

Prediksi *Failure Degradation Bearing* dengan Pendekatan Gabungan Model

Syamil Anwa Zahra Zakaria *, Sutawanir Darwis

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,,
Universitas Islam Bandung, Indonesia.

syamilanwazahra@gmail.com, std.darwis@gmail.com

Abstract. The reliability of machine components is a very important aspect to consider, one of the critical components in a machine is the bearing, which serves as a support and allows rotating movement between two components. Failure in bearings can result in serious damage to the machine and stop operations, this has a major impact on productivity and maintenance costs. Predictions were made using the Support Vector Regression method with the Cox Proportional Hazard model as the target vector. The Cox Proportional Hazard model is used to see the effect of covariates on bearing survival time, while Support Vector Regression with RBF kernel is used to predict survival chances based on the risk from the Cox model. In this study, the Weibull distribution hazard function is used as the baseline hazard function in the Cox model, to calculate the value of the shape and scale parameters in the Weibull distribution, the Least Squares method is used until the respective values of the shape parameter and scale parameter are 1.02045 and 7.3732. The results show that combining the Cox Proportional Hazard model and Support Vector Regression can provide a high average accuracy value of 99.83% on training data and 98.82% on testing data. As in the journal article (Caesarendra et al., 2010) with a high accuracy value, it can be said that the model can be used.

Keywords: *Bearing, Support Vector Regression (SVR), Model Cox Proportional Hazard (CPH).*

Abstrak. Keandalan komponen mesin menjadi aspek yang sangat penting untuk diperhatikan, salah satu komponen kritis dalam mesin adalah *bearing*, yang berfungsi sebagai penopang dan memungkinkan pergerakan berputar antara dua komponen. Kegagalan dalam *bearing* dapat mengakibatkan kerusakan serius pada mesin dan menghentikan operasi, hal ini berdampak besar pada produktivitas dan biaya pemeliharaan. Dilakukan prediksi menggunakan metode *Support Vector Regression* dengan model *Cox Proportional Hazard* sebagai target vektornya. Model *Cox Proportional Hazard* digunakan untuk melihat pengaruh dari kovariat terhadap waktu *survival bearing*, sedang *Support Vector Regression* dengan kernel RBF digunakan untuk memprediksi peluang *survival* berdasarkan risiko dari model Cox. Dalam penelitian ini digunakan fungsi *hazard* berdistribusi Weibull sebagai fungsi *baseline hazard* dalam model Cox, untuk menghitung nilai parameter bentuk dan skala pada distribusi Weibull digunakan metode Kuadrat Terkecil, hingga diperoleh nilai masing-masing dari parameter bentuk dan parameter skala adalah 1,02045 dan 7,3732. Hasil menunjukkan bahwa kombinasi model *Cox Proportional Hazard* dan *Support Vector Regression* dapat memberikan nilai rata-rata akurasi yang tinggi sebesar 99,83% pada data *training* dan 98,82% pada data *testing*. Seperti pada artikel jurnal (Caesarendra et al., 2010) dengan nilai akurasi yang tinggi, maka dapat dikatakan bahwa model dapat digunakan.

Kata Kunci: *Bearing, Support Vector Regression (SVR), Model Cox Proportional Hazard (CPH).*

A. Pendahuluan

Dalam dunia industri, keandalan komponen mesin menjadi aspek yang sangat penting untuk diperhatikan. salah satu komponen kritis dalam mesin adalah *bearing*, yang berfungsi sebagai penopang dan memungkinkan pergerakan berputar antara dua komponen. Kegagalan dalam *bearing* dapat mengakibatkan kerusakan serius pada mesin dan menghentikan operasi, hal ini berdampak besar pada produktivitas dan biaya pemeliharaan. Distribusi Weibull merupakan salah satu model yang paling umum digunakan dalam analisis keandalan dan analisis *survival*. Fleksibilitas distribusi Weibull memungkinkan model ini untuk menangani berbagai bentuk data kegagalan, mulai dari data yang menunjukkan kegagalan dini hingga data yang menunjukkan ketahanan jangka panjang. Distribusi Weibull biasanya digunakan untuk menyelesaikan masalah yang berhubungan dengan lamanya waktu suatu objek bisa bertahan dimulai dari waktu berfungsi sampai tidak dapat digunakan lagi (Ota, 2016).

Kerusakan suatu mesin dapat diperkirakan melalui getaran *bearing*, dengan begitu perlu dilakukan *machine health prognostic* untuk mengamati kondisi mesin di masa yang akan datang. Salah satu analisis yang digunakan untuk prognostik pada kesehatan mesin adalah *Support Vector Machine* (SVM). *Support Vector Machine* (SVM) adalah teknik machine learning untuk klasifikasi. Dalam kasus regresi, terdapat pengembangan dari metode *Support Vector Machine* yaitu *Support Vector Regression* (SVR). Model *Cox Proportional Hazard* digunakan untuk menentukan kegagalan dan menggunakan RMS sebagai fitur untuk menilai kondisi *bearing*. Kondisi kesehatan *bearing* akan ditunjukkan oleh nilai degradasi kegagalan yang berkisar antara 0 dan 1, di mana nilai 1 menunjukkan kondisi sehat dan nilai 0 menunjukkan kondisi kegagalan akhir (Caesarendra et al., 2010).

Berdasarkan hal tersebut, penelitian ini dilakukan untuk mengkaji ulang mengenai akurasi *Support Vector Regression* dengan model *Cox Proportional Hazard* untuk analisis *survival* pada data *bearing*. Data yang akan digunakan dalam penelitian ini adalah data vibrasi *bearing* yang diperoleh dari *Prognostics and Health Management*, FEMTO-ST Institute.

B. Metode

Distribusi Weibull

Distribusi Weibull banyak digunakan pada persoalan keandalan (*reliability*) dan analisis perawatan (*maintainability*) terutama dalam masalah pengujian kehidupan seperti waktu untuk kegagalan atau panjang umur suatu produk. Distribusi ini berguna sebagai model kegagalan dalam menganalisis keandalan berbagai jenis sistem. Fungsi densitas peluang dan fungsi *hazard* menurut Ebeling (1997), adalah:

$$f(t) = \frac{b}{\theta} \left(\frac{t}{\theta}\right)^{b-1} \exp\left(-\left(\frac{t}{\theta}\right)^b\right) \quad \dots (1)$$

$$h(t) = \frac{b}{\theta} \left(\frac{t}{\theta}\right)^{b-1} \quad \dots (2)$$

Distribusi ini memiliki 2 parameter yaitu parameter skala (θ) dan parameter bentuk (b), untuk menaksir parameter pada distribusi Weibull digunakan metode kuadrat terkecil (*Least Square Method*). Metode ini merupakan salah satu metode yang sering digunakan untuk menaksir parameter model regresi, pada distribusi Weibull perlu dilakukan transformasi ke model regresi (Razali et al., 2009).

$$\hat{b}_0 = \underline{y} - \hat{b}_1 \underline{x} \quad \dots (3)$$

$$\hat{b}_1 = \frac{\sum_{i=1}^n (x_i - \underline{x})(y_i - \underline{y})}{\sum_{i=1}^n (x_i - \underline{x})^2} \quad \dots (4)$$

Maka persamaannya adalah:

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_i$$

Dengan,

$$x_i = \ln t_i \quad ; y_i = \ln \ln \left[\frac{1}{1-F(t_i)} \right] \quad ; F(t_i) = \frac{i-0.3}{n+0.4}$$

$$b_I = b \quad ; b_0 = -b \ln \theta$$

Dengan begitu, persamaannya dapat ditulis kembali menjadi:

$$y_i = bx - b \ln \theta ; b_I = b \quad ; \theta = \exp \left(-\frac{b_0}{b} \right)$$

Cox Proportional Hazard

Cox Proportional Hazard adalah model regresi yang digunakan untuk melihat faktor-faktor yang menyebabkan suatu peristiwa terjadi, model ini menunjukkan bahaya (risiko) yang dimiliki suatu objek dengan karakteristik tertentu yang disebut kovariat (Cox dan Oakes, 1984). Salah satu tujuan dari model *Cox Proportional Hazard* adalah untuk memodelkan hubungan antara waktu bertahan hidup dan variabel-variabel yang diduga dapat mempengaruhi waktu bertahan hidup juga digunakan sebagai bagian dari pendekatan untuk memprediksi degradasi *bearing*, Secara umum model Cox dapat ditulis sebagai berikut (Kleinbaum & Klein, 2012):

$$h(t, x) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \quad \dots(5)$$

Di mana,

$h_0(t)$: fungsi baseline hazard

$\beta_1, \beta_2, \dots, \beta_p$: parameter regresi

$x_1, x_2, \dots, x_p$: nilai dari kovariat

Parameter regresi pada persamaan *Cox Proportional Hazard* dapat diestimasi menggunakan metode *partial likelihood*. dapat ditulis sebagai berikut (Nikulin & Wu, 2016):

$$L(\beta) = \prod_{t=l}^n \frac{\exp(\beta' x_t)}{\sum_{l \in R(t_l)} \exp(\beta' x_l)} \quad \dots(6)$$

fungsi *ln-likelihood* dari persamaan (6) adalah:

$$\begin{aligned} \ln L(\beta) &= \sum_{t=l}^n (\beta' x_t) - \sum_{t=l}^n \ln [\sum_{l \in R(t_l)} \exp(\beta' x_l)] \\ &= \sum_{t=l}^n \sum_{i=l}^p (\beta_i x_{it}) - \sum_{t=l}^n \ln \left[\sum_{l \in R(t_l)} \exp \sum_{i=l}^p (\beta_i x_{il}) \right] \\ &= \sum_{t=l}^n \left(\sum_{i=l}^p (\beta_i x_{it}) - \ln \left[\sum_{l \in R(t_l)} \exp \sum_{i=l}^p (\beta_i x_{il}) \right] \right) \quad \dots(7) \end{aligned}$$

Digunakan iterasi numerik metode Newton-Raphson karena turunan pertama implisit menurut Collet (2003) dalam (Khotimah, 2018). Jika $g(\beta)$ merupakan vektor berukuran $p \times 1$ yang merupakan turunan pertama dari fungsi $\ln L(\beta)$ terhadap parameter β . $H(\beta)$ merupakan matriks hessian berukuran $p \times p$ yang terdapat turunan kedua dari fungsi *ln-likelihood*, dengan begitu estimasi parameter pada iterasi ke $(l+1)$ adalah $\beta^{(l+1)} = \beta^l - H^{-1}(\beta^l)g(\beta^l)$. Iterasi akan berhenti ketika $\|\beta^{(l+1)} - \beta^l\| \leq \varepsilon$, di mana ε merupakan suatu bilangan yang sangat kecil.

Support Vector Regression

Support Vector Machine (SVM) merupakan metode *machine learning* yang diperkenalkan oleh Vapnik pada tahun 1972, sangat populer dan banyak digunakan baik untuk klasifikasi maupun regresi. Pada penelitian ini, Support Vector Machine (SVM) dideskripsikan sebagai model regresi nonlinier yang digunakan dalam bidang diagnosis dan prognosis kerusakan mesin yang dapat disebut juga sebagai *Support Vector Regression* (SVR). Dalam konteks penelitian ini, SVR digunakan untuk

mengidentifikasi pola atau tren dalam data kesehatan *bearing*, yang kemudian dapat digunakan untuk memprediksi kapan kegagalan *bearing* kemungkinan akan terjadi. Menurut Cortez dan Vapnik dalam (Hedianti, 2019), SVR memetakan vektor *input* ke dalam ruang fitur *high - dimensional*. Rumus *Support Vector Regression* dapat ditulis seperti berikut:

$$f(x) = w^T \phi(x) + b \quad \dots(8)$$

Dengan,

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

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

b : bias (*threshold*)

Fungsi Kernel

Fungsi kernel digunakan untuk menyelesaikan masalah non linier dalam SVR, dan fungsi ini akan dimasukkan ke dalam algoritma SVR. Pada proses *input*, data x yang berada di *input space* akan dipetakan ke *feature space* dengan dimensi yang lebih tinggi melalui *map* $\phi(\phi: x \rightarrow \phi(x))$. Dalam menemukan titik-titik *support vector* bergantung pada *dot product* dari data yang sudah ditransformasikan pada ruang baru yang berdimensi lebih tinggi, yaitu $\phi(x_i) \cdot \phi(x)$. Secara implisit fungsi kernel $K = (x_i, x)$ didefinisikan sebagai transformasi ϕ , hal ini disebut sebagai Kernel Trick yang dirumuskan sebagai berikut (Awad & Khanna, 2015):

$$K(x_i, x) = \phi(x_i) \cdot \phi(x) \quad \dots(9)$$

Fungsi kernel yang digunakan dalam penelitian ini adalah fungsi basis radial, juga dikenal sebagai kernel RBF (*Radial Basis Function*) atau kernel Gaussian. Fungsi ini didefinisikan sebagai berikut:

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right), \sigma > 0 \quad \dots(10)$$

Untuk mengevaluasi kinerja model prediksi terhadap data aktual menggunakan data training, menggunakan Root Mean Square Error (RMSE) dan koefisien determinasi (R^2). Nilai RMSE yang rendah menunjukkan bahwa model memiliki akurasi yang lebih baik, karena kesalahan prediksi lebih rendah. Koefisien determinasi digunakan untuk mengukur atau melihat kualitas dari suatu model, juga dapat menunjukkan tingkat akurasi dan ketika nilai koefisien determinasi mendekati 1 maka model semakin baik.

$$RMSE = \left(\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2\right)^{\frac{1}{2}} \quad \dots(11)$$

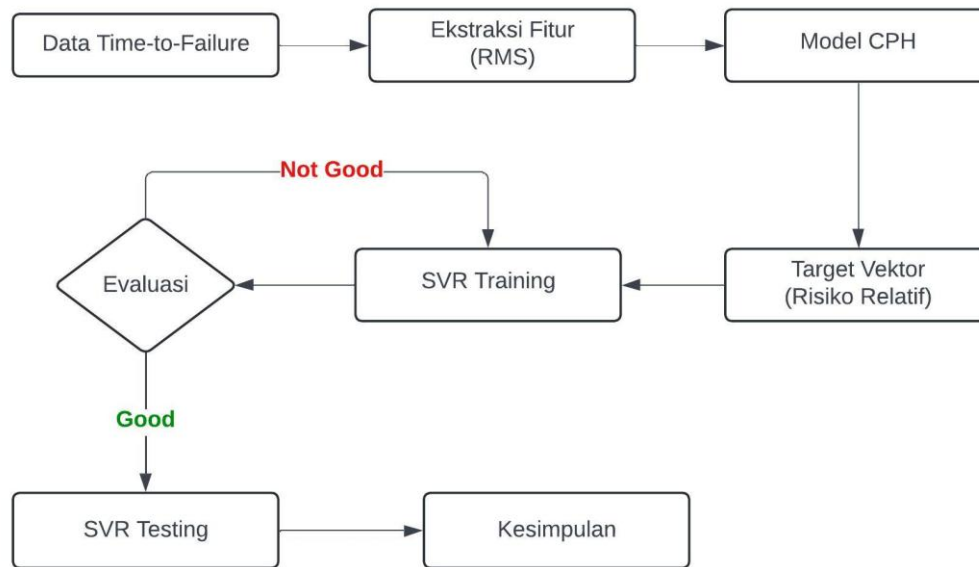
$$R^2 = \left(\frac{cov(y_t, \hat{y}_t)}{S_{y_t} S_{\hat{y}_t}}\right)^2 \quad \dots(12)$$

Di mana,

$$cov(y_t, \hat{y}_t) = \frac{1}{N} \sum_{t=1}^N (y_t - \bar{y}_t) (\hat{y}_t - \bar{\hat{y}}_t)$$

sedangkan S_{y_t} dan $S_{\hat{y}_t}$ merupakan standar deviasi dari nilai aktual dan nilai yang divalidasi.

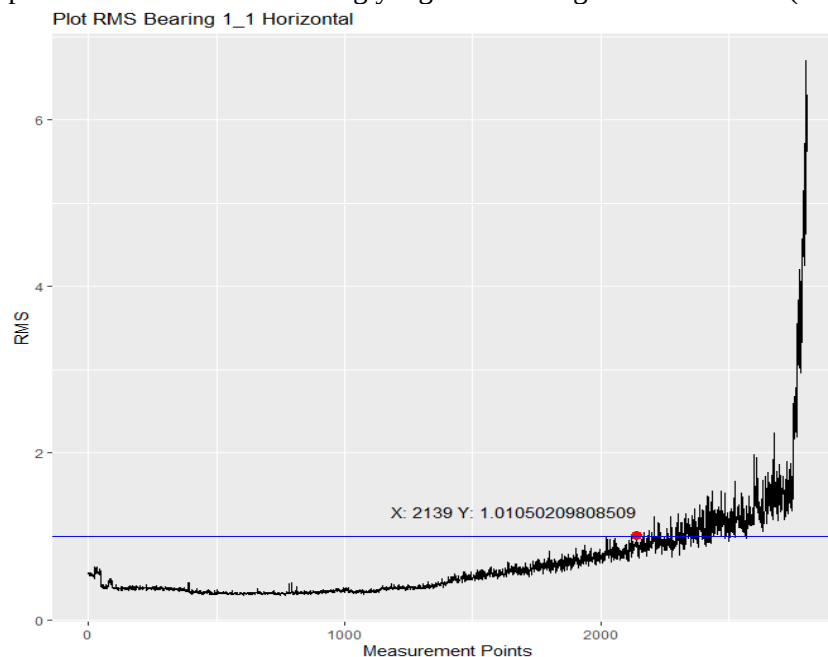
Data yang digunakan diunduh dari *Institute of Electrical and Electronic Engineers* melalui penyimpanan data oleh Patrick Nectoux, Rafael Gouriveau, Kamal Medjaher, Kamran Javed, Ahmed Mosallam, Emmanuel Ramasso, dan Nouredine Zerhouni dari *Prognostics and Health Management, FEMTO-ST Institute*, terdapat 3 data set dengan total 24 *bearing*. Adapun tahapan penelitian yang akan dilakukan adalah sebagai berikut:



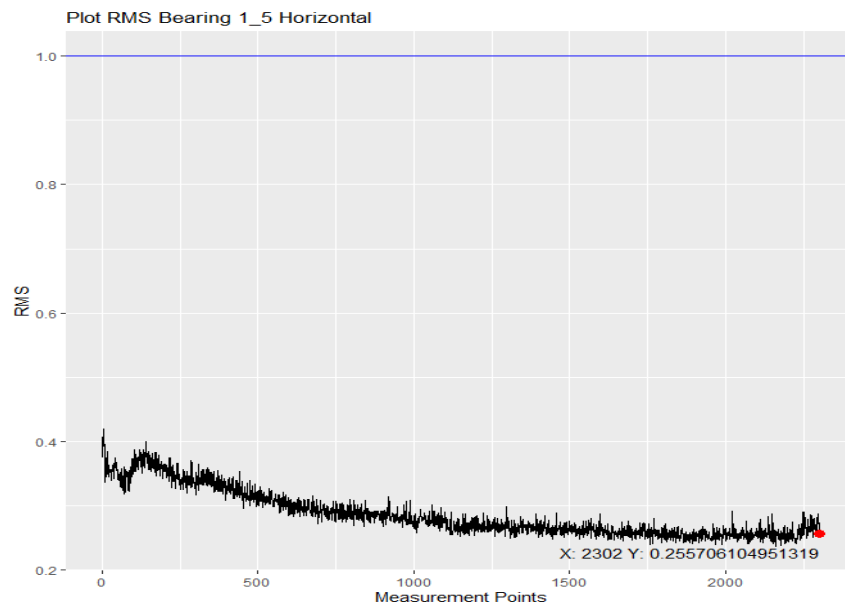
Gambar 1. Tahapan Penelitian

C. Hasil Penelitian dan Pembahasan

Dalam penelitian ini, digunakan 24 data getaran *bearing* yang diambil dari FEMTO-ST *Institute*. Untuk memudahkan dalam mendeteksi kerusakan pada *bearing* dengan menganalisis pola getaran, data tersebut akan diubah menjadi RMS. Untuk menentukan *time-to-failure* pada data *bearing*, dapat menggunakan *threshold* dilihat dari plot RMS terhadap nilai *measurement point*, hasilnya terdapat 17 *bearing* rusak dan 7 *bearing* belum mengalami kerusakan. Sebagai contoh, akan ditunjukkan plot RMS tersebut untuk *bearing* yang mengalami kerusakan (tidak tersensor) pada Gambar 1. dan pada Gambar 2. untuk *bearing* yang belum mengalami kerusakan (tersensor).



Gambar 1. Plot RMS Tidak Tersensor



Gambar 2. Plot RMS Tersensor

Fungsi *baseline hazard* pada persamaan (5) diasumsikan berdistribusi Weibull, sehingga 17 *bearing* yang mengalami kerusakan atau dapat dikatakan sebagai data *time-to-failure* akan dicari nilai parameter bentuk dan skalanya menggunakan metode kuadrat terkecil. Adapun perhitungan untuk menaksir parameter tersebut sebagai berikut:

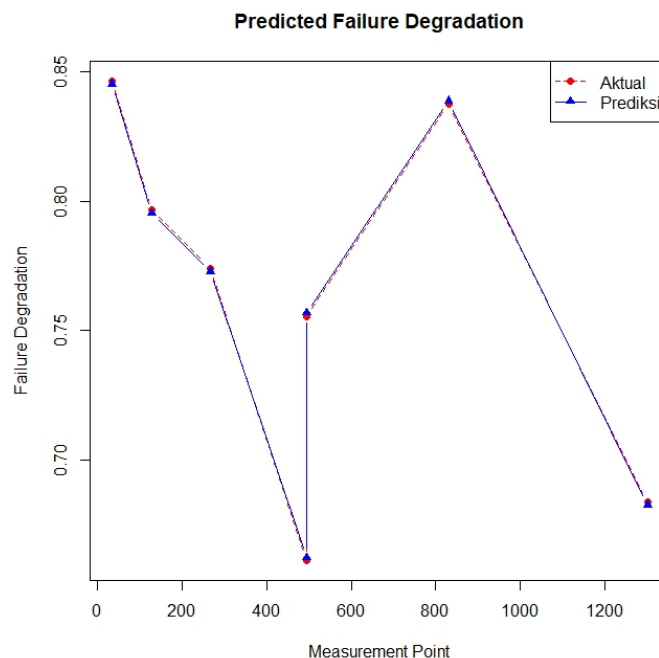
Tabel 1. Perhitungan Parameter pada Distribusi Weibull

i	x_i	$\hat{F}(t_i)$	y_i
1	-1,47331	0,04023	-3,19268
2	-0,12563	0,09770	-2,27488
3	0,61368	0,15517	-1,78009
4	0,66140	0,21264	-1,43098
5	0,67916	0,27011	-1,15560
6	1,23070	0,32759	-0,92412
7	1,23070	0,38506	-0,72108
8	1,69460	0,44253	-0,53726
9	1,75041	0,5	-0,36651
10	2,02228	0,55747	-0,20426
11	2,02320	0,61494	-0,04671
12	2,19954	0,67241	0,10975
13	2,40607	0,72989	0,26919
14	2,42037	0,78736	0,43705
15	2,42591	0,84483	0,62231

i	x_i	$\hat{F}(t_i)$	y_i
16	2,50609	0,90230	0,84408
17	2,69781	0,95977	1,16725

Berdasarkan tabel di atas, diperoleh nilai parameter bentuk (b) dan parameter skala (θ) pada distribusi Weibull adalah 1,02045 dan 7,3732. Nilai parameter regresi pada persamaan (5) diperoleh menggunakan metode *partial likelihood* dengan bantuan software R Studio, sehingga nilai untuk β pada persamaan (5) adalah 1,2645.

Data risiko yang akan digunakan akan dibagi ke dalam data *training* dan data *testing*. Data *training* akan digunakan untuk menentukan parameter C dan σ dalam pembentukan model SVR, dengan 12 data pertama dijadikan data *training* dan 5 data lainnya dijadikan data *testing*. Pada pembentukan model SVR, Fungsi kernel yang digunakan dalam penelitian ini adalah Kernel RBF, dengan parameter C yang digunakan adalah $C = [0.1, 1, 10, 100, 1000, 10000]$, $\sigma = [0.5, 1, 1.5, 2, 2.5]$ dan $\varepsilon = [0.0001, 0.001, 0.02]$. Menggunakan *software R 4.4.1*, hasil parameter yang terbaik untuk nilai C adalah 10, nilai σ adalah 0.5, dan nilai ε adalah 0,02. Pada pembentukan SVR didapatkan plot seperti gambar di bawah:



Gambar 3. Plot Prediksi Data *Training*

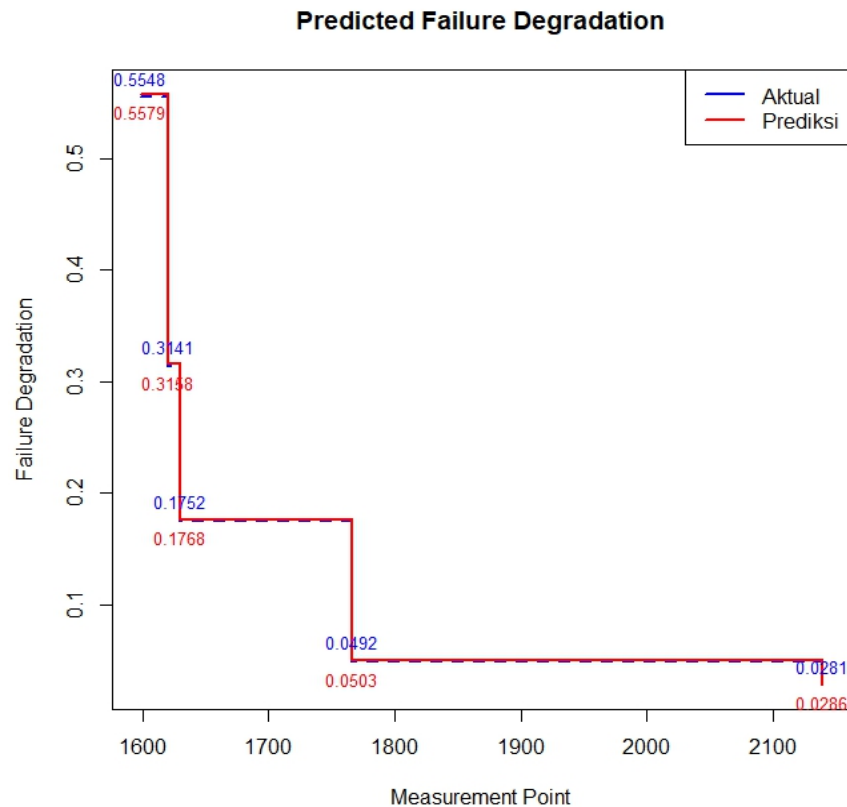
Dari **Gambar 3**, terlihat bahwa hasil prediksi dapat menjelaskan secara baik nilai aktual, dan diperoleh nilai RMSE dan R^2 adalah 0,01248 dan 0,99965. Selanjutnya akan dihitung rerata akurasi pada model SVR dengan data *training*.

Tabel 2. Nilai Akurasi SVR Data *Training*

No	Aktual	Prediksi	Akurasi
1	0,84665	0,84525	99,834
2	0,79670	0,79538	99,834
3	0,77419	0,77289	99,832

No	Aktual	Prediksi	Akurasi
4	0,66104	0,66216	99,831
5	0,75555	0,75683	99,830
6	0,83743	0,83883	99,834
7	0,68383	0,68266	99,829

Dari hasil perhitungan tersebut, didapatkan rata-rata nilai akurasi adalah 99,832%. Kemudian model SVR yang sudah dievaluasi dapat digunakan untuk melakukan prediksi menggunakan data *testing*, dengan menggunakan *software R Studio 4.4.1* diperoleh plot seperti di bawah:



Gambar 4. Plot Prediksi Data *Testing*

Dari **Gambar 4.** terlihat bahwa hasil prediksi dapat menjelaskan secara baik nilai aktual, langkah selanjutnya dapat dihitung rerata akurasi pada model SVR dengan data *testing*.

Tabel 3. Nilai Akurasi SVR Data *Testing*

No	Aktual	Prediksi	Akurasi
1	0,55485	0,55789	99,452
2	0,31414	0,31579	99,475
3	0,17515	0,17682	99,047
4	0,04916	0,05025	97,776
5	0,02814	0,02861	98,350

Dari hasil perhitungan tersebut, didapatkan rerata nilai akurasi adalah sebesar 98,82%. Hasil rerata akurasi data *training* dan data *testing* akan ditunjukkan pada tabel di bawah sebagai berikut:

Tabel 4. Nilai Akurasi SVR-CPH

		Akurasi
SVR-CPH	Data Training	99,832%
	Data Testing	98,82%

Seperti yang terlihat dalam tabel, tingkat akurasi pada model SVR dengan menjadikan nilai risiko dari model Cox sebagai target vektor menggunakan data *training* dan juga data *testing* menunjukkan hasil yang sama-sama tinggi.

D. Kesimpulan

Berdasarkan hasil penelitian ini, model yang digunakan untuk memprediksi *failure degradation* pada *bearing* dengan SVR dan CPH, memberikan hasil prediksi yang cukup akurat, sebagaimana ditunjukkan dalam perbandingan nilai aktual dan prediksi pada data *training* maupun data *testing*, yang memiliki nilai *error* rendah. Saat validasi metode SVR menggunakan data *training*, diperoleh nilai *error* yang rendah yaitu 0,0124 dan nilai koefisien determinasi yang tinggi sebesar 0,9996, parameter dalam model tersebut dapat digunakan untuk SVR data *testing*. Model yang diterapkan menunjukkan kemampuan yang baik dalam memprediksi hasil dengan tingkat akurasi yang tinggi, dengan rerata nilai akurasinya sebesar 99,83% pada data *training* dan 98,82% pada data *testing*. Ketika menggunakan model *Cox Proportional Hazard*, penting untuk memastikan bahwa asumsi *proportional hazard* terpenuhi. Jika asumsi tidak terpenuhi, disarankan untuk mempertimbangkan metode alternatif.

Daftar Pustaka

- Awad, M., & Khanna, R. (2015). *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Apress.
- Caesarendra, W., Widodo, A., & Yang, B. S. (2011). Combination of probability approach and support vector machine towards machine health prognostics. *Probabilistic Engineering Mechanics*, 26(2), 165-173. <https://doi.org/10.1016/j.probengmech.2010.09.008>.
- Cox, D.R., Oakes, D., & Oakes, D. (1984). *Analysis of survival data*. Taylor & Francis.
- Ebeling, C. E. (1997). *An Introduction to Reliability and Maintainability Engineering*. McGraw Hill.
- Hedianti, E. S. (2019). Peramalan Harga Saham Dengan Menggunakan Metode Support Vector Regression (SVR).
- Khotimah, C. (2018). Additive Survival Least Square SVM Untuk Analisis Data Survival (Studi Simulasi dan Studi Kasus pada Data Pasien Kanker Serviks di RSUD dr. Soetomo Surabaya).
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis: A Self-Learning Text, Third Edition*. Springer New York.
- Nikulin, M., & Wu, H.-D. I. (2016). *The Cox Model and Its Applications*. Springer Berlin Heidelberg. <https://link.springer.com/book/10.1007/978-3-662-49332-8>.

- Otaya, L. G. (2016). Distribusi Probabilitas Weibull Dan Aplikasinya (Pada Persoalan Keandalan (Reliability) dan Analisis Rawatan (Maintainability). *Tadbir: Jurnal Manajemen Pendidikan Islam*, 4(2), 44-66.
<https://journal.iaingorontalo.ac.id/index.php/tjmpi/article/view/438>.
- Razali, A. M., Salih, A. A., & Mahdi, A. A. (2009). Estimation Accuracy of Weibull Distribution Parameters. *The Journal of Applied Sciences Research*, 5, 790-795.