



Diagnosa Kerusakan *Bearing* Menggunakan *Principal Component Analysis* (PCA) dan *Naïve Bayes Classifier*

Dwi Pudyastuti^{a,*}, Toni Prahasto^b, Achmad Widodo^b

^a Magister Sistem Informasi Universitas Diponegoro

^bJurusan Teknik Mesin, Fakultas Teknik, Universitas Diponegoro

Naskah Diterima : 13 Juni 2016; Diterima Publikasi : 12 Oktober 2016

DOI: 10.21456/vol6iss2pp114-123

Abstract

The research aims to build an information system which used to diagnose damage to the bearings based on data mining. Bearing was one of the essential parts in industry machinery. Bearing was used to reduce machines frictions or could be a moving component which oppressed each other. This fault diagnosis can avoid loss and damage of other machines components. This research was started with data preprocessing using wavelet discrete transformation, feature extraction, feature reduction using Principal Component Analysis, and classification process using Naïve Bayes classifier methods. Naïve Bayes Classifier is a classification method which based on probability and Bayesian theorem. The results of the systems shows that Naïve Bayes classification have a good performance which shown by a good accuracy in each data test.

Keywords: Data mining; principal component analysis; Naïve Bayes classification

Abstrak

Penelitian ini bertujuan untuk membangun sistem informasi yang digunakan untuk mendiagnosa kerusakan yang terjadi pada bearing berbasis *data mining*. *Bearing* merupakan salah satu komponen penting dalam mesin-mesin industri. *Bearing* berfungsi untuk mengurangi gesekan pada mesin atau komponen-komponen yang bergerak dan saling menekan antara satu dengan yang lainnya. Diagnosis kerusakan ini dapat menghindari terjadinya kerugian dan kerusakan komponen lain pada suatu mesin. Tahapan penelitian dimulai dengan prapemrosesan data menggunakan transformasi *wavelet* diskret, ekstraksi fitur, reduksi fitur menggunakan *Principal Component Analysis* dan proses klasifikasi menggunakan metode klasifikasi *Naïve Bayes*. *Naïve Bayes* adalah metode klasifikasi yang berdasarkan probabilitas dan Teorema Bayes. Hasil penggunaan sistem ini menunjukkan bahwa klasifikasi *Naïve Bayes* memiliki performa yang cukup bagus terlihat dari akurasi yang dihasilkan dari setiap data yang diuji.

Kata kunci: *Data mining*; ; principal component analysis; klasifikasi *Naïve Bayes*

1. Pendahuluan

Pemantauan kondisi mesin dan diagnosa kerusakan mesin sangat penting bagi industri untuk menjamin fungsionalitas performansinya. Keandalan yang tinggi diperlukan untuk mengurangi penurunan baik laba maupun produk yang disebabkan berhentinya produksi karena kerusakan mesin.

Keandalan sistem operasi dari mesin-mesin industri dan sistem produksi memiliki pengaruh besar terhadap keuntungan perusahaan dan kemampuan daya saing dalam dunia industri. Hal tersebut semakin menekankan pentingnya aktivitas pemantauan kondisi mesin secara *online*, diagnosis dan prognosis terhadap mesin-mesin, serta sistem dan proses produksi dalam industri (Helle, 2006).

Bearing merupakan salah satu komponen utama dalam mesin-mesin industri. Fungsi *bearing* selain untuk mengurangi gesekan, panas dan aus, juga berfungsi untuk menahan beban poros dan mesin, menjaga kekencangan dan mempermudah pergantian serta mengurangi biaya operasional.

Metode *Naïve Bayes* yang dipilih untuk melakukan klasifikasi merupakan bentuk sederhana dari sebuah jaringan *Bayes* yang merupakan salah satu metode *data mining*. Distribusi prior dalam metode klasifikasi *Naïve Bayes* adalah suatu distribusi yang menggambarkan pengetahuan awal tentang parameter sebelum pengamatan dilakukan (Bolstad, 2007).

Penelitian ini bertujuan untuk membangun sistem informasi yang digunakan untuk mendiagnosa kerusakan yang terjadi pada bearing berbasis *data mining* Manfaat

*Penulis korespondensi: dwi.pudyastuti.2015@gmail.com

yang diharapkan dari hasil penelitian ini adalah untuk mendapatkan data dan informasi tentang karakterisasi kerusakan berdasarkan hasil pengukuran getaran serta mempelajari kecocokan perilaku klasifikasi *Naïve Bayes* untuk data getaran *bearing*.

2. Kerangka Teori

Diagnosis kerusakan merupakan serangkaian proses yang meliputi proses deteksi dan proses mengidentifikasi kerusakan yang terjadi pada komponen. Proses diagnosis ini biasanya dilakukan pada komponen yang masih beroperasi seperti biasa namun sudah terlihat adanya penurunan kinerja sehingga tidak dapat berfungsi secara optimal. Diagnosis dapat digunakan sebagai umpan balik informasi untuk mendesain ulang sebuah sistem.

2.1 Pengecekan Severitas

Sebelum dilakukan klasifikasi, data yang akan digunakan untuk penelitian terlebih dahulu dilakukan pengecekan severitas dari data yang akan diolah. Pengecekan ini dilakukan dengan menghitung nilai *Vrms* sinyal, dan dibandingkan dengan ISO 10816. Standar ISO 10816 diperlihatkan seperti Gambar 1 berikut ini:

VIBRATION SEVERITY PER ISO 10816					
Machine		Class I small machines	Class II medium machines	Class III large rigid foundation	Class IV large soft foundation
in/s	mm/s				
Vibration Velocity Vrms	0.01	0.28			
	0.02	0.45			
	0.03	0.71			
	0.04	1.12			
	0.07	1.80			
	0.11	2.80		good	
	0.18	4.50			
	0.28	7.10		satisfactory	
	0.44	11.2			
	0.70	18.0		unsatisfactory	
	0.71	28.0			
	1.10	45.0		unacceptable	

Gambar 1. ISO 10816

2.2 Prapemrosesan Data

Setelah itu, data akan dilakukan prapemrosesan data tahap kedua, yaitu dengan menggunakan transformasi *wavelet* diskret. *Wavelet* mengklasifikasikan sinyal dalam versi penskalaan dan pergeseran (*scalling and shifting*) masing-masing dari sinyal sumber/*mother wavelet*. Proses *wavelet* diskret diawali dengan melewati sinyal pada seperangkat tapis frekuensi tinggi (*highpass filter*) dan tapis frekuensi rendah (*lowpass filter*). Kemudian dilanjutkan dengan operasi *sub-sampling* dengan mengambil masing-masing setengah dari keluaran filter.

Wavelet mengkonsentrasikan energinya dalam ruang dan waktu sehingga cocok untuk menganalisis sinyal yang sifatnya sementara. *Wavelet* populer pada aplikasi-aplikasi teknik, terutama pada analisa sinyal getaran/*vibration*, dan *wavelet* digunakan juga untuk

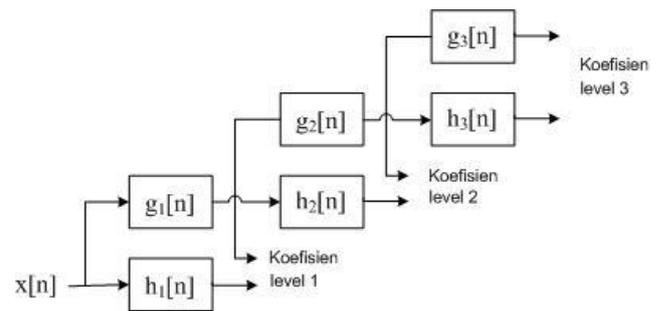
memonitor kondisi dan mendiagnosa kerusakan mesin. *Wavelet* diaplikasikan untuk menganalisa sinyal getaran pada roda gigi untuk mendeteksi kerusakan yang akan terjadi (Newland, 1994).

Wavelet sebagai *data cleansing* juga digunakan secara luas untuk pra-pemrosesan sinyal di bidang pemantauan dan diagnosa kerusakan (Menon et al., 2000). Transformasi *wavelet* mendekomposisi sebuah sinyal ke dalam kombinasi linier unit skala waktu. Sinyal asli dianalisa dan dibagi ke dalam beberapa komponen sinyal berdasarkan translasi atau pergeseran dari *mother wavelet* (atau fungsi dasar *wavelet*) sehingga terjadi perubahan skala dan diperlihatkan transisi dari setiap komponen frekuensi (Daubechies, 1992).

Metode transformasi *wavelet* dapat digunakan untuk menapis data atau meningkatkan mutu kualitas data dan untuk mendeteksi kejadian-kejadian tertentu serta dapat digunakan untuk pemampatan data (Foster et al., 1994).

Transformasi *wavelet* juga dapat digunakan untuk analisis sinyal-sinyal non-stasioner (yaitu sinyal yang kandungan frekuensinya bervariasi terhadap waktu), karena berkaitan dengan kemampuannya untuk memisah-misahkan berbagai macam karakteristik pada berbagai skala (Anant et al., 1997).

Pemisahan sinyal menjadi frekuensi tinggi dan frekuensi rendah melalui tapis frekuensi tinggi dan tapis frekuensi rendah disebut sebagai dekomposisi (Teriza, 2006). Proses dekomposisi diperlihatkan pada Gambar 2 berikut ini:



Gambar 2. Dekomposisi wavelet 3 tingkat

Proses dekomposisi dimulai dengan melewati sinyal asal melewati tapis frekuensi tinggi dan tapis frekuensi rendah. Setelah filterisasi, setengah dari sampel atau salah satu subsampling dapat dieliminasi berdasarkan aturan Nyquist (Teriza, 2006). Sehingga sinyal selalu di-subsampling oleh 2 ($\downarrow 2$) dengan cara mengabaikan setiap sampel yang kedua. Proses dekomposisi ini dapat melalui satu atau lebih tingkatan. Dekomposisi satu tingkat ditulis dengan ekspresi matematika pada persamaan (2) dan (3).

$$y_{k,tinggi} = \sum_n x_n h_{2k-n} \tag{1}$$

$$y_{k,rendah} = \sum_n x_n g_{2k-n} \tag{2}$$

$y_{k,tinggi}$ dan $y_{k,rendah}$ merupakan hasil dari tapis frekuensi tinggi dan tapis frekuensi rendah. $y_{k,tinggi}$ disebut sebagai koefisien transformasi *wavelet* diskret (Polikar, 1998). $y_{k,tinggi}$ merupakan detil dari informasi sinyal, sedangkan $y_{k,rendah}$ merupakan taksiran kasar dari fungsi pensakalaan. Dengan menggunakan koefisien *wavelet* diskret ini maka dapat dilakukan proses *Inverse Discrete Wavelet Transform* (IDWT) untuk merekonstruksi menjadi sinyal asal.

Proses rekonstruksi diawali dengan menggabungkan koefisien *wavelet* dari yang berada pada akhir dekomposisi dengan sebelumnya meng-upsampling oleh 2 ($\uparrow 2$) melalui tapis frekuensi tinggi dan tapis frekuensi rendah. Proses rekonstruksi ini sepenuhnya merupakan kebalikan dari proses dekomposisi sesuai dengan tingkatan pada proses dekomposisi. Sehingga persamaan rekonstruksi pada masing-masing tingkatan dapat ditulis:

$$x_n = \sum_k (y_{k,tinggi} h_{2k-n} + y_{k,rendah} g_{2k-n}) \quad (3)$$

2.3 Ekstraksi Fitur

Data hasil transformasi *wavelet* tersebut kemudian akan dilakukan ekstraksi fitur, yang berfungsi untuk mengelompokkan data sinyal getaran berdasarkan kesamaan fitur.

Ekstraksi fitur melibatkan transformasi linier atau non linier dari ruang fitur asli ke dalam dimensi baru yang lebih kecil. Ekstraksi fitur juga tidak mengurangi dimensi dari vektor yang diumpankan ke sistem klasifikasi, sedangkan seleksi fitur biasanya mengurangi jumlah fitur asli dengan memilih subset dari fitur tersebut, sehingga seleksi fitur masih mempertahankan cukup informasi untuk proses klasifikasi (Niu dkk., 2005).

Ekstraksi fitur adalah proses untuk menghitung fitur-fitur statistika yang diperlukan untuk mengelompokkan data-data getaran berdasarkan kemiripan atau kesamaan ciri yang berguna untuk proses identifikasi.

Fitur statistik yang digunakan ada 21 fitur, yaitu:

1. Mean
2. Root Mean Square (RMS)
3. Shape Factor (SP)
4. Skewness (c3)
5. Kurtosis (c4)
6. Crest Factor
7. Estimation
8. Error
9. Histogram Upper
10. Histogram Lower
11. Root Mean Square Frequency (RMSF)
12. Frequency Center (FC)
13. Root Varian Frequency (RVF)
14. Auto Regression (Ar) 2
15. Ar 3
16. Ar 4
17. Ar 5

18. Ar 6
19. Ar 7
20. Ar 8
21. Ar 9

2.4 Reduksi Fitur menggunakan PCA

Data hasil ekstraksi fitur tersebut kemudian akan direduksi dimensinya menggunakan *Principal Component Analysis* (PCA). PCA bisa dipandang sebagai metode klasik dari metode-metode analisis statistik multi variabel untuk memperoleh matrik yang tereduksi dimensinya. Karena adanya fakta bahwa sekumpulan data variabel yang tidak terkorrelasi dengan dimensi yang lebih kecil lebih mudah dipahami dan digunakan untuk analisa lebih jauh lagi dibandingkan data yang berdimensi lebih besar, maka teknik kompresi data ini banyak dipakai untuk analisa klaster, visualisasi dari data dimensi tinggi, regresi, kompresi data, dan pengenalan pola.

Prosedur dalam melakukan PCA adalah sebagai berikut:

- a. Menghitung matrik kovarian

input vektor $x_t (t=1, \dots, l)$ dan $x_t = 0$) dengan dimensi m $x_t = [x_t(1), x_t(2), \dots, x_t(m)]^T$ biasanya $m < l$, setiap vector x_t ditransformasikan secara linear kedalam satu vektor baru s yang dinyatakan sebagai :

$$s_t = U^T \cdot x_t \quad (4)$$

U adalah matrik orthogonal $m \times m$ dengan kolom ke i , u_i adalah nilai *eigenvector* dari sampel matrik kovarian:

$$C = \frac{1}{l} \sum_{t=1}^l x_t \cdot x_t^T \quad (5)$$

- b. Menghitung *eigenvalue* dan *eigenvector* dari matrik kovarian

$$\lambda_i U_i = C \cdot u_i, \quad i = 1, \dots, m \quad (6)$$

dimana i adalah salah satu *eigenvalue* dari C , u_i adalah nilai *eigenvector*.

- c. Berdasarkan nilai estimasi u_i dan komponen s_t , yang kemudian dihitung sebagai transformasi orthogonal dari x_t

$$s_t(i) = u_i^T x_t, \quad i = 1, \dots, m \quad (7)$$

- d. Komponen yang baru tersebut disebut dengan *principal component*. Dengan menggunakan hanya beberapa nilai pertama *eigenvector* yang telah diurutkan berdasarkan nilai *eigennya*, jumlah *principal component* dari s_t dapat direduksi. *Principal component* dari PCA mempunyai properti $s_t(i)$ yang tidak saling terkorrelasi, mempunyai varian maksimum yang berurutan dan estimasi error rata-rata dari representasi data input asli adalah minimal.

2.5 Klasifikasi Naïve Bayes

Melalui klasifikasi *Naïve Bayes*, setelah dilakukan pengamatan, informasi dalam distribusi prior

dikombinasikan dengan informasi melalui data sampel, kemudian hasilnya dinyatakan dalam bentuk distribusi posterior yang menjadi dasar klasifikasi *Naïve Bayes* untuk inferensi (Berger, 1990). Klasifikasi *Naïve Bayes* dapat digunakan untuk memprediksikan objek kelas yang labelnya tidak diketahui atau dapat memprediksikan data yang akan muncul di masa depan.

Klasifikasi *Naïve Bayes* merupakan suatu metode klasifikasi yang berdasarkan probabilitas dari teorema *Bayesian* dengan asumsi bahwa setiap variabel *X* bersifat bebas (*independent*). Dengan kata lain, klasifikasi *Naïve Bayes* mengasumsikan bahwa keberadaan sebuah atribut (variabel) tidak ada kaitannya dengan keberadaan atribut yang lain.

Persamaan *Naïve Bayes* dapat dituliskan sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (8)$$

dimana:

X= Data sampel dengan kelas yang tidak diketahui

H= Hipotesa bahwa *X* adalah data dengan kelas/label

P(X|H)= Peluang data sampel *X*, bila diasumsikan hipotesa *H* benar

P(H)= Peluang dari hipotesa *H* (*likelihood*)

P(X)= Peluang data sampel yang diamati (*evidence*)

3. Metodologi

3.1 Alat Penelitian

Alat yang digunakan dalam penelitian ini diantaranya: perangkat keras berupa Laptop Asus Prosesor Intel® Core™ i3-3217U CPU @ 1.80GHz 1.70GHz, RAM 2.00 GB, Hardisk 464 GB dan perangkat lunak yang digunakan adalah Microsoft Windows 7 ultimate (64-bit) dan Matlab 7.10.0 R2014a (64-bit).

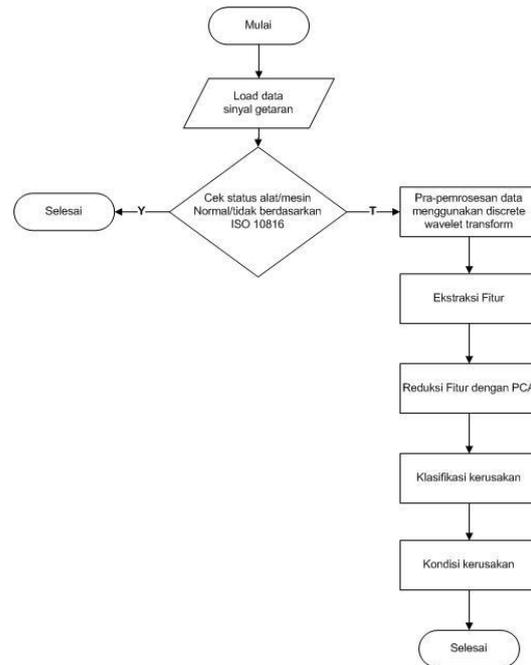
3.2 Bahan Penelitian

Bahan penelitian yang digunakan adalah berupa data-data getaran mesin yang diakuisisi dari sensor pada contoh mesin yang mengalami kerusakan dan mesin yang berada dalam kondisi normal. Data tersebut adalah:

- Data getaran mesin hasil percobaan di Fakultas Teknik Mesin Universitas Diponegoro.
- Data getaran mesin yang diambil *Bearing Data Center*
- Data getaran mesin yang diambil dari *data set bearing MPFT*.

3.2 Prosedur Penelitian

Prosedur penelitian diagnosa kerusakan *bearing* menggunakan *Principal Component Analysis (PCA)* dan Klasifikasi *Naïve Bayes* diperlihatkan pada Gambar 3.



Gambar 3. Prosedur Penelitian

Penjelasan tahapan prosedur penelitian:

a. Data Sinyal Getaran

Data sinyal getaran adalah data hasil pengukuran mesin yang diperoleh dalam 4 kondisi bantalan yaitu *ball fault bearing*, *innerace fault bearing*, *outerace fault bearing* dan *normal bearing*.

b. Pengecekan Severitas

Pengecekan severitas adalah proses awal untuk mengecek kondisi data getaran mesin berdasarkan nilai *vrms*-nya. Pengecekan ini dilakukan dengan cara membandingkan nilai *vrms* data dengan Tabel ISO 10816.

c. Pra-pemrosesan Data

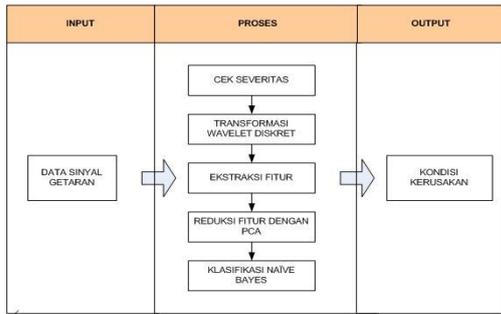
Pra-pemrosesan data adalah proses membersihkan data, mengintegrasikan data, mentransformasikan data dan mereduksi data. Pra-pemrosesan data yang dilakukan pada penelitian ini terdiri dari proses transformasi wavelet diskret, ekstraksi fitur dan reduksi fitur.

d. Klasifikasi Kerusakan

Data getaran mesin yang digunakan untuk penelitian akan dilakukan proses klasifikasi dengan menggunakan klasifikasi *Naïve Bayes*. Proses ini dilakukan dengan cara melakukan pengenalan pembelajaran sistem dengan menggunakan data pelatihan. Kemudian dilakukan proses pengujian yang berfungsi untuk mengukur keakuratan dari sistem yang sudah dilatih dengan data pelatihan.

3.3 Kerangka Sistem Informasi

Kerangka sistem informasi pada penelitian ini dapat dilihat pada Gambar 4 berikut ini.



Gambar 4. Kerangka sistem informasi

Penjelasan kerangka sistem informasi diagnosa kerusakan *bearing* menggunakan *principal component analysis* (PCA) dan klasifikasi Naïve Bayes adalah sebagai berikut:

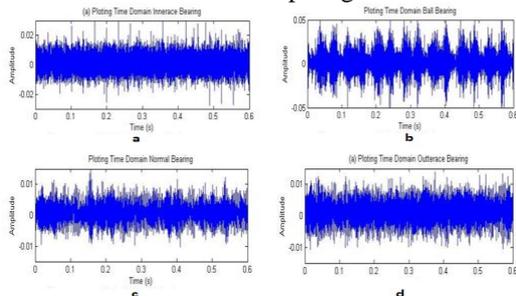
1. Input atau Masukan
Pada tahap ini pengguna memasukkan data getaran mesin yang akan dilakukan klasifikasi.
2. Proses
Data getaran mesin akan dicek tingkat severitasnya terlebih dahulu berdasarkan tabel standar ISO 10816 dengan cara menghitung nilai *vrms*-nya. Jika data dalam kondisi tidak normal, maka dilanjutkan dengan proses transformasi data menggunakan *wavelet* diskret, kemudian diekstraksi fitur. Fitur-fitur yang sudah dihitung tersebut kemudian direduksi menggunakan PCA. Selanjutnya data hasil PCA tersebut sebagian digunakan sebagai data pelatihan untuk pembelajaran sistem.
3. Output atau Keluaran
Setelah semua proses dilakukan, kemudian dilakukan pengujian data sehingga dapat ditampilkan kondisi kerusakan pada data uji tersebut.

4. Hasil dan Pembahasan

4.1 Skenario Penelitian Pertama

4.1.1 Data Sinyal Getaran

Data yang digunakan pada percobaan ini adalah data getaran *bearing* yang diperoleh dari Fakultas Teknik Mesin Universitas Diponegoro.



Gambar 5. Data sinyal getaran skenario pertama, (a) Innerace bearing fault, (b) Ball bearing fault, (c) Normal bearing, (d) Outerace bearing fault

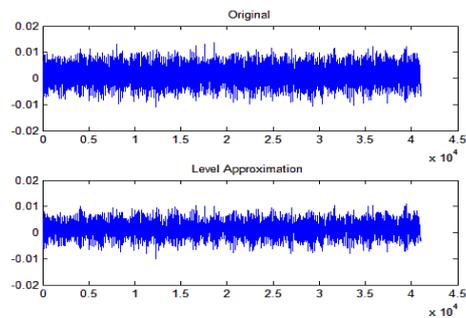
Data sinyal getaran diperoleh dari pengukuran pada 4 kondisi *bearing* yaitu *ball fault bearing*,

innerace fault bearing, *outerace fault bearing* dan *normal bearing*. Data diambil pada *sampling rate* 68kS/s. Setiap kondisi diukur sebanyak 20 kali, dimana dalam setiap pengukuran diperoleh data getaran sebanyak 40960 data.

4.1.2 Pra-pemrosesan Data Menggunakan Transformasi Wavelet Diskret

Pra-pemrosesan data menggunakan metode *wavelet* ini bertujuan untuk melakukan *denoising* sinyal. Sinyal pertama-tama dilewatkan pada rangkaian filter *high-pass* dan *low-pass*, kemudian setengah dari masing-masing keluaran diambil sebagai sampel melalui operasi *sub-sampling*. Proses ini disebut sebagai proses dekomposisi satu tingkat. Keluaran dari filter *low-pass* digunakan sebagai masukan di proses dekomposisi tingkat berikutnya.

Proses ini diulang sampai tingkat dekomposisi yang diinginkan. Gabungan dari keluaran-keluaran filter *high-pass* dan satu keluaran filter *low-pass* yang terakhir, disebut sebagai koefisien *wavelet*, yang berisi informasi sinyal hasil transformasi yang telah terkompresi. Hasil transformasi *wavelet* diskret pada sinyal getaran ditunjukkan pada Gambar 6.

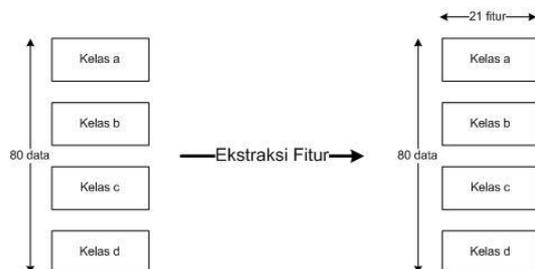


Gambar 6. Hasil transformasi *wavelet* diskret skenario pertama

4.1.3 Ekstraksi Fitur

Ekstraksi fitur bertujuan untuk mengelompokkan data getaran berdasarkan kesamaan fiturnya. Fitur statistik yang digunakan berjumlah 21 fitur.

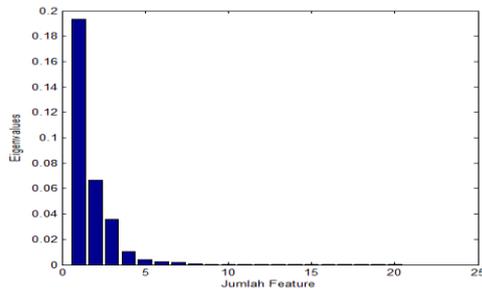
Sehingga dari data sinyal getaran sebanyak 80 buah tersebut setelah dilakukan ekstraksi fitur akan menghasilkan data dengan dimensi matrik [80x21], seperti ditunjukkan pada Gambar 7.



Gambar 7. Hasil ekstraksi fitur

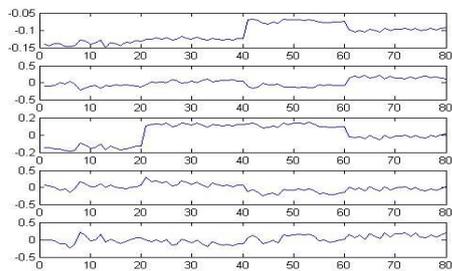
4.1.4 Pricipal Component Analysis (PCA)

PCA bertujuan untuk mereduksi fitur sehingga semakin sedikit data, algoritma data mining akan semakin cepat, dan akurasi dari proses data mining menjadi lebih tinggi. Reduksi fitur ini dapat dilakukan dengan cara memilih 5 nilai *eigen* yang terbesar.

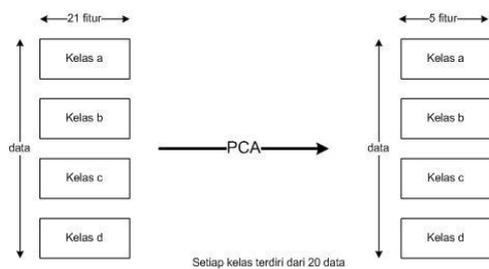


Gambar 8. Lima nilai *eigen* terbesar hasil PCA skenario pertama

Principal component dari sinyal getaran diperlihatkan pada Gambar 9. Setiap *principal component* terdiri atas 4 kelas atau kondisi *bearing* dengan pengukuran sebanyak 20 kali.



Gambar 9. Lima *principal component* skenario pertama



Gambar 10. Dimensi data setelah dilakukan PCA skenario pertama

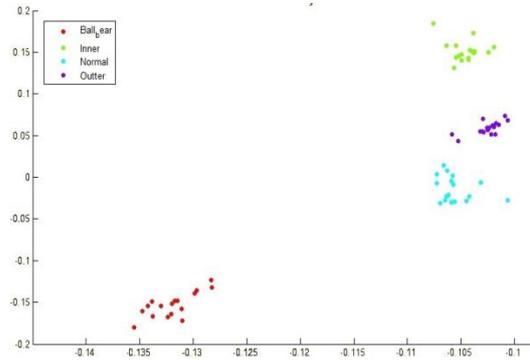
Dimensi data baru yang berukuran [80x5] inilah yang selanjutnya akan digunakan dalam proses pelatihan dan pengujian.

4.1.5 Pelatihan dan pengujian

a. Pelatihan

Data pelatihan yang digunakan untuk membangun model klasifikasi *Naive Bayes* ini diambil sebanyak 17 data pada setiap kelas atau kondisi. Sehingga total data training yang digunakan adalah 68 data. Kelas a

adalah kondisi *Ball bearing fault*, kelas b *Innerace bearing fault*, kelas c *Normal bearing*, dan kelas d *Outterace bearing fault*. Setelah diumpankan ke sistem diperoleh hasil seperti pada Gambar 11. dibawah ini.



Gambar 11. Hasil klasifikasi data pelatihan pada penelitian skenario pertama

b. Pengujian

Data pengujian yang digunakan pada penelitian skenario kedua ini adalah sebanyak 3 data pada setiap kelas atau kondisi. Sehingga total data pengujian yang digunakan adalah 3data x 4 kelas = 12 data. Hasil pengujian terhadap data uji ditunjukkan pada Tabel 1. Dari tabel tersebut dapat dihitung nilai akurasi dari tingkat pengenalan data ujinya adalah sebesar 100% benar, karena tidak ada data yang salah terklasifikasi. Akurasi digunakan juga untuk mengevaluasi performa model yang dibangun. Akurasi dapat dihitung dengan menggunakan persamaan :

$$Akurasi = \frac{\text{jumlah prediksi benar}}{\text{jumlah total prediksi}} \tag{9}$$

Keterangan:

jumlah prediksi benar = hanya data yang diklasifikasikan benar atau seluruh jumlah data *true* (*true positive* dan *true negative*)

jumlah total prediksi = seluruh jumlah data, baik yang diklasifikasikan *true* maupun *false* (*true positive*, *true negative*, *false positive*, dan *false negative*)

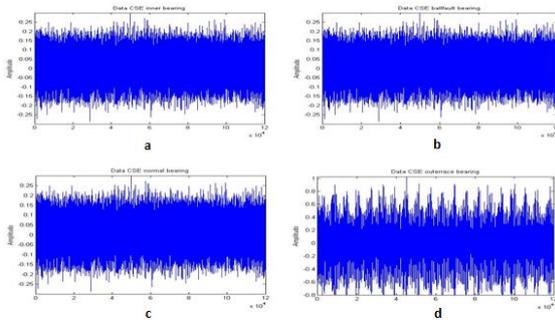
Tabel 1. Hasil pengujian penelitian skenario pertama

No	Kelas yang diuji	Hasil Prediksi
1	<i>ballbear</i>	<i>ballbear</i>
2	<i>ballbear</i>	<i>ballbear</i>
3	<i>ballbear</i>	<i>ballbear</i>
4	<i>inner</i>	<i>inner</i>
5	<i>inner</i>	<i>inner</i>
6	<i>inner</i>	<i>inner</i>
7	<i>normal</i>	<i>normal</i>
8	<i>normal</i>	<i>normal</i>
9	<i>normal</i>	<i>normal</i>
10	<i>outer</i>	<i>outer</i>
11	<i>outer</i>	<i>outer</i>
12	<i>outer</i>	<i>outer</i>

4.2 Skenario Penelitian Kedua

4.2.1 Data Sinyal Getaran

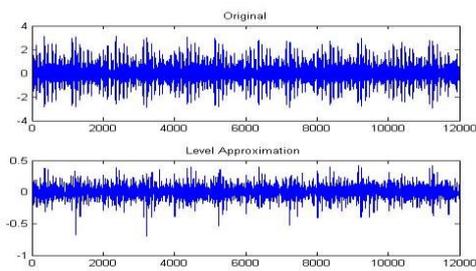
Data yang digunakan pada percobaan skenario kedua ini adalah Data CSE (*Columbia Southern Education*) yang diambil dari *Bearing Data Center* di <http://csegroups.case.edu/bearingdatacenter/>. Data sinyal getaran diperoleh dari pengukuran 4 kondisi *bearing* yaitu *ball fault bearing*, *innerace fault bearing*, *outterace fault bearing* dan *normal bearing*. Total pengukuran dilakukan sebanyak 50 kali, dimana setiap pengukuran diperoleh data getaran sebanyak 12000 data. Setiap kelas kondisi kerusakan memiliki karakteristik getaran sendiri-sendiri seperti yang diperlihatkan pada Gambar 12.



Gambar 12. Data sinyal getaran skenario kedua, (a) Innerace bearing fault, (b) Ball bearing fault, (c) Normal bearing, (d) Outterace bearing fault

4.2.2 Pra-pemrosesan Data Menggunakan Transformasi Wavelet Diskret

Hasil transformasi *wavelet* diskret sinyal getaran pada penelitian skenario kedua ditunjukkan pada Gambar 13.



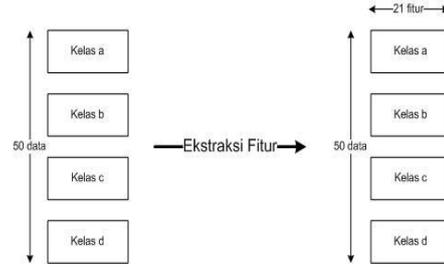
Gambar 13. Hasil transformasi *wavelet* diskret data skenario kedua

4.2.3 Ekstraksi Fitur

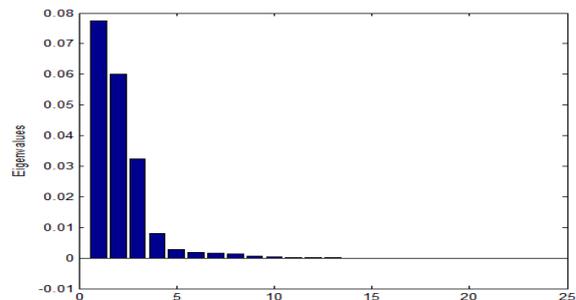
Hasil ekstraksi fitur dari data sinyal getaran sebanyak 50 buah data tersebut akan menghasilkan data dengan dimensi matrik [50x21], seperti ditunjukkan pada Gambar 14.

4.2.4 Principal Component Analysis (PCA)

Pada penelitian dipilih 5 fitur berdasarkan nilai *eigen* terbesar dari matrik kovarians. Gambar 15. menunjukkan lima nilai *eigen* terbesar hasil PCA pada Data CSE.



Gambar 14. Hasil ekstraksi fitur untuk data skenario kedua



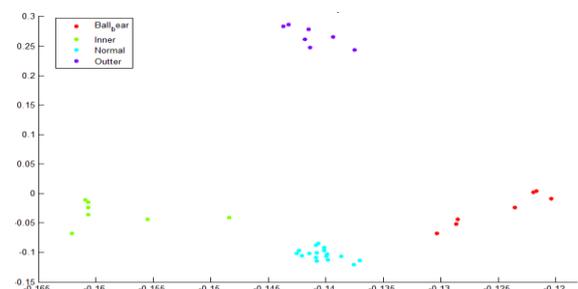
Gambar 15. Lima nilai *eigen* terbesar hasil PCA pada data skenario kedua

4.2.5 Pelatihan dan Pengujian

a. Pelatihan

Data pelatihan yang digunakan untuk membangun model klasifikasi *Naïve Bayes* diambil sebanyak 38 data dari seluruh kelas atau kondisi. Kelas a adalah kondisi *Ball bearing fault*, kelas b *Innerace bearing fault*, kelas c *Normal bearing*, dan kelas d *Outterace bearing fault*.

Data pelatihan tersebut kemudian diumpankan ke sistem dan diperoleh gambar seperti ditunjukkan pada Gambar 16.



Gambar 16. Hasil klasifikasi data pelatihan pada penelitian skenario kedua

b. Pengujian

Data pengujian yang digunakan pada penelitian skenario kedua ini adalah sebanyak 3 data pada setiap kelas atau kondisi. Sehingga total data pengujian yang digunakan adalah 3data x 4 kelas = 12 data.

Hasil pengujian terhadap data uji ditunjukkan pada Tabel 2. Dari tabel tersebut dihitung nilai akurasi dari tingkat pengenalan data ujinya adalah sebesar 100% benar, karena tidak ada data yang salah terklasifikasi.

Tabel 2. Hasil pengujian penelitian skenario kedua

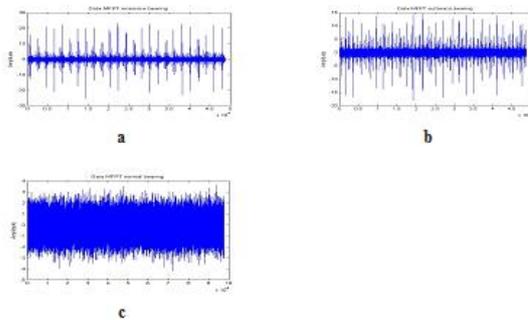
No	Kelas yang diuji	Hasil Prediksi
1	<i>ballbear</i>	<i>ballbear</i>
2	<i>ballbear</i>	<i>ballbear</i>
3	<i>ballbear</i>	<i>ballbear</i>
4	<i>inner</i>	<i>inner</i>
5	<i>inner</i>	<i>inner</i>
6	<i>inner</i>	<i>inner</i>
7	normal	normal
8	normal	normal
9	normal	normal
10	<i>outer</i>	<i>outer</i>
11	<i>outer</i>	<i>outer</i>
12	<i>outer</i>	<i>outer</i>

4.3 Skenario Penelitian Ketiga

4.3.1 Data Sinyal Getaran

Data yang digunakan pada percobaan skenario ketiga ini adalah data yang diambil dari *data set bearing MFPT (Machinery Failure Prevention Technology)* <http://data-acoustics.com/measurement/bearing-faults>.

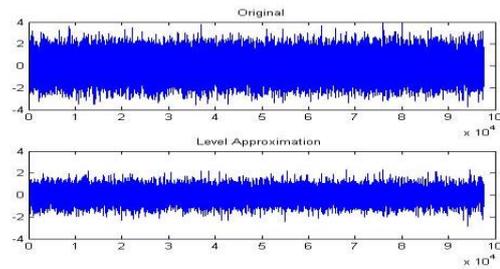
Data sinyal getaran ini diperoleh dari pengukuran 3 kondisi *bearing* yaitu *innerace fault bearing*, *outterace fault bearing* dan *normal bearing*. Ada total 78 kali pengukuran, dimana setiap pengukuran diperoleh data getaran sebanyak 48828 untuk data *innerace fault* dan *outterace fault*, serta 97656 data untuk kondisi *normal bearing*. Setiap kelas kondisi kerusakan memiliki karakteristik getaran sendiri-sendiri seperti yang diperlihatkan pada Gambar 17.



Gambar 17. Data sinyal getaran skenario ketiga, (a) *Innerace bearing fault*, (b) *Outterace bearing fault*, (c) *Normal bearing*

4.3.2 Pra-pemrosesan Data Menggunakan Transformasi Wavelet Diskret

Hasil transformasi *wavelet* diskret untuk data getaran pada skenario ketiga ditunjukkan pada Gambar 18.



Gambar 18. Hasil transformasi *wavelet* diskret skenario ketiga

4.3.3 Ekstraksi Fitur

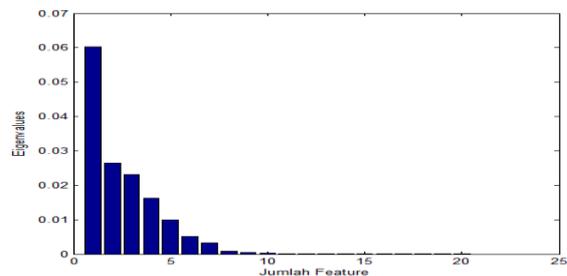
Hasil ekstraksi fitur dari data sinyal getaran sebanyak 78 buah data tersebut akan menghasilkan data dengan dimensi matrik [78x21], seperti ditunjukkan pada Gambar 19.



Gambar 19. Hasil ekstraksi fitur data skenario ketiga

4.3.4 Pricipal Component Analysis (PCA)

Hasil PCA untuk Data MFPT dengan cara memilih 5 fitur berdasarkan nilai *eigen* terbesar dari matrik kovariansnya ditunjukkan pada Gambar 20.



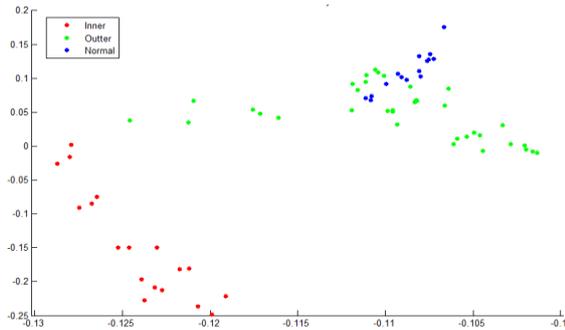
Gambar 20. Lima nilai *eigen* terbesar hasil PCA skenario ketiga

4.3.5 Pelatihan dan Pengujian

a. Pelatihan

Data pelatihan yang digunakan untuk membangun model klasifikasi *Naïve Bayes* diambil sebanyak 69 data dari total seluruh kelas atau kondisi. Kelas a adalah kondisi *Innerace bearing fault*, kelas b *Outterace bearing fault*, dan kelas c *Normal bearing*.

Data pelatihan tersebut kemudian diumpankan ke sistem dan diperoleh gambar seperti ditunjukkan pada Gambar 21.



Gambar 21. Hasil klasifikasi data pelatihan skenario ketiga

b. Pengujian

Data pengujian yang digunakan pada penelitian skenario ketiga ini adalah sebanyak 3 data pada setiap kelas atau kondisi. Sehingga total data pengujian yang digunakan adalah 3 data x 3 kelas = 9 data.

Hasil pengujian terhadap data uji ditunjukkan pada Tabel 3. Dari tabel tersebut dihitung nilai akurasi dari tingkat pengenalan data ujinya adalah sebesar 100% benar.

Tabel 3. Hasil pengujian skenario ketiga

No	Kelas yang diuji	Hasil Prediksi
1	inner	inner
2	inner	inner
3	inner	inner
4	outer	outer
5	outer	outer
6	outer	outer
7	normal	normal
8	normal	normal
9	normal	normal

4.4 Pembahasan

Berdasarkan hasil percobaan yang telah dilakukan dengan tiga skenario penelitian diatas maka dapat dihitung tingkat pengenalan pembelajaran dan tingkat pengujiannya. Pada skenario penelitian pertama didapatkan hasil tingkat pengenalan data ujinya adalah 100% karena tidak ada satupun data yang salah terklasifikasi. Sedangkan pada skenario penelitian yang kedua dan ketiga juga didapatkan hasil tingkat pengenalan data latihnya sebesar 100% karena tidak ada satupun data yang salah terklasifikasi.

Pre-processing data dengan menggunakan metode *wavelet* dan juga penerapan metode PCA menghasilkan tingkat akurasi yang tinggi.

5. Kesimpulan

Naïve Bayes Classifier sangat baik digunakan untuk mendiagnosis kerusakan *bearing*. Dengan menggunakan metode klasifikasi *Naïve Bayes* dan dikombinasi dengan penggunaan metode PCA (*Principal Component Analysis*) pada penelitian ini dapat diperoleh tingkat akurasi yang mencapai 100%.

PCA selain digunakan untuk mengurangi dimensi data dengan mempertahankan sebanyak mungkin

informasi dari dataset yang asli, juga berguna untuk mendapatkan proses klasifikasi yang lebih efisien dan lebih cepat.

Daftar Pustaka

- Anant, K.S., Dowla, F.U., 1997. Wavelet Transform Methods for Phase Identification in Three-Component Seismograms, *Bulletin of Seismological Society of America*, Vol. 87, No. 6, 1598 – 1612.
- Bengtsson, M., Olsson, E., Funk, P., Jackson, M., 2004. Technical design of condition based maintenance system-a case study using sound analysis and case based reasoning, *The 8th Maintenance and Reability Conference*, Knoxville USA, 2-5 Mei.
- Berger, C., 1990. Statistical Inference, Pacific Grove, California.
- Berry, M., Linoff, G.S., 2004. Data Mining Techniques Second Edition, Wiley Publishing, Inc., New York.
- Bolstad, W.M., 2007. Introduction to Bayesian Statistics Second Edition, A John Wiley & Sons, Inc., New York.
- Chu, P.S., Xin, Z., 2011. Bayesian analysis for extreme climatic events: A review, Hawaii, *Journal of Atmospheric Research* 102, 243-262
- Daubechies, C.I., 1992. Ten Lectures on Wavelet, SIAM, Philadelphia.
- Foster, D.J., Mosher, C.C., Hassanzadeh, S., 1994. Wavelet Transform Methods for Geophysical Applications, *Proceeding of SPIE* 2033, San Diego, 11 Juli, 1465–1468.
- Hand, J., Kamber M. 2001. Data Mining: Concepts and Techniques, Academic Press, Cambridge.
- Helle, A., 2006, Development of prognostic concepts and tools, *VTT Symposium* 243, Espoo, 12 Desember, 5-12.
- Hsu, C.C., Ping, H.Y., Keng-Wei, C., 2008. Extended Naïve Bayes classifier for mixed data, *Journal of Informations Sciences* 163, 103-122
- ISO 13373-1, 2002. Condition Monitoring and Diagnostics of Machines-Vibration Condition Monitoring, First Edition, International Organization Standart, Geneva.
- Jiang, X., Sankaran, M., 2007. Bayesian risk-based decision method for model validation under uncertainty, *Journal of Reliability Engineering & System Safety* 92, 707-718
- Joliffe, I. J., 1986. Principal Component Analysis, Springer, New York.
- Koc, L., Thomas, M., Shahram, S., 2012. A network intrusion detection system based on a hidden Naïve Bayes multiclass classifier, *Journal of Expert Systems with Applications* 39, 13492-13500
- Li, D., Zhen, Y.H., Feng, L.X., 2013. Prediction analysis of a wastewater treatment system using a

- Bayesian network, *Journal of Environmental Modelling & Software* 40, 140-150
- Li, G., Jing, S., 2012. Application of Bayesian methods in wind energy conversion systems, *Journal of Renewable Energy* 43, 1-8
- Menon, S., Schoes, J. N., Hamza, R., Bush, D., 2000. Wavelet-based acoustic emission detection method with adaptive thresholding, *Proceeding of SPIE 3986*, Newport Beach, 6 Maret, 71-77.
- Newland, D., E., 1994. Wavelet analysis of vibration, *Journal of Vibration and Acoustics*, 409-417
- Niu, X., Zhu, L., dan Ding, H., 2005. New Statistical Moments for Detection of Defects in Rolling Element Bearing, *Journal of Manufacture Technology*, 1268-1274.
- Polikar, R., 1998. Multi Resolution Analysis: The Discrete Wavelet Transform, Durham Computation Center, Iowa.
- Sripathi, D., 2003. Efficient Implementations of Discrete Wavelet Transform using FPGAs, Florida State University, Florida.
- Terzija, N., 2006. Robust Digital Image Watermarking Algorithms for Copyright Protection, Universität Duisburg, Essen.
- Thorp, B., 2007. What is predictive maintenance and how it has benefitted Seminole Electric Corp. Inc., *Journal of Asset Management and Maintenance*, 20(2), 14-21.
- Widodo, A., 2011. Application of Intelligent System for Machine Fault Diagnosis and Prognosis, Badan Penerbit Universitas Diponegoro, Semarang.