BAB II

TINJAUAN PUSTAKA

2.1. Pendahuluan

Mesin merupakan alat mekanik atau elektrik yang bisa mengirim atau mengubah untuk melakukan dan melaksanakan tugas manusia. Mesin terdiri dari beberapa elemen yang masing-masing elemennya memiliki fungsi tersendiri agar mesin tersebut bisa berjalan dengan lancar. Dalam penggunaan elemen mesin bisa berfungsi sebagai elemen pengikat, elemen transmisi, elemen penyangga dan sebagainya. Salah satu elemen penyangga yaitu *bearing*. *Bearing* adalah elemen mesin yang mendukung elemen mesin lainnya bergerak, berfungsi sebagai pembatas gerak *relative* antara dua atau lebih komponen agar selalu bergerak pada arah yang diinginkan.

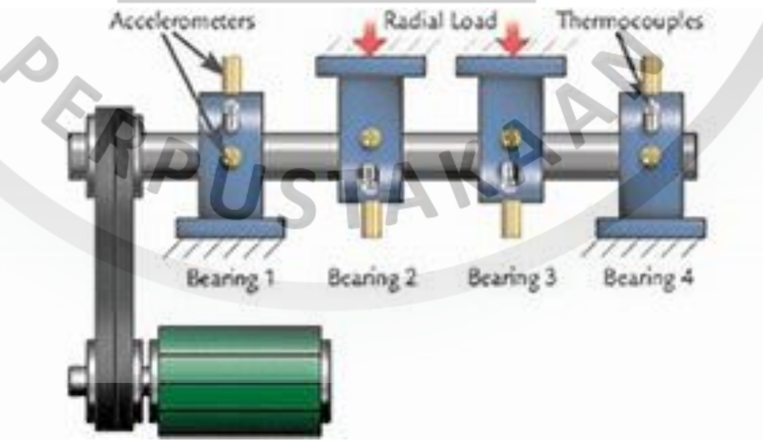
2.1.1. Pengertian *Bearing*

Bearing dalam bahasa Indonesia berarti *Bearing*. Menurut Sularso (1997), *bearing* adalah elemen mesin yang mampu menumpu poros berbeban, sehingga gesekan bolak-baliknya dapat berlangsung secara halus, aman dan panjang usia pemakaiannya. *Bearing* harus cukup kokoh untuk memungkinkan poros suatu mesin bekerja dengan baik. Menurut Shigley dan Mitchell (1983), *bearing* dibuat untuk menerima beban radial murni, beban aksial murni, atau gabungan kedua-duanya. Gambar 2.1 merupakan gambar untuk salah satu contoh *bearing.*

Gambar 2.1 *Bearing*

Dalam ilmu mekanika *bearing* adalah sebuah elemen mesin yang berfungsi untuk membatasi gerak *relative* antara dua atau lebih komponen mesin agar selalu bergerak pada arah yang diinginkan. Gambar 2.2 adalah alat yang digunakan untuk mengukur vibrasi *bearing* dengan cara kerja rotor diputar oleh motor listrik, rotor ditumpu oleh 4 buah *bearing*. Vibrasi diukur dengan tiga buah sensor akselerometer yang dipasang pada *bearing* secara *horizontal*, *vertical*, dan *aksial* yang seperti disajikan pada Gambar 2.2 sebagai berikut:

Gambar 2.2 Alat pengukur vibrasi *bearing*



Gambar 2.3 Ilustrasi akselerometer pada *bearing*

Umur dari suatu *bearing* dinyatakan sebagai jumlah putaran total, atau jumlah jam pada suatu kecepatan putar tertentu, dari operasi *bearing* diperlukan untuk mengembangkan kriteria kegagalan. Dibawah kondisi ideal kegagalan lelah akan berupa penghancuran permukaan yang menerima beban. *Standard The Anti-Fiction Bearing Menufactures Association* (AFBMA) menyatakan bahwa kriteria kegalan adalah suatu bukti awal dari kelelahan logam.

2.1.2. Jenis-jenis *Bearing*

Berikut merupakan jenis-jenis *bearing* berdasarkan gerakan terhadap poros.

1) *Rolling Bearing*

Pada *bearing* ini terjadi gesekan gelinding antara bagian yang berputar dengan yang diam melalui elemen gelinding seperti bola (peluru), rol atau rol jarum dan rol bulat. Berikut merupakan sifat-sifat dari *rolling bearing*:

a) Gesekan mula yang jauh lebih kecil dan pengaruh yang lebih kecil dari jumlah putaran terhadap gesekan.
b) Gesekan kerja lebih kecil sehingga penimbulan panas lebih kecil pada pembebanan yang sama.
c) Penurunan waktu pemasukan dan pengaruh dari bahan poros.
d) Pelumasan terus menerus dengan jumlah bahan pelumas yang jauh lebih sedikit.
e) Memiliki kemampuan dukung yang lebih besar setiap lebar *bearing*.

Meskipun *bearing* gelinding menguntungkan, orang tetap menggunakan *sliding bearing* dalam hal tertentu, contohnya kalau kebisingan *bearing* mengganggu, selanjutnya pada kejutan yang kuat dalam putaran bebas, contohnya pada mesin cadangan; kemudian pada *bearing* yang dipecah-pecah, pada *bearing* radial yang sangat besar(harganya relatif mahal), pada *bearing* aksial kecepatan tinggi dari generator dan turbin dan *bearing* yang sangat kecil dalam peralatan yang murah (contohnya *bearing* luncur-plastik dalam mesin rumah tangga).

Jenis-jenis dari *rolling bearing* yaitu *bearing* peluru, *bearing* silinder, *bearing* jarum, *bearing* rol kerucut, *bearing* rol bulat simetris, dan *bearing* rol bulat tidak simetris.

2) *Sliding Bearing*

Pada *bearing* ini terjadi gesekan luncur antara poros dan *bearing* karena permukaan poros ditumpu oleh permukaan *bearing* dengan perantaraan lapisan pelumas. Poros yang ber*bearing* luncur mencapai putaran yang paling tinggi. Berikut merupakan sifat-sifat dari *bearing* luncur:

a) Tingkat kebisingan lebih kecil.
b) Putaran tinggi.
c) Ketepatan tinggi.
d) Guncangan dan getaran kuat.
e) *Bearing*-nya berdiameter kecil.

Jenis-jenis *bearing* luncur terbagi menjadi beberapa klasfikasi yaitu arah gaya, penggunaan, desain, bahan, dan pelumas. Berikut merupakan jenis-jenis dari *sliding bearing*:

a) Menurut arah gaya: *Bearing radial* dan *bearing aksial.*
b) Menurut penggunan: *Bearing* mesin perkakas, *bearing* kotak roda gigi, *bearing* motor, *bearing* transmisi, *bearing* turbin, *bearing* pekerjaan gilas.
c) Menurut desain: *Bearing* mata, *bearing* penutup, *bearing* tetap, *bearing* gantung, *bearing* ayun, *bearing* kotak, *bearing* cakram, *bearing* terpasang.

d) Menurut bahan: *Bearing* logam putih, *bearing* perunggu, *bearing* besi tuang merah, *bearing* logam ringan, *bearing* logam sinter, *bearing* bahan pres, *bearing* berbahan banyak.

e) Menurut pelumas: *bearing* gemuk, *bearing* minyak, *bearing* air, *bearing* udara, *bearing* pelumasan cincin, *bearing* pelumasan aliran, *bearing* hidrostatis atau aerostatis (Niemann, 1981).

2.1.3. Penyebab Kerusakan *Bearing*

Kerusakan pada *bearing* dapat disebabkan oleh:

a) Karena kesalahan bahan (penggerusan, pengaluran, pelapisan dengan plat), kesalahan pembuatan (peretakan berat atau halus, kesalahan toleransi atau kesalahan celah *bearing*, panas-bising).

b) Karena kesalahan pemasangan: Pasan terlalu longgar (cincin-cincin terputar, kehausan pasan), pasan terlalu erat (ventilasi kurang, suhu meningkat, penengangan lebih), pembenjolan (kehausan jalur jalan dan rol atau pencekungan), montasi salah (penekanan yang tertinggal, pematahan).

c) Karena kesalahan operasi: bahan pelumas yang tidak sesuai (korosi, pelarutan), pengotoran (kehausan, berjalan bising, penekanan dalam jalur rol dan badan gelinding, peremukan), guncangan pada putaran bebas dan pelintasan arus (berbagai jenis dari pembentukan lekukan-lekukan).

2.2. *Root Mean Square* (RMS)

Kerusakan mesin akan meningkatkan nilai vibrasi mesin. Beberapa ciri ekstraksi statistik dapat menjadi sangat signifikan pada penggunaan teknik analisa. Kondisi kerusakan mesin dapat dibedakan menggunakan representasi kuantitatif ciri domain waktu. Salah satu ciri ekstraksi statistik adalah *Root Mean Square* (RMS). *Root Mean Square* (RMS) merupakan ciri statistik yang mengukur komposisi karakteristik energi dari sinyal vibrasi. Ciri ini baik dalam mengidentifikasi tingkat kebisingan secara keseluruhan. Berikut merupakan rumus untuk RMS:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{2.1}$$

2.3. Analisis Survival

2.3.1. Pengertian Analisis Survival

Analisis survival merupakan salah satu teknik statistik untuk menganalisis data dengan variabel *outcame* yang akan diteliti adalah waktu (*time)* hingga suatu peristiwa yang muncul. Analisis survival merupakan analisis statistika khusus yang membantu menganalisis suatu kasus yang tidak dapat diselesaikan dengan analisis statistika pada umumnya. Analisis ini digunakan ketika kasus berkaitan dengan waktu dan lama waktu hingga terjadi peristiwa tertentu dan kemungkinan adanya data tersensor merupakan karakteristik khas yang membedakan dengan analisis lain. Misalnya peristiwa timbulnya suatu penyakit, kambuhnya penyakit, kesembuhan dan kematian (Kleinbaum dan Klein, 2012). Dalam analisis survival, sering menggunakan istilah '*failure'* untuk menentukan terjadinya suatu peristiwa kegagalan dan istilah '*survival time'* untuk menentukan lamanya waktu hingga terjadinya kegagalan. Analisis survival bertujuan untuk menaksir probabilitas survival, kematian, kegagalan, dan juga peristiwa lainnya sampai periode waktu tertentu. Perbedaan antara analisis survival dengan analisis statistik lainnya adalah adanya data tersensor. Data dikatakan tersensor jika pengamatan waktu survival hanya sebagian, tidak sampai *failure time*.

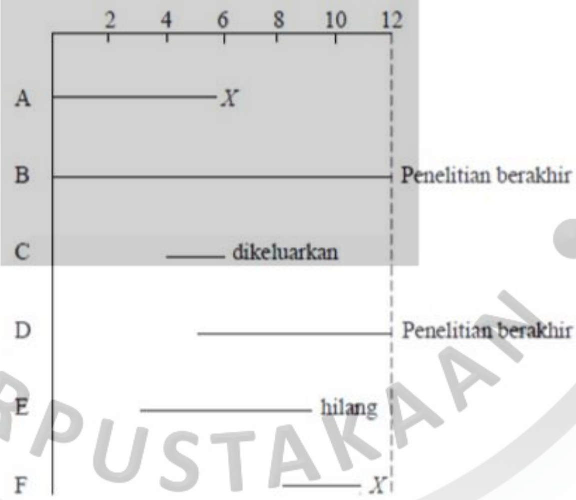
Menurut Collect (2003) dalam menentukan waktu survival, terdiri dari tiga faktor, yaitu:

1. Waktu awal pencatatan (*start point*) yang merupakan waktu awal dimana dilakukannya pencatatan untuk menganalisis suatu kejadian.

2. Waktu akhir pencatatan (*end point)* yang merupakan waktu dimana pencatatan berakhir. Waktu ini berguna untuk mengetahui status tersensor atau tidak tersensor seorang pasien untuk bisa melakukan analisis.
3. Dan skala pengukuran sebagai batas dari waktu kejadian dari awal sampai akhir kejadian. Skala dihitung dalam hari, minggu, atau tahun.

2.3.2. Penyensoran

Penyensoran adalah salah satu langkah yang harus dilakukan untuk mengatasi ketidaklengkapan suatu data pengamatan. Data dikatakan tersensor apabila data tidak dapat diamati secara lengkap karena subjek penelitian hilang atau mengundurkan diri atau sampai akhir penelitian subjek tersebut belum mengalami kejadian tertentu, sedangkan data yang dapat diamati secara lengkap sampai penelitian berakhir disebut data yang tidak tersensor.



Gambar 2.4 Grafik data tersensor

Gambar 2.4 merupakan gambar pencatatan sebuah kejadian dari awal penelitian sampai akhir waktu penelitian. Skala waktu diatas berdasarkan minggu dan setiap individu memiliki *failure* yang berbeda-beda dalam pencatatan.

Menurut Kleinbaum dan Klein (2005) tiga penyebab data dikatakan tersensor antara lain:

1. *Loss to follow up,* yaitu subjek menghilang selama masa pengamatan, missal subjek pindah atau menolak untuk diamati.
2. Subjek tidak mengalami kejadian selama penelitian.
3. Subjek terpaksa diberhentikan dari pengamatan karena meninggal sebelum pengamatan berakhir atau alasan lain.

Pada *threshold* untuk menentukan *time-to-failure* pada *bearing* dalam analisis yang dilihat dari plot nilai RMS terhadap nilai *measurement point*-nya, menurut Widodo dan Yang (2011) jika RMS melebihi *threshold* 1 maka menandakan bahwa *bearing* rusak sedangkan jika RMS tidak melebihi *threshold* 1 maka menandakan bahwa *bearing* tidak rusak atau tersensor. Menurut Collect (2003) dalam analisis survival terdapat 3 tipe penyensoran yaitu:

a. Sensor kanan (*right censoring)*

Sensor yang terjadi dikarenakan objek pengamatan belum mengalami kejadian hingga akhir periode pengamatan, sedangkan waktu awal dari objek pengamatan dapat diamati secara penuh. Atau waktu survivalnya lebih lama daripada waktu sensor. Misalkan suatu individu diamati selama lima tahun dari awal pengamatan, kemudian pada tahun ketiga individu tersebut pindah ke wilayah lain dan tidak dapat diamati lagi (*lost to follow up)*. Individu ini memiliki waktu survival dalam penelitian setidaknya dua tahun, sehingga waktu pengamatan individu tersebut dikatakan tersensor kanan.

b. Sensor kiri (*left censoring)*

Sensor yang terjadi dikarenakan waktu awal dari subjek pengamatan tidak terdapat teramati pada awal pengamatan, sementara kegagalan dapat diamati secara penuh sebelum penelitian berakhir. Contohnya, peneliti mengamati pasien penyakit kanker, peneliti dapat mencatat kejadian tepatnya seseorang tersebut positif kanker di tes pertamanya, namun peneliti tidak memiliki catatan tentang waktu tepatnya seseorang tersebut mulai mengidap kanker, dengan demikian pasien kanker tersebut tersensor kiri.

c. Sensor interval (*interval censoring)*

Sensor interval adalah sensor yang waktu survivalnya berada dalam suatu selang tertentu. Sebagai contohnya, jika catatan medis menunjukkan bahwa pada usia 45 tahun pasien yang mengidap kanker masih berkondisi sehat dan belum mengidap kanker, kemudian pasien tersebut melakukan tes pertama saat berumur 50 tahun dan terdiagnosis terkena penyakit kanker, dengan demikian usia saat didiagnosis posistif kanker antara 45 sampai 50 tahun.

2.3.3. Fungsi Survival

Penaksir peluang hidup dapat digunakan untuk membantu menaksirkan daya tahan hidup suatu unit atau individu pada suatu keadaan tertentu usia manusia untuk hidup, sebagai landasan perhitungan premi dalam asuransi, dan menaksir pertumbuhan atau pengurangan populasi. Alat untuk menaksir peluang hidup dikenal dengan fungsi survival. Biasanya fungsi tersebut dibahas dalam dunia asuransi jiwa (Keinbaum dan Klein, 2012). Misalkan waktu ketahanan T mempunyai fungsi distribusi peluang dengan fungsi densitas $f(x)$. Fungsi distribusi kumulatif bagi T ditulis sebagai berikut:

$$F(t) = P(T < t) = \int_0^t f(x)\, dx \tag{2.2}$$

Menyatakan peluang waktu ketahanan hidup bernilai lebih kecil dari t. Sesuai dengan definisi fungsi distribusi kumulatif $F(t)$ dari T, fungsi survival dari T adalah sebagai berikut:

$$S(t) = P(T \geq t) = 1 - F(t) \tag{2.3}$$

Menyatakan peluang individu bertahan melebihi waktu t yakni $S(t)$ adalah peluang bahwa variabel acak T melebihi t.

Fungsi survival juga dapat dinyatakan dalam bentuk fungsi kepadatan peluang yaitu:

$$S(t) = P(T \geq t) = \int_t^\infty f(x)\, dx \tag{2.4}$$

Diperoleh hubungan antara fungsi kepadatan peluang, fungsi distribusi kumulatif dari T, dan fungsi survival yaitu:

$$F(t) = 1 - S(t) = \frac{d\big(F(t)\big)}{dt} = \frac{d\big(1 - S(t)\big)}{dt} = F'(t) = f(t) = -S'(t) \tag{2.5}$$

2.3.4. Distribusi Weibull

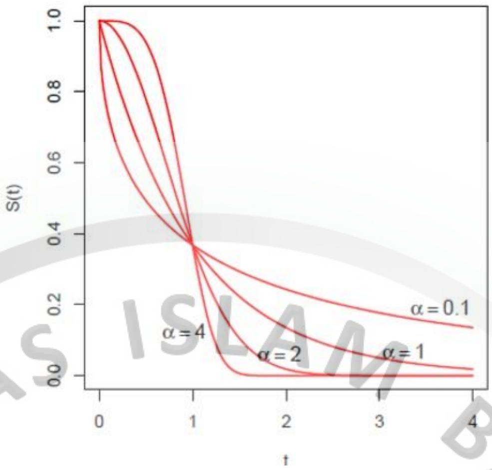
Distribusi Weibull banyak digunakan dalam analisis reliabilitas yang berkaitan dengan umur (rentang waktu). Contohya rentang waktu dimana sebuah mesin mungkin akan rusak (tidak berfungsi). Variabel random kontinyu T berdistribusi Weibull, dengan parameter $\alpha > 0$ dan $\beta > 0$ dengan $t > 0$, dimana fungsi densitas peluangnya adalah sebagai berikut:

$$f(t) = \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^{\alpha}} \tag{2.6}$$

Berdasarkan persamaan (2.6) dapat dicari fungsi survival distribusi Weibull sebagai berikut:

$$S(t) = \int_t^\infty f(x)\, dx = \int_t^\infty \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}} dx = -\, e^{-\left(\frac{x}{\beta}\right)^{\alpha}}\Bigg|_t^\infty = e^{-\left(\frac{t}{\beta}\right)^{\alpha}} \tag{2.7}$$

Berikut merupakan contoh kurva survival distribusi Weibull dengan beberapa nilai parameter α pada Gambar 2.5 sebagai berikut:



Gambar 2.5 Contoh kurva distribusi Weibull

2.3.5. Estimasi Kaplan-Meier

Estimasi fungsi survival dengan menggunakan metode Kaplan-Meier disebut juga estimasi *product limit*. Kaplan dan Meier (1985) adalah orang pertama yang membahas estimasi fungsi ini. Penaksir Kaplan-Meier (1985) sangat popular untuk analisis survival yang paling cocok digunakan ketika ukuran sampel kecil. Lama pengamatan masing-masing subjek disusun dari yang terpendek sampai yang terpanjang dengan catatan subjek yang tersensor diikutsertakan. Pada dasarnya metode Kaplan-Meier sama dengan metode *life-table*, metode ini memiliki kelebihin yaitu dapat memberikan proporsi ketahanan hidup yang pasti karena menggunakan waktu ketahanan hidup secara tepat bukan berdasarkan kelas interval.

Kaplan-Meier adalah komputasi stastistik untuk menghitung peluang survival. Metode ini didasarkan pada waktu kelangsungan hidup individu dan mengasumsikan bahwa data data sensor adalah independen berdarkan waktu kelangsungan hidup.

Estimasi Kaplan-Meier untuk fungsi survival didefinisikan sebagai:

$$\hat{S}(t) = \prod_{j=1}^{m} \left(\frac{n_j - d_j}{n_j} \right) \tag{2.8}$$

Keterangan: d_j : banyaknya kematian pada saat $t_j (j = 1,2, \ldots, m)$

n_j: banyaknya individu yang masih hidup sesaat sebelum t_j jumlah individu berisiko pada saat t_j) termasuk yang meninggal pada saat t_j dengan $\hat{S}(t) = 1$ untuk $t < t_1$ dan $t_{r+1} = \infty$.

Jika unit j telah mencapai kegagalan sebelum perbaikan atau penggantian pada mesin, maka probabilitas kelangsungan hidup ditetapkan dengan nilai "1" atau kelangsungan hidupnya sebesar 100% hingga waktu kegagalan, berikut merupakan fungsi survivalnya.

$$\hat{S}(t + k\Delta) = \begin{cases} 1, & 0 \leq t + k\Delta < T \\ 0, & t + k\Delta > T \end{cases} \tag{2.9}$$

Dimana T adalah waktu kegagalan, k adalah jumlah interval masa depan dan Δ adalah interval waktu tetap, dimana pada penelitian ini ditetapkan $\Delta = 1$.

Kumpulan data dianggap tersensor jika komponen mesin belum mencapai *Threshold* saat dikeluarkan dari mesin. Dalam hal ini, rumus standar estimasi Kaplan-Meier dimodifikasi untuk komponen mesin individu/unit.

$$\hat{S}(t + k\Delta) = \begin{cases} 1, & 0 \leq t + k\Delta < L \\ \displaystyle\prod_{j=1}^{m} \left(\frac{n_j - d_j}{n_j} \right), & t + k\Delta > L \end{cases} \tag{2.10}$$

Dimana L merupakan waktu bertahan hidup terakhir yang diamati dari komponen unit mesin.

2.3.6. Estimasi Survival *Probability Density Fucntion* (PDF)

Estimasi Survival *Probability Density Function* (PDF) digunakan untuk memperkirakan fungsi survival dari masing-masing unit j. Dalam hal ini, estimasi probabilitas survival merupakan perkalian berturut-turut unit yang telah bertahan pada interval sebelumnya, dan memiliki keadaan indeks yang lebih tinggi dari indeks yang diamati item j tetapi lebih rendah dari *threshold*. Estimasi Kegagalan *Probability Density Function* (PDF) didefinisikan sebagai berikut:

$$\hat{S}(t+k\Delta) = \prod_{j=1}^{m} \frac{pdf_{obs}}{pdf_{ovr}} = \prod_{j=1}^{m} \frac{\int_{y_{i,t+k\Delta}}^{Y_{threshold}} f(y|t+k\Delta)dy}{\int_{y_{i,t+k\Delta}}^{\infty} f(y|t+k\Delta)dy} \tag{2.11}$$

Keterangan: k: angka untuk interval masa depan

Δ: waktu tetap interval

y_i: kondisi indeks obesrvasi, $i = 1,2, \dots, m$

Kondisi indeks observasi merupakan kondisi nilai yang mewakili keseluruhan objek sebelum menyentuh *Threshold.* Pada penelelitian kali ini, untuk mengetahui kondisi indeks observasi, seluruh objek yang dibawah *Threshold* akan dicari nilai rata-rata-ratanya.

Dimana pdf_{obs} merupakan integral dari PDF antara kondisi indeks observasi pada unit i dan *Threshold*, sedangkan pdf_{ovr} adalah integral dari PDF semua nilai yang mungkin sama atau lebih tinggi dari kondisi indeks observasi pada unit i.

Heng (2008) dalam Widodo dan Yang (2011) mengatakan bahwa estimasi akhir untuk probabilitas survival adalah rata-rata dari kedua estimasi survival diatas, yaitu estimasi Kaplan-Meier dan estimasi kegagalan PDF.

2.4. *Support Vector Regression* (SVR)

Support Vector Regression (SVR) merupakan pengembangan dari *Support Vector Machine* (SVM) untuk kasus regresi. Tujuan dari SVR adalah untuk menemukan sebuah fungsi *f(x)* sebagai suatu *hyperplane* (garis pemisah) berupa fungsi regresi yang mana sesuai dengan semua input data dengan sebuah *error* ε dan membuat ε setipis mungkin (Scholkopf dan Smola, 2002). SVR bisa menentukan suatu fungsi $f(x)$ yang mempunyai deviasi ε paling besar dari target, untuk semua data *training*. Jika nilai ε sama dengan 0 maka diperoleh suatu persamaan regresi yang sempurna (Santosa, 2007).

Menurut Abe (2005) dalam Septiningrum dkk. (2015), tujuan dari SVR ini adalah untuk memetakan vektor input ke dalam dimensi yang lebih tinggi. Misalkan sebuah fungsi berikut adalah garis regresi sebagai *optimal hyperplane*:

$$f(x) = w^T \varphi(x) + b \tag{2.12}$$

Keterangan: w: vektor bobot berdimensi l (l = banyaknya data *training*)

$\varphi(x)$: fungsi yang memetakan x pada ruang dimensi l

b: bias

Masalah optimasi pada persamaan *hyperplane* SVR dilakukan dalam bentuk *Quadratic Programming:*

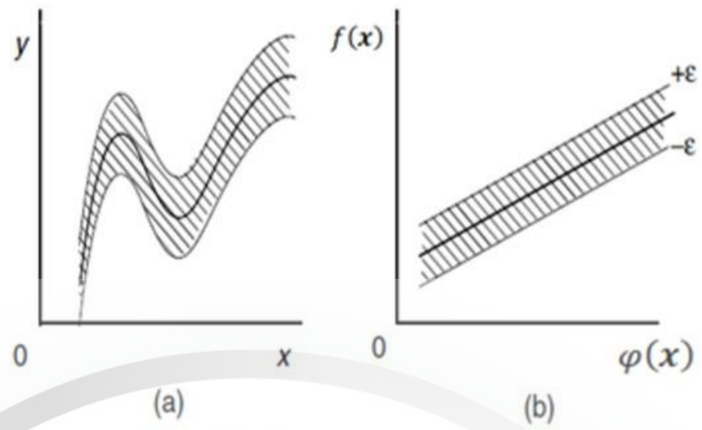
$$min \frac{1}{2} \|w\|^2 \tag{2.13}$$

Dengan syarat:

$$\begin{gathered} y_i - w^T \varphi(x_i) - b \leq \varepsilon \\ w^T \varphi(x_i) - y_i + b \leq \varepsilon \end{gathered} \tag{2.14}$$

Keterangan: x_i: vektor input dengan $i = 1,2, \dots, l$

y_i: output scalar dengan $i = 1,2, \dots, l$



Gambar 2.6 *Intensive zone* (a) *original input space* dan (b) *feature space*

Faktor $\|w\|^2$ dinamakan regulasi. Menimimalkan $\|w\|^2$ akan membuat suatu fungsi setipis (*flat*) mungkin, sehingga bisa mengontrol kapasitas fungsi (*function capacity*). Pada persamaan (2.13) diasumsikan bahwa semua titik ada dalam rentang $f(x) \pm \varepsilon$ (*feasible*), dalam hal ketidaklayakan (*infiesibility*), dimana ada beberapa titik yang mungkin keluar dari rentang $f(x) \pm \varepsilon$ maka ditambahkan variabel *slack* ξ dan ξ^* untuk mengatasi masalah pembatasan yang tidak layak (*infeasible constraints*) dalam masalah optimasi (Santosa, 2007).

Semua titik yang berada diluar margin akan dikenai pinalti. Selanjutnya masalah optimasi diatas bisa diformulasikan sebagai berikut:

$$min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{2.15}$$

Dengan syarat:

$$\begin{aligned} & y_i - w^T\varphi(x_i) - b - \xi_i \le \varepsilon, \qquad i = 1,2,\dots,l \\ & w^T\varphi(x_i) - y_i - b - \xi_i^* \le \varepsilon, \qquad i = 1,2,\dots,l \\ & \xi_i, \xi_i^* \ge 0 \end{aligned} \tag{2.16}$$

Solusi optimasi untuk persaman (2.15) dengan batas bawah (2.16) adalah dengan fungsi *Lagrange* berikut:

$$L(w,b,\xi_i,\xi_i^*,\alpha,\alpha^*,\eta,\eta^*) = \ min\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i+\xi_i^*) - \sum_{i=1}^{l}\alpha_i(\varepsilon+\xi_i-y_i+w^T\varphi(x)+b) - \sum_{i=1}^{l}\alpha_i^*(\varepsilon+\xi_i^*-y_i+w^T\varphi(x)+b) - \sum_{i=1}^{l}(\eta_i\xi_i+\eta_i^*\xi_i^*) \quad (2.17)$$

L dinamakan *Lagrange*, $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$ adalah *Lagrange Multiplier*. *Lagrange Multiplier* digunakan untuk menyelesaikan masalah optimasi dengan kendala (*constrined optimization*). Dimana kendalanya akan dikonversi menjadi masalah optimasi tanpa kendala (*unconstrained optimization*).

Untuk mencari nilai parameter persamaan *optimal hyperlane* dapat ditentukan dari turunan *Lagrange Multiplier* terhadap w, b, ξ_i, ξ_i^*.

$$\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0 \quad (2.18)$$

$$\frac{\partial L}{\partial b} = \sum_{i-1}^{l}(\alpha_i^*-\alpha_i) = 0 \quad (2.19)$$

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)\varphi(x_i) = 0 \quad (2.20)$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \eta_i = 0 \quad (2.21)$$

$$\frac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \quad (2.22)$$

Dari persamaan (2.20) *w* dapat ditulis dengan $w = \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)\varphi(x_i)$, sehingga fungsi regresi secara eksplisit dirumuskan sebagai berikut:

$$f(x) = \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)\,K(x_i,x)+b \quad (2.23)$$

Dimana hasil selisih antara α_i dan α_i^* menghasilkan nilai beta.

Optimasi permasalahan dual adalah:

$$Q(\alpha, \alpha^*) = -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\varphi^T(x_i)\varphi(x_j) -$$

$$\varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) \quad (2.24)$$

Menurut Scholkopf dan Smola (2003) solusi optimal untuk bias (b) dapat dihitung menggunakan KKT (Kaush-Kuhn-Tucker) sebagai berikut:

$$\begin{gathered} \alpha_i(\varepsilon + t_i - y_i + w^T\varphi(x_i) + b) = 0 \\ \alpha_i^*(\varepsilon + t_i^* - y_i + w^T\varphi(x_i) - b) = 0 \\ (C - \alpha_I)t_i = 0 \\ (C - \alpha_i^*)t_i^* = 0 \end{gathered} \quad (2.25)$$

Sehingga didapat:

$$\begin{gathered} b = y_i - w^T\varphi(x_i) - \varepsilon \text{ untuk } 0 < \alpha_i < C \\ b = y_i - w^T\varphi(x_i) + \varepsilon \text{ untuk } 0 < \alpha_i < C \end{gathered} \quad (2.26)$$

Data *training* yang memiliki nilai $\alpha_i > 0$ merupakan sebuah *support vector* sedangkan sisanya memiliki nilai $\alpha_i = 0$. Dengan demikian fungsi keputusan yang dihasilkan hanya dipengaruhi oleh *support vector*.

2.5. Fungsi Kernel

Menurut Santosa (2007) banyak teknik data *mining* yang dikembangkan dengan asumsi kelinearan, sehingga algoritma yang dihasilkan terbatas untuk kasus-kasus linier. Secara umum, kasus-kasus di dunia nyata adalah kasus yang tidak linear sehingga untuk mengatasi masalah ketidaklinearan yang sering terjadi, dapat diterapkan fungsi kernel. Dengan fungsi kernel suatu data x di *input space* dipetakan ke *feature space* dengan dimensi yang lebih tingga melalui φ.

$$\varphi: x \rightarrow \varphi(x)$$

Nilai $K(x_i, x)$ merupakan fungsi kernel yang menunjukan pemetaan linier pada *feature space*. Perlu dijelaskan disini bahwa nilai $K(x_i, x)$ tidak selalu bisa diekspresikan secara eksplisit sebagai kombinasi antara α, y dan $\varphi(x)$, karena dalam banyak kasus $\varphi(x)$ tidak diketahui atau sulit dihitung.

Fungsi kernel yang akan digunakan dalam penelitian ini adalah:

1. Kernel Polynomial

$$\varphi(x) = K(x_i, x) = \left((x^T x) + 1\right)^d, \;\; d > 0 \tag{2.27}$$

2. Kernel Gaussian RBF

$$\varphi(x) = K(x_i, x) = exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right), \;\; \sigma > 0 \tag{2.28}$$

2.6. *Root Mean Square Error* (RMSE)

Root Mean Square Error (RMSE) adalah metode alternatif untuk mengevaluasi teknik prediksi. RMSE merupakan nilai rata-rata dari jumlah kuadrat kesalahan, juga dapat menyatakan ukuran besarnya kesalahan yang dihasilkan oleh suatu model prakiraan. Nilai RMSE rendah menunjukkan bahwa variasi nilai yang dihasilkan oleh suatu model prakiraan mendekati variasi nilai observasinya. Berikut merupakan rumus RMSE yang ditulis dalam persamaan berikut ini:

$$RMSE = \left(\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2\right)^{\frac{1}{2}} \tag{2.29}$$

2.7. Koefisien Determinasi (R^2)

Koefisien determinasi (R^2) merupakan salah satu ukuran untuk mengukur/melihat kualitas dari suatu model. Semakin besar nilai R^2 maka model semakin baik. Koefisien determinasi juga dapat menunjukkan tingkat akurasi. Koefisien determinasi merupakan nilai dari koefisien korelasi R yang dikuadratkan. Berikut merupakan rumus koefisen determinasi:

$$Koefisien\ determinasi = R^2$$

$$R = \frac{Cov(y_t, \hat{y}_t)}{S_{y_t} S_{\hat{y}_t}} \tag{2.30}$$