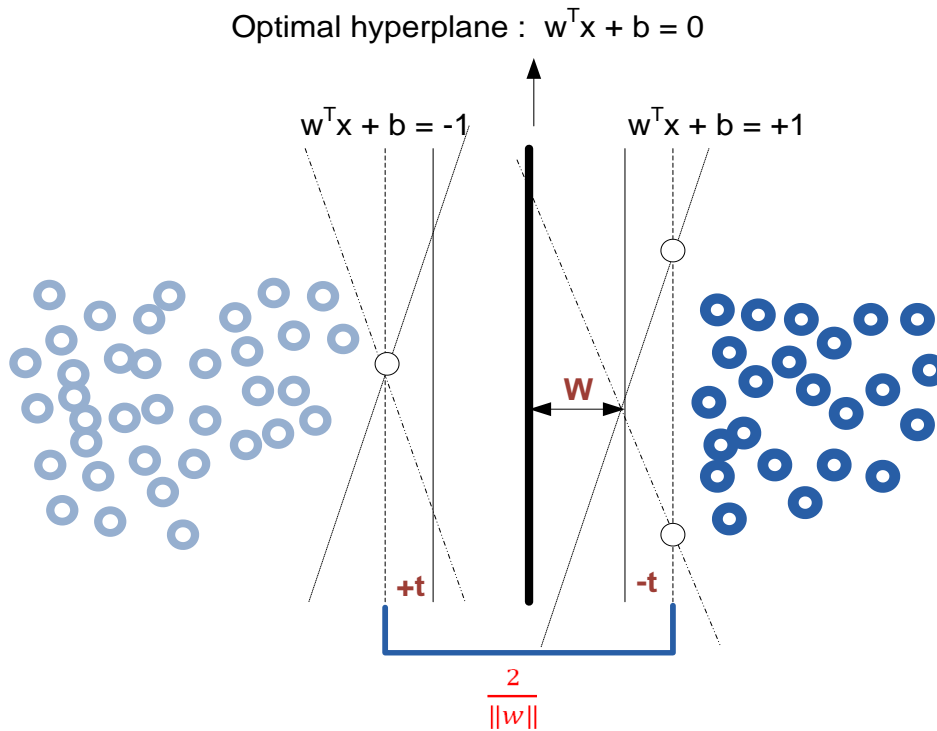
$$g(x) := \text{sgn}(f(x)) \quad (2.8)$$

Dengan  $f(x) = w^T x + b$ ,

Dimana  $x, w \in \mathfrak{R}^n$  dan  $b \in \mathfrak{R}$ . Masalah klasifikasi ini dapat dirumuskan sebagai berikut :  $f(x_i) = \langle w, x \rangle + b = y_i$ , untuk semua  $i$ . Dalam teknik ini diupayakan untuk menemukan fungsi pemisah (*hyperplane*) terbaik diantara fungsi yang tidak terbatas jumlahnya untuk memisahkan dua macam obyek. *Hyperplane* terbaik adalah *hyperplane* yang terletak ditengah-tengah antara dua set obyek dari dua kelas. Mencari *hyperplane* yang terbaik ekuivalen dengan memaksimalkan margin atau jarak antara dua set obyek dari kelas yang berbeda. Jika  $w x_1 + b = +1$  adalah *hyperplane* pendukung dari kelas  $+1$  ( $w x_1 + b = +1$ ) dan  $w x_2 + b = -1$  dan *hyperplane* pendukung dari kelas  $-1$  ( $w x_2 + b = -1$ ), margin antara dua kelas dapat dihitung dengan mencari jarak antara kedua *hyperplane* pendukung dari kedua kelas. Secara spesifik, margin dihitung dengan cara :

$$\begin{aligned} & (w x_1 + b = +1) - (w x_2 + b = -1) & (2.9) \\ \Rightarrow & w(x_1 - x_2) = 2 \\ \Rightarrow & \left( \frac{w}{\|w\|} (x_1 - x_2) \right) = \frac{2}{\|w\|} \end{aligned}$$

Gambar 2.1. memperlihatkan bagaimana SVM bekerja untuk menemukan suatu fungsi pemisah dengan margin yang maksimal. Untuk membuktikan bahwa memaksimalkan margin antara dua set obyek akan meningkatkan probabilitas pengelompokkan secara benar dari data *testing*. Pada dasarnya jumlah fungsi pemisah ini tidak terbatas banyaknya. Misalkan dari jumlah data yang tidak terbatas ini kita ambil dua saja, yaitu  $f_1(x)$  dan  $f_2(x)$  (gambar....). fungsi  $f_1$  mempunyai margin yang lebih besar dari pada fungsi  $f_2$ . Setelah menemukan dua fungsi ini, sekarang dua data baru masuk dengan keluaran  $-1$  atau  $+1$ . Maka harus dikelompokkan apakah data itu ada dalam kelas  $-1$  atau  $+1$  menggunakan fungsi pemisah yang sudah ditentukan. Dengan menggunakan  $f_1$ , akan dikelompokkan data baru ini dikelas  $-1$  yang berarti pengelompokkannya benar. Jika kita gunakan  $f_2$ , data akan ditempatkan di kelas  $+1$  yang berarti salah.



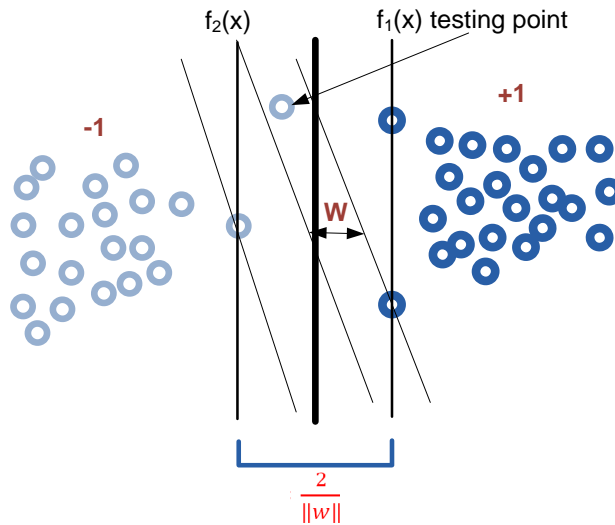
Gambar 2.1. Mencari fungsi pemisah yang optimal secara linear

Secara matematis, formulasi problem optimasi SVM untuk kasus klasifikasi linear di dalam *primal space* adalah:

$$\min \frac{1}{2} \|w\|^2 \tag{2.10}$$

$$\text{Subject to } y_i(w x_i) \geq 1, i = 1, \dots, \ell$$

dimana  $x_i$  adalah data *input*,  $y_i$  keluaran dari data  $x_i$ ,  $w, b$ , adalah parameter – parameter yang dicari nilainya. Dalam formulasi diatas, diminimalkan fungsi tujuan (*objectif function*)  $\frac{1}{2} \|w\|^2$  atau memaksimalkan kuantitas  $\|w\|^2$  atau  $w^T w$  dengan memperhatikan pembatas  $y_i(w x_i + b) \geq 1$ . Bila *output* data  $y_i = +1$ , maka pembatas menjadi  $(w x_i + b) \geq 1$ .



Gambar 2.2 Memperbesar margin untuk meningkatkan probabilitas

Sebaliknya bila  $y_i = -1$ , pembatas menjadi  $(wx_i + b) \leq -1$ . Dalam kasus yang tidak *feasible (infeasible)* dimana beberapa data mungkin tidak bisa dikelompokkan secara benar, maka formulasi matematikanya menjadi :

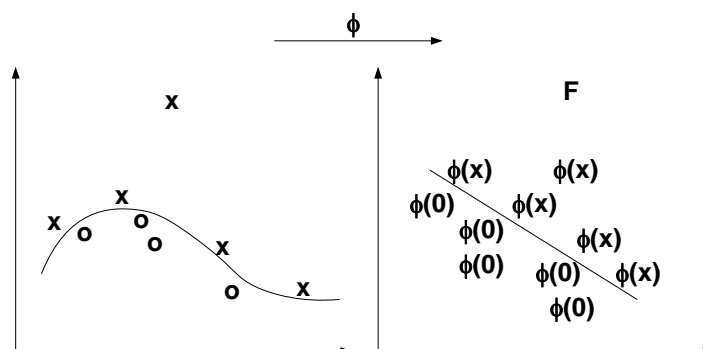
$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\ell} t_i. \quad (2.11)$$

$$\text{Subject to } y_i(wx_i + b) + t_i \geq 1$$

$$t_i \geq 0, i = 1, \dots, \ell$$

Dimana  $t_i$  adalah variabel *slack*. Dengan formulasi tersebut maka untuk memaksimalkan margin antara dua kelas dapat dilakukan dengan meminimalkan  $\|w\|^2$ . Dengan formulasi tersebut dapat diminimalkan kesalahan klasifikasi yang dinyatakan dengan adanya variabel *slack*  $t_i$  sementara dalam waktu yang sama dimaksimalkan margin  $\frac{1}{\|w\|}$ . Penggunaan variabel *slack*  $t_i$  adalah untuk mengatasi kasus ketidaklayakan dari pembatas  $y_i(wx_i + b) \geq 1$  dengan cara memberi pinalti untuk data yang tidak memenuhi pembatas tersebut. Untuk meminimalkan nilai  $t_i$  ini, diberikan pinalti dengan konstanta  $C$ . Vektor  $w$  tegak lurus terhadap fungsi :  $wx + b = 0$ . Konstanta  $b$  menentukan lokasi fungsi pemisah relatif terhadap titik asal.

Untuk mengatasi masalah ketidaklinearan yang sering terjadi dalam kasus yang nyata, bisa diterapkan metode kernel. Metode kernel memberikan pendekatan alternatif dengan cara melakukan mapping data  $x$  dari input space ke *feature space*  $F$  melalui suatu fungsi  $\phi$  sehingga  $\phi : x \rightarrow \phi(x)$ . Karena itu suatu titik  $x$  dalam *input space* menjadi  $\phi(x)$  dalam *feature space*. Sering kali fungsi  $\phi(x)$  tidak tersedia atau tidak bisa dihitung. Tetapi *dot product* dari dua vector dapat dihitung baik dalam *input space* maupun dalam *feature space*. Dengan kata lain, sementara  $\phi(x)$  mungkin tidak diketahui, *dot product*  $\langle \phi(x_1), \phi(x_2) \rangle$  masih bisa dihitung dalam *feature space*. Untuk bisa menggunakan metode kernel, pembatas perlu diekspresikan dalam bentuk *dot product* dari vektor data  $x_i$ . Sebagai konsekuensi, pembatas yang menjelaskan permasalahan dalam klasifikasi harus diformulasikan kembali sehingga menjadi bentuk *dot product*. Dalam *feature space* ini, *dot product*  $\langle \cdot \rangle$  menjadi  $\langle \phi(x), \phi(x') \rangle$ . Suatu fungsi kernel,  $k(x, x')$ , bisa untuk menggantikan *dot product*  $\langle \phi(x), \phi(x') \rangle$ . Kemudian di *feature space*, dibuatkan suatu fungsi pemisah yang linear yang mewakili fungsi yang nonlinear di *input space*. Gambar 2.3 mendeskripsikan suatu contoh *feature mapping* dari ruang dua dimensi ke *feature space* dua dimensi. Dalam *input space*, data tidak dapat dipisahkan secara linear, tetapi dapat dipisahkan di *feature space*. Karena itu dengan memetakan data ke *feature space* menjadikan tugas klasifikasi menjadi lebih mudah.



Gambar 2.3 Kernel map yang mengubah masalah tidak linear menjadi linear



Fungsi kernel yang biasanya dipakai dalam literatur SVM adalah :

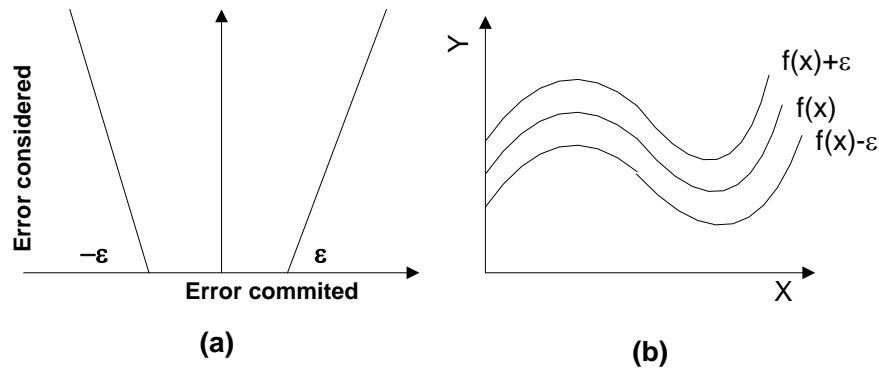
- a. linear :  $x^T x$
- b. Polynomial :  $(x^T x_i + 1)^p$ ,
- c. *Radial Basis Function* (RBF) :  $\exp\left(-\frac{1}{2\sigma^2} \|x - x_1\|^2\right)$ ,
- d. *Tangent hyperbolic* (sigmoid) :  $\tanh(\beta x^T x_i + \beta_i)$ , dimana  $\beta, \beta_i \in \mathfrak{R}$

Fungsi kernel mana yang harus digunakan untuk substitusi *dot product* di *feature space* sangat bergantung pada data. Biasanya metode *cross validation* digunakan untuk pemilihan fungsi kernel ini. Pemilihan fungsi kernel yang tepat adalah hal yang sangat penting karena fungsi ini akan menentukan *feature space* dimana fungsi klasiner akan dicari. Sepanjang fungsi kernelnya *legitimate*, SVM akan beroperasi secara benar meskipun kita tidak tahu seperti apa map yang akan digunakan. Fungsi kernel yang *legitimate* diberikan oleh Teori Mercer (Vapnik, 1995) dimana fungsi itu harus memenuhi syarat *continous and positive definite*. Lebih mudah menemukan fungsi kernel daripada mencari map  $\phi$  seperti apa yang tepat untuk melakukan mapping dari *input space* ke *feature space*. Pada penerapan metode kernel, kita tidak perlu tahu map apa yang digunakan untuk satu per satu data, tetapi lebih penting mengetahui bahwa *dot product* dua titik di *feature space* bisa digantikan oleh fungsi kernel.

SVR merupakan penerapan SVM untuk kasus regresi. Perbedaan antara SVM dan SVR adalah penerapan SVM dilakukan untuk menemukan fungsi pemisah (*hyperplane*) terbaik diantara fungsi yang tidak terbatas jumlahnya untuk memisahkan dua objek, misalkan objek tersebut akan termasuk dalam kelas -1 atau kelas +1 dengan cara memaksimalkan margin atau jarak antara dua set objek dari kelas yang berbeda. Sedangkan penerapan SVR dilakukan untuk menemukan suatu fungsi yang memiliki deviasi paling besar  $\epsilon$  dari target aktual  $y_i$ .

Dalam kasus regresi *output* berupa bilangan riil. SVR merupakan metode yang dapat mengatasi *overfitting*, sehingga akan menghasilkan performansi yang baik dan untuk mendapatkan suatu fungsi dengan tingkat kesalahan paling kecil sehingga menghasilkan sebuah prediksi yang bagus. SVR menentukan set data

yang dibagi menjadi set training dan set validasi. Kemudian dari set training tersebut ditentukan suatu fungsi regresi dengan batasan deviasi tertentu sehingga dapat menghasilkan prediksi yang mendekati target aktual.



Gambar 2.4. Batas *error* dalam *Support Vector Regression*

Gambar 2.4(a) menunjukkan bagaimana *error* dalam SVR dihitung. Sampai dengan garis batas *error*  $\epsilon$ , nilai *error* sama dengan 0, sedangkan di luar batas tersebut, nilai *error* akan dihitung sebagai *error-epsilon*. Oleh karena itu, dari sebuah garis, akan dibentuk sebuah tabung yang memiliki toleransi terhadap *error* seperti yang ditunjukkan pada gambar 2.4(b).

Fungsi regresi tersebut akan sempurna apabila batas deviasinya sama dengan 0 sehingga dapat dituliskan sebagai berikut :

$$f(x) = w^T \varphi(x) + b \quad (2.12)$$

dimana  $\varphi(x)$  menunjukkan suatu titik didalam *feature space*  $F$  yang merupakan hasil pemetaan  $x$  didalam *input space*. Koefisien  $w$  dan  $b$  diestimasi dengan cara meminimalkan fungsi resiko (*risk function*) yang didefinisikan dalam persamaan :

$$\min \frac{1}{2} \|w\|^2 C \frac{1}{\lambda} \sum_{i=1}^{\lambda} L_{\epsilon}(y_i, f(x_i)) \quad (2.13)$$

$$\text{subject to } \begin{cases} y_i - W\varphi(x_i) - b \leq \epsilon \\ W\varphi(x_i) - y_i + b \leq \epsilon, i = 1, 2, \dots, \lambda \end{cases}$$

Dimana :

$$L_{\epsilon}(y_i, f(x_i)) = \begin{cases} |y_i - f(x_i)| - \epsilon & |y_i - f(x_i)| \geq \epsilon \\ 0 & \text{untuk yang lain} \end{cases} \quad (2.14)$$

Faktor  $\|w\|^2$  inilah yang dinamakan regularisasi. Meminimalkan  $\|w\|^2$  akan membuat suatu fungsi setipis mungkin, sehingga bisa mengontrol kapasitas fungsi. Faktor lain yang dapat mendukung ketelitian fungsi ini adalah dengan memperhitungkan kesalahan empirik (*empirical error*) yang diukur dengan besarnya  $\varepsilon$ -insensitive loss function. Nilai  $\varepsilon$  mendefinisikan derajat toleransi terhadap *error*. Oleh karena itu untuk mendapatkan generalisasi yang baik dari fungsi regresi  $f$  maka dengan menggunakan ide  $\varepsilon$ -insensitive loss function bisa diminimalkan norm dari  $w$  seperti penyelesaian masalah optimasi berikut ini:

$$\min \frac{1}{2} \|w\|^2 \quad (2.15)$$

$$\text{Subject to } \begin{cases} y_i - W\varphi(x_i) - b \leq \varepsilon \\ w\varphi(x_i) - y_i + b \leq \varepsilon, i = 1, 2, \dots, \lambda \end{cases}$$

$$\begin{aligned} \sum e^2 &= \sum (y - \hat{y})^2 \\ &= \sum (y - (y - (b_0 + b_1x)))^2 \end{aligned}$$

Suatu fungsi dianggap layak dan baik apabila semua titik ada dalam rentang yang seharusnya, sebaliknya bila ada beberapa titik yang keluar dari rentang yang seharusnya maka fungsi tersebut tidak layak dilakukan untuk sebuah prediksi.

Jika diasumsikan bahwa ada suatu fungsi  $f$  yang dapat mengaproksimasi semua titik  $(x_i, y_i)$  dengan presisi  $\varepsilon$ . Apabila semua titik ada dalam rentang  $f \pm \varepsilon$  maka fungsi tersebut layak tetapi dalam hal ketidaklayakan berarti ada beberapa titik yang mungkin keluar dari rentang  $f \pm \varepsilon$ . Untuk mengatasi masalah ketidaklayakan pembatas dalam optimasi tersebut maka perlu ditambahkan sebuah variabel yang disebut variabel *slack* ( $\xi, \xi^*$ ) seperti pada gambar 3.

Selanjutnya masalah optimasi diatas bisa diformulasikan sebagai berikut :

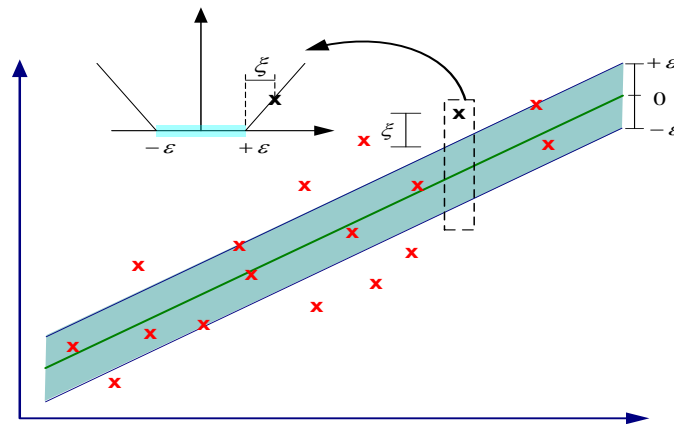
$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\lambda} (\xi_i + \xi_i^*) \quad (2.16)$$

$$\text{subject to } \begin{cases} y_i - w^t \varphi(x_i) - b - \xi_i \leq \varepsilon, i = 1, 2, \dots, \lambda \\ w\varphi(x_i) - y_i + b - \xi_i^* \leq \varepsilon, i = 1, 2, \dots, \lambda \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

Konstanta  $C > 0$  menentukan *trade off* antara ketipisan fungsi  $f$  dan batas atas deviasi lebih dari  $\varepsilon$  masih ditoleransi. Semua deviasi lebih besar dari pada  $\varepsilon$  akan dikenakan pinalti sebesar  $C$ .

Dalam SVR  $\varepsilon$  ekuivalen dengan akurasi dari aproksimasi terhadap data training. Nilai  $\varepsilon$  yang kecil terkait dengan nilai yang tinggi pada variabel *slack*  $\xi_i^*$  dan akurasi aproksimasi yang tinggi sebaliknya nilai yang tinggi untuk  $\varepsilon$  berkaitan dengan nilai  $\xi_i^*$  yang kecil dan aproksimasi yang rendah. Menurut persamaan 2.12, nilai yang tinggi untuk variabel *slack* akan membuat kesalahan empirik mempunyai pengaruh yang besar terhadap faktor regulasi.

Dalam SVR, *support vector* adalah data *training* yang terletak pada dan diluar batas  $f$  dari fungsi keputusan, karena itu jumlah *support vector* menurun dengan naiknya  $\varepsilon$ .



Gambar 2.5. Variabel *Slack* untuk mengatasi masalah ketidaklayakan pembatas

Fungsi yang menunjukkan hubungan antara *error* dengan bagaimana *error* ini dikenakan penalti disebut *Loss Function*. Perbedaan *loss function* akan menghasilkan formulasi SVR yang berbeda. Macam-macam *loss function* antara lain Quadratic, Laplace, Huber dan  $\varepsilon$ -insensitive (Gun, 1998). Yang paling sederhana adalah  $\varepsilon$ -insensitive. Dengan menggunakan *loss function*  $\varepsilon$ -insensitive, kita akan peroleh formulasi matematikanya sebagai berikut :

$$L_{\varepsilon}(y) = \begin{cases} 0, & \text{untuk } |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon, & \text{untuk yang lain} \end{cases} \quad (2.17)$$

### 2.2.3. Pengertian Prognosis dan Bantalan Gelinding

Prognosis berasal dari kata dalam bahasa Yunani *prognosticos*. *Pro* artinya sebelum dan *gnosis* artinya mengetahui. Hippocrates menggunakan kata prognosis, sama artinya dengan yang kita gunakan saat ini yaitu untuk menunjukkan atau meramalkan suatu kesalahan yang akan terjadi pada beberapa waktu kedepan.

Tingginya kompetisi pasar saat ini mengakibatkan banyaknya sektor industri saat ini yang mencoba meminimalkan modal dan biaya operasional dengan cara memanfaatkan siklus hidup dari mesin produksi tanpa mengorbankan pekerja (manusia), hasil produksi maupun keselamatan lingkungan. *Condition-based maintenance* adalah bagian yang paling berguna untuk memprediksi kegagalan suatu mesin atau peralatan dan untuk menghindari perawatan yang tidak perlu terhadap mesin. Prognosis merupakan kemampuan untuk menilai baik atau tidaknya kondisi suatu bagian dari mesin dan memprediksi sisa umur fungsi dari sebuah mesin dalam horison waktu yang tetap serta memprediksi waktu kegagalan dari sebuah mesin. Prognosis juga merupakan hal yang penting dalam keberadaan suatu mesin karena dengan adanya prognosis terhadap mesin, dapat meningkatkan keselamatan kerja, perencanaan misi yang baik, penjadwalan biaya pemeliharaan dan penggunaan waktu yang efisien.

Prognosis dari kinerja suatu mesin digunakan dengan tujuan untuk mendeteksi, menganalisis, dan memperbaiki masalah sebelum kegagalan terjadi. Prognosis lebih difokuskan pada pemantauan penurunan kinerja dan penilaian, sehingga kegagalan dapat diprediksi dan dicegah. Untuk memenuhi tujuan dari prognosis, ada tiga langkah penting yang harus dilakukan yaitu pertama, cacat atau kelainan yang terjadi pada kondisi mesin harus dapat dideteksi pada tahap awal. Kedua, mesin harus terus dipantau secara terus-menerus dan yang terakhir, sisa umur pakai dan kemungkinan kegagalan yang mungkin terjadi pada sebuah mesin harus dapat diprediksi secara efektif (Widodo, 2009).

*Bearing* atau bantalan gelinding adalah suatu komponen yang berfungsi untuk mengurangi gesekan pada mesin atau komponen-komponen yang bergerak dan saling menekan satu dengan yang lainnya.



Gambar 2.6. Bantalan Gelinding

Bila gerakan dua permukaan yang saling berhubungan terhambat, maka akan menimbulkan panas. Hambatan ini dikenal sebagai gesekan (*friction*). Gesekan yang terus menerus akan menyebabkan panas yang makin lama semakin meningkat dan menyebabkan keausan pada komponen tersebut. Gesekan yang tidak terkontrol dapat menyebabkan kerusakan pada komponen dan alat tidak bisa bekerja.

Bantalan gelinding digunakan untuk menahan atau menyangga komponen-komponen yang bergerak. Bantalan gelinding biasanya dipakai untuk menyangga perputaran pada *shaft*, dimana terjadi sangat banyak gesekan. Fungsi dari Bantalan gelinding antara lain mengurangi gesekan, panas dan aus, menahan beban *shaft* dan mesin, menahan *radial load* dan *thrust load*, menjaga toleransi kekencangan serta mempermudah pergantian dan mengurangi biaya operasional.