

# Application of Gaussian Mixture and Hidden Markov Model (GM-HMM) for Prediction of Machinery Failure

Achmad Widodo<sup>1\*</sup>, Toni Prahasto<sup>1</sup> dan Yasir Abdur Rohman<sup>2\*</sup>

<sup>1</sup>Departemen Teknik Mesin, Fakultas Teknik, Universitas Diponegoro

<sup>2</sup>Prodi Sarjana Departemen Informatika, Fakultas Sains dan Matematika, Universitas Diponegoro

\*Corresponding author: achmadwidodo@lecturer.undip.ac.id

**Abstract.** Machinery failure is one of problems which must be avoided in industry because of its impact in production capacity and loss of profit. Sudden failure and catastrophic will give impacts more severe than ordinary failure due to those failures are unpredictable and will give a longer downtime. This research deals with implementation of machine learning (ML) for prediction of machinery failure. The information of failure in rotating machinery may be transferred through bearings. So, it is relevant to research bearings failure prediction and is regarded as failure prediction in machinery. Symptoms of failures in bearing can be noise and/or vibration in high level. If bearing is experienced with high level vibration in long time duration so it leads machinery failure. The proposed method is composed of two steps: first is a feature extraction from original vibration data that was acquired by accelerometer, and second is a building model of ML thru training of GM-HMM with extracted feature data. A trained GM-HMM model is then used to predict time to failure or remaining useful life (RUL) of bearing as well as assessment of bearing condition. The proposed method is validated using bearing run-to-failure vibration data which is an open source data hosted in NASA prognostic data repository. Results show that the prediction of failure time is good based on confidence level and error prediction.

**Abstrak.** Kerusakan mesin adalah salah satu masalah di industri yang harus dihindari karena akan berdampak adanya kerugian akibat berkurangnya produksi. Kerusakan mesin yang bersifat mendadak, katastropik dan tanpa bisa diprediksi sebelumnya menyebabkan waktu menganggur (*downtime*) yang lebih lama, hal ini karena perawatan yang dilakukan akan bersifat seketika dan tanpa perencanaan yang baik. Pada artikel ini akan dibahas sebuah metode *machine learning* (ML) untuk memprediksi kerusakan yang terjadi pada salah satu elemen mesin rotasi yaitu bantalan (*bearing*). Salah satu indikator kerusakan pada bantalan adalah kondisi tingkat getaran dan derau yang tinggi, apabila hal ini dibiarkan akan menyebabkan kerusakan yang lebih parah pada mesin. Metode yang diusulkan terdiri dari dua tahap, pertama ekstraksi fitur data yang diperoleh dari sensor getaran, dan tahap kedua adalah proses pembuatan model dengan cara melatih algoritma ML yaitu GM-HMM menggunakan fitur data yang sudah diekstrak pada tahap pertama. Selanjutnya, model yang sudah dihasilkan digunakan untuk memprediksi waktu kerusakan atau sisa umur (*remaining useful life – RUL*) sekaligus melakukan asesmen terhadap kondisi kesehatan mesin melalui data sensor getaran yang dipasang pada rumah bantalan mesin tersebut. Metode yang diusulkan diuji dengan data getaran bantalan yang ada pada laman “*NASA prognostic data repository*”, di mana data tersebut adalah data pengujian getaran *run-to-failure* bantalan dengan berbagai kondisi operasi. Hasil menunjukkan prediksi waktu kerusakan/RUL yang cukup baik berdasarkan tingkat kepercayaan dan kesalahan (*error*) yang bisa diterima.

**Kata kunci:** prognosis, getaran, *machine learning*, *hidden markov model*, *gaussian mixture*.

---

© 2019. BKSTM-Indonesia. All rights reserved

## Pendahuluan

Perawatan/pemeliharaan (*maintenance*) mesin dalam kondisi beroperasi adalah hal vital bagi industri yang menuntut pemenuhan kriteria dan pemilihan metode yang tepat agar tujuan perawatan tersebut dapat tercapai dengan optimal. Perawatan mesin dapat dilakukan secara kuratif,

preventif ataupun prediktif tergantung dari bermacam-macam kriteria pemilihan metode perawatan mesin. Perawatan kuratif dilakukan manakala mesin sudah mengalami kegagalan atau disebut juga dengan *breakdown maintenance*. Metode perawatan preventif dilakukan untuk mencegah kejadian gagal pada sistem mesin.

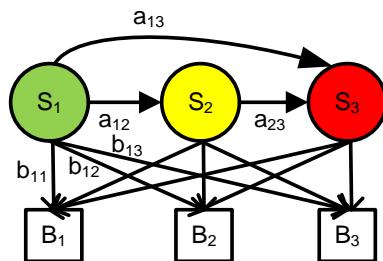
Metode perawatan ini biasanya berbasis waktu yang tetap termasuk di dalamnya adalah *overhaule*. Metode perawatan prediktif sering disebut juga dengan metode berbasis kondisi atau *condition based maintenance* – CBM. Metode perawatan ini diklaim dapat menghasilkan keuntungan karena dapat melakukan prediksi kondisi mesin berbasis pengukuran parameter mesin melalui sensor, secara tepat dan *realtime* sehingga kerusakan yang bersifat katastropis dapat dihindari, hal ini tentunya dapat juga menghindari *downtime* yang lama [1].

Kerangka kerja metode perawatan berbasis kondisi (prediktif) memasukkan cara diagnosis dan prognosis untuk mendiagnosa jenis kerusakan dan memprediksi sisa umur (RUL) dari sebuah mesin ataupun sistem [2]. Beberapa metode pendekatan yang digunakan adalah: 1) berbasis model degradasi (baik fisik maupun matematik); 2) berbasis data-driven, yaitu data pengukuran oleh sensor digunakan untuk memodelkan perilaku degradasi mesin atau elemen mesin yang terkait; 3) berbasis experience, data dari sensor dikumpulkan dalam periode waktu yang lama kemudian data ini digunakan untuk memperkirakan parameter yang terkait dengan kehandala. Selanjutnya ekstrapolasi maupun proyeksi dilakukan untuk menentukan RUL [3].

Pada penelitian ini, perkiraan waktu kegagalan setelah terdeteksi adanya abnormalitas sistem atau RUL dilakukan dengan memakai data pengukuran sensor getaran (akselerometer). Ekstraksi fitur dilakukan pada data getaran asli (*raw data*) untuk mendapatkan fitur yang sensitif terhadap fenomena degradasi mesin. Fitur ini selanjutnya dipakai untuk mengajari/melatih algoritma ML yang dipakai. Selanjutnya, model MG-HMM yang telah dilatih digunakan untuk mengasess kondisi terkini mesin dan selanjutnya memprediksi kondisi mesin pada jangka waktu mendatang. Prediksi ini juga meliputi estimasi RUL dengan tingkat kepercayaan tertentu yang sudah diset sebelumnya. Informasi tentang RUL cukup penting dalam rangka membuat keputusan perencanaan kegiatan perawatan mesin, misalnya untuk menunda (*delay*) program perawatan ataupun menghentikan operasi mesin karena ditemui *trend* kegagalan lebih dini.

## Metode Penelitian

**Hidden Markov Model (HMM).** HMM didefinisikan sebagai sebuah model statistik yang digunakan untuk menyajikan proses stokastik di mana status dari sebuah sistem yang ditinjau tidak secara langsung diobservasi [4]. Gambar 1 menunjukkan arsitektur HMM dengan tiga buah status dengan konstruksi alur perubahan status dari kiri-ke-kanan.



Gambar 1. Arsitektur HMM untuk alur perubahan tiga status

Arsitektur HMM secara lengkap didefinisikan dengan beberapa parameter di bawah ini:

1. N : jumlah status pada model HMM;
2. M : jumlah observasi berbeda pada masing-masing status;
3. A : probabilitas distribusi observasi masing-masing status;
4.  $\pi$  : distribusi status awal

Secara sederhana, HMM dapat dinyatakan dengan notasi sebagai berikut ( $\lambda = \pi, A, B$ ) yang dipakai pada setiap model HMM. Pada implementasinya, HMM dapat digunakan untuk memecahkan tiga jenis permasalahan sebagai berikut:

1. Permasalahan deteksi, diberikan model HMM  $\lambda$  dan matrik urutan observasi  $O$  untuk menghitung probabilitas bersyarat  $P(O|\lambda)$  dari urutan data observasi yang diberikan;
2. Permasalahan *decoding*, diberikan matrik data observasi  $O$  untuk menentukan urutan status tersembunyi (*hidden state*) yang berisi probabilitas urutan data observasi. Permasalahan ini diselesaikan dengan algoritma Viterbi [3];
3. Permasalahan pelatihan (*training*), menentukan parameter model yang sesuai dengan urutan matrik observasi  $O$  dengan memaksimalkan probabilitas bersyarat  $P(O|\lambda)$ . Permasalahan ini diselesaikan dengan algoritma Baum-Welch [4].

HMM diskrit dapat dipakai pada observasi atribut diskrit dan menggunakan probabilitas diskrit untuk memodelkan transisi probabilitas observasi. Pada pendekatan ini, pemantauan kondisi mesin dapat dianggap sebagai obervasi melalui sinyal yang bersifat kontinyu. Agar dapat menggunakan observasi kontinyu ini beberapa batasan digunakan untuk menjamin bahwa parameter fungsi probabilitas densitas dapat diperkirakan [5].

$$b_i(O) = \sum_{m=1}^M C_{jm} \xi(O, \mu_{jm}, U_{jm}), \quad (1)$$

$$1 \leq j \leq N$$

Dimana  $O$  adalah vector observasi,  $C_{jm}$  koefisien *mixture* untuk *mixture* ke- $m$  pada status  $i$  dan  $\xi$  adalah logaritmik konkaf atau densitas eliptik simetris dengan mean vektor  $\mu_{jm}$  dan matrik kovarian  $U_{jm}$  untuk mixture komponen ke  $m$  pada status  $j$ . Densitas Gaussian biasanya digunakan untuk  $\xi$  dan model yang berkaitan disebut model GM-HMM yang secara lengkap didefinisikan dengan matrik A, matrik B dan probabilitas awal  $\pi$ . Pada model GM-HMM matrik observasi B dimodelkan dengan densitas Gaussian dengan mean  $\mu$ , standar deviasi  $\sigma$  dan matrik *mixture* M.

**Metode Diagnosis dan Prognosis.** Beberapa hasil riset yang terkait dengan penggunaan HMM dalam bidang dianalisa kerusakan mesin dilaporkan pada artikel [6,7]. Pada penelitian ini, metode prediksi RUL dilakukan dengan dua tahap yaitu pertama tahap pelatihan (training) dan kedua tahap prediksi kondisi mesin dan estimasi RUL.

Pada tahap pelatihan, fitur yang diekstrak dari data pemantauan kondisi mesin melalui sensor getaran ditransformasi dan dilatihkan pada HMM dengan algoritma Baum-Welch untuk menghasilkan model prediksi [4]. Model HMM yang sudah dibuat dengan beberapa jenis data (termasuk data kegagalan) kemudian disimpan dalam database yang berisi seluruh model HMM dengan label/atribut serta riwayat pelatihannya.

Tahap prediksi dapat dilakukan secara online dengan cara memberikan input data fitur pada model yang sudah dilatih dengan tujuan melakukan prediksi status (kesehatan) mesin dan estimasi RUL. Prediksi dilakukan dengan cara mencari model yang paling baik kesesuaiannya dengan urutan (sekuen) data input (aktual) dengan menghitung probabilitas bersyarat  $P(O|\lambda)$ . Selanjutnya, pada tahap kedua ini fitur yang terpilih digunakan untuk menggambarkan kondisi status mesin terkini dan memperkirakan RUL. Algoritma Viterbi [3] digunakan untuk menemukan urutan status yang paling gigih muncul (persisten) pada observasi terakhir dan urutan tersebut dipertahankan sebagai kondisi status terkini mesin, sebagaimana ditunjukkan pada Pers. 2.

$$\begin{aligned} \text{urutan status} &= (s_1, s_2, \dots, s_t), \\ \text{status akhir} &= (s_{t-l}, \dots, s_{t-2}, s_{t-1}, s_t) \end{aligned} \quad (2)$$

dimana  $l$  = observasi lampau dan  $t$  = waktu kini.

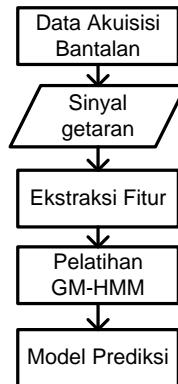
Selanjutnya, dengan mengetahui kondisi status aktual kesehatan mesin persamaan Chapman-Kolmogorov [3] digunakan untuk memperkirakan kembali status kesehatan mesin setelah  $n$  kali iterasi. Pada saat probabilitas status kesehatan yang diprediksi mencapai batas nilai  $\varepsilon$  yang telah

ditetapkan sebelumnya, maka RUL dapat dihitung dengan rumus

$$\hat{P}(n) = \pi_t A^n \quad (3)$$

$$RUL = n \Leftrightarrow \hat{P}(s = s_N | n) = \varepsilon \quad (4)$$

Diagram alir penelitian ini ditunjukkan pada Gambar 2.



**Gambar 2.** Proses prediksi dengan GM-HMM

Terkait dengan implementasi algoritma Viterbi pada perkiraan perhitungan RUL, mean dan standar deviasi yang merupakan fungsi dari durasi visit dinyatakan dengan persamaan berikut:

$$\mu(D(s_i)) = \frac{1}{K} \sum_{\theta=1}^K D(s_{i\theta}) \quad (5)$$

$$\sigma(D(s_i)) = \sqrt{\frac{1}{K} \sum_{\theta=1}^K [D(s_{i\theta}) - \mu(D(s_i))]^2} \quad (6)$$

dimana  $D(\bullet)$  adalah durasi visit,  $i$  indeks status dan  $K$  adalah total visit. Selanjutnya, representasi GM-HMM secara lengkap dapat dinyatakan sebagai berikut

$$\lambda = (\pi, A, B, \mu(D(s_i)), \sigma(D(s_i)), s_{final}) \quad (7)$$

Selanjutnya estimasi RUL dilakukan dengan formula sebagai berikut

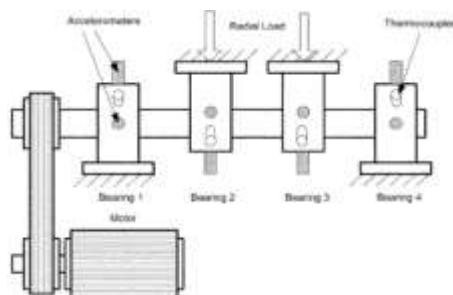
$$RUL_{upr} = \sum_{i=crstt}^N [\mu(D(s_i)) + n \sigma(D(s_i))] \quad (8)$$

$$RUL_{mean} = \sum_{i=crstt}^N \mu(D(s_i)) \quad (8)$$

$$RUL_{lwr} = \sum_{i=crstt}^N [\mu(D(s_i)) - n \sigma(D(s_i))] \quad (8)$$

**Data Eksperimen.** Data eksperimen yang dipakai untuk validasi metode yang diusulkan adalah data getaran bantalan dari sebuah test rig yang ditampilkan secara skematis pada Gambar 3. Test rig tersebut terdiri dari empat buah bantalan gelinding yang dipasang pada satu poros. Kecepatan putar diatur tetap sebesar 2000 RPM dan beban 6000 lb diberikan pada poros dan bantalan seperti ditunjukkan pada gambar skema. Pada masing-masing bantalan dipasang dua buah

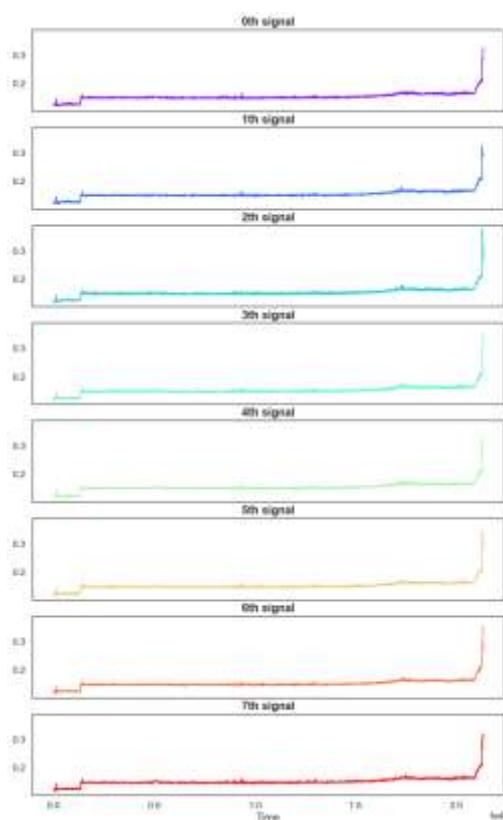
sensor akselerometer dengan arah pemasangan horizontal dan vertikal, jadi seluruhnya terdapat 8 sensor getaran. Pada alat data akuisisi diatur frekuensi sampling 20 kHz saat mengambil data getaran bantalan.



Gambar 3. Testrig bantalan gelinding

## Hasil dan Pembahasan

Data getaran bantalan yang digunakan pada penelitian ini ditampilkan pada Gambar 4 yang berupa fitur root-means-square (rms) sinyal getaran bantalan. Fitur tersebut digunakan untuk melatih (training) GM-HMM dalam rangka membentuk model prediksi.



Gambar 4. Fitur rms sinyal getaran dataset#1

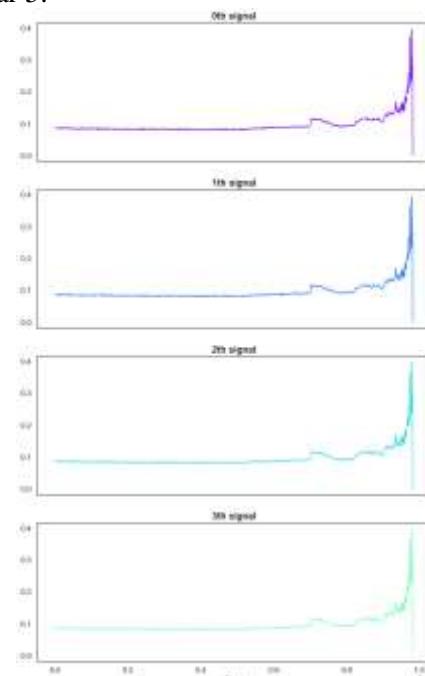
Sebanyak 8 buah fitur rms signal yang masing-masing diperoleh dari 8 buah sensor getaran digunakan untuk melatih GM-HMM. Selanjutnya, model yang dihasilkan digunakan untuk memprediksi kondisi status mesin dan RUL nya.

Pada saat tahap pelatihan, ada tiga status dibuat yaitu normal, degradasi dan gagal (failure).

Parameter GM-HMM pada tahap awal diinisiasi secara acak, kemudian fitur rms pada Gambar 4 diumpulkan kepada algoritma pelatihan untuk memperkirakan kembali parameter ( $\lambda = \pi, A, B$ ). Jumlah parameter *mixture* untuk masing-masing GM-HMM dipilih nilai dua, yang memungkinkan trade-off antara waktu komputasi dan kepresisan perhitungan. Ada 14 model GM-HMM yang dihasilkan dari proses pelatihan dengan algoritma Baum-Welch. Parameter yang diperkirakan tersebut adalah

$$\pi = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, A = \begin{pmatrix} 0,98 & 0,02 & 0 \\ 0,01 & 0,99 & 0 \\ 0,02 & 0,03 & 0,95 \end{pmatrix}, M = \begin{pmatrix} 0,68 & 0,32 \\ 0,28 & 0,72 \\ 0,99 & 0,01 \end{pmatrix} \quad (9)$$

Data online yang digunakan sebagai data uji prediksi diperoleh dari test rig yang sama namun dengan kondisi operasi yang berbeda baik putaran dan bebananya. Data tersebut ditampilkan pada Gambar 5.



Gambar 5. Data uji prediksi dan perkiraan RUL

Hasil prediksi RUL disajikan pada Tabel 1 di bawah ini

Tabel 1 Hasil prediksi RUL

Data	RUL_upper	RUL_mean	RUL_lower	Status
0	6424	6430	6436	S3
1	9045	9050	9055	S3

Dari hasil prediksi pada Tabel 1 didapatkan nilai yang cukup dekat untuk RUL upper, mean dan

lower. Dari kedua uji di atas memperlihatkan konsistensi prediksi RUL.

### Kesimpulan

Dari penelitian ini didapat hasil prediksi kondisi mesin berdasarkan data getaran yang diambil dari *testrig* menunjukkan hasil yang cukup baik yaitu status S3 (gagal). Sedangkan prediksi RUL menunjukkan nilai yang konsisten dari dua data uji yang diambil. Dapat disimpulkan bahwa metode GM-HMM dapat dipakai untuk prediksi status kondisi mesin dan prediksi RUL.

### Referensi

- [1] Lebold, M et al., 2001. Open standard for condition-based maintenance and prognostic system. Maintenance and Reliability Conference (MARCON).
- [2] Jardine, A.K. et al., 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical System and Signal Processing 20(7), 1483-1510.
- [3] Viterbi, A., 1967. Error bounds for convolutional codes and asymptotically optimal decoding algorithm. IEEE Transaction Information and Theory 13, 260-269.
- [4] Dempster, A. et al., 1977. Maximum likelihood from incomplete data via EM algorithm. Journal of Royal Statistical Society 39, 1-38.
- [5] Rabiner, L.R. et al., 1989. A tutorial on hidden markov models and selected applications in speech recognition. Proceeding of the IEEE 77, 257-286.
- [6] Dong, M. et al., 2007. A segmental semi-hidden markov model based diagnostics and prognostics framework and methodology. Mechanical System and Signal Processing 21, 2248-2266.
- [7] Ocak, H. et al., 2007. Online tracking of bearing wear using wavelet packet decomposition and probabilistics modeling: A method fro bearing prognostics. Journal of Sound and Vibration 302, 951-961.