

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Metode Penelitian**

Penelitian mengenai peningkatan kecepatan prediksi produksi susu sapi ini menggunakan metode eksperimen dengan metode sebagai berikut:

##### **a. Pengumpulan data**

Data untuk penelitian prediksi produksi susu ini didapatkan dari Balai Besar Pemuliaan Ternak Unggul Sapi Perah Baturraden. Data berupa catatan hasil perah susu yang dilakukan setiap hari sebanyak dua kali pagi dan sore selama masa laktasi.

##### **b. Pengolahan data awal**

Data yang telah didapatkan selanjutnya dilakukan pengolahan awal hingga siap untuk dilakukan analisis. Dimana langkah-langkah yang dilakukan meliputi proses input data ke format yang dibutuhkan, penentuan atribut yang data dibutuhkan, serta pemecahan data yang digunakan untuk keperluan proses training dan testing.

##### **c. Eksperimen dan pengujian metode**

Pada tahap ini dari metode yang di usulkan dilakukan percobaan hingga membuktikan metode yang diusulkan pada tahap *pre-processing* memiliki keunggulan dari segi kecepatan pemrosesan.

##### **d. Evaluasi dan validasi**

Setelah dilakukan eksperimen, selanjutnya dilakukan evaluasi dan validasi untuk menguji tingkat kehandalan dari metode yang diusulkan. Proses evaluasi dan validasi yang digunakan adalah kecepatan yang dibutuhkan oleh model dalam satuan detik, *Root Mean Square Error* (RMSE) dan Prediction Trend Accuracy.

### 3.2 Metode pengumpulan data

### 3.3 Metode pengolahan data awal

Data yang didapat dari BBPTU Sapi Perah Baruttaden berupa catatan harian pemerahan susu sapi yang dilaksanakan tiap pagi dan sore. Data catatan tersebut terdiri dari kolom ID Sapi, tanggal perah, laktasi ke, pemerahan pagi, pemerahan sore, dan kolom kandang.

1	id sapi	tanggal	laktasi	pagi	sore	kandang
168362	GID012212071#1886-07	25/04/2011	1	8	7	G
168363	GID011805031#1521-03	25/04/2011	5	8.2	8.1	G
168364	GID012001081#1889-08	25/04/2011	1	10	6.5	A
168365	GID012111081#1912-08	25/04/2011	1	7.1	5	A
168366	GID060212081#1913-08	25/04/2011	1	6	4.5	G
168367	GID012104031#1506-03	25/04/2011	5	5	5	A
168368	GID011404081#1895-08	25/04/2011	1	6	5	A
168369	GID013108041#004	19/04/2011	4	8	9.2	A
168370	GID021308041#008	19/04/2011	4	12	13.4	A
168371	GID011508041#009	19/04/2011	5	6	6	A
168372	GID021009041#027	19/04/2011	4	12	11.6	A
168373	GID011005041#028	19/04/2011	4	7	5.2	A
168374	GID020905041#030	19/04/2011	4	9	8	A
168375	GID011306041#034	19/04/2011	3	14	13.4	A
168376	GID011003041#045	19/04/2011	4	8	7	A
168377	GID012902041#1637-04	25/04/2011	4	11.7	12	A
168378	GID011703041#046	19/04/2011	5	9.4	9	A
168379	GID012704041#050	19/04/2011	4	8	7.2	A
168380	GID010204021#0024	25/04/2011	6	5.2	4.9	A
168381	GID020505041#054	19/04/2011	4	10	7.4	A
168382	GID013010041#1693-04	25/04/2011	3	7	7	A
168383	GID011611031#055	19/04/2011	3	12	13	A
168384	GID010703041#059	19/04/2011	5	1	4	A
168385	GID021708081#1905-08	25/04/2011	1	7	7.3	A
168386	GID020804061#1779-06	19/04/2011	3	8	8	A
168387	GID021203041#068	19/04/2011	4	7.6	8	A

Gambar 9. Data catatan harian pemerahan susu sapi.

Data mentah tersebut yang diambil dari tanggal 1 Oktober 2007 hingga 29 Juli 2011 terdiri dari 179.449 baris dari sapi yang berjumlah 338 ekor. Dari sekian banyak sapi, dipilih seekor sapi secara acak yang telah mengalami masa laktasi hingga 5 kali. Sapi tersebut memiliki kode GID010405041#073 dengan catatan pemerahan harian sebanyak 1163 record. Selanjutnya dilakukan data cleaning diantaranya adalah:

- Dari sebanyak 5 kali laktasi yang memiliki rekord sebanyak 1163 baris selanjutnya diambil hanya pada laktasi pertama saja, sehingga jumlah *record*-nya menjadi 251 baris.

- b. Membuang beberapa variabel yang tidak diperlukan yaitu ID sapi, laktasi ke, kandang. Sedangkan kolom tanggal perah, pemerahan pagi, pemerahan sore masih tersisa untuk masuk ke tahap pengolahan berikutnya.
- c. Pada variabel pemerahan pagi dan sore terdapat beberapa baris yang tidak diisi dengan lengkap. Keadaan ini menjadikan validitas data berkurang. Untuk itu perlu ada perlakuan khusus terhadap *record* yang tidak terisi lengkap. Menurut Larose [26] dalam bukunya mengatakan bahwa ada beberapa cara untuk memperlakukan record yang tidak berisi nilai. Diantaranya adalah dengan memasukkan nilai tertentu secara langsung. Cara kedua adalah dengan mencari nilai rata-rata dari pada kolom pagi atau sore. Namun metode ini mengandung resiko karena jika data null ada di bagian bawah maka akan mendapatkan nilai yang cukup tinggi. Hal ini tentu akan bertentangan dengan grafik laktasi seperti terlihat pada gambar 2 dimana terjadi penurunan yang cukup drastis pada tahap-tahap akhir. Cara yang ketiga adalah dengan mengisikan nilai secara acak dari nilai yang ada pada kolom pagi dan sore. Hal ini juga memiliki resio yang sama seperti pada cara sebelumnya. Untuk itu penulis memilih cara yang pertama yaitu dengan memasukkan nilai tertentu. Nilai tertentu disini diambil dari baris yang ada di atasnya yang terisi.
- d. Pada tahap ini dilakukan *replace missing value* dengan pola *replacement previous value*.

	A	B	C	D
1	tanggal	xlaktasi	pagi	sore
384	21/11/2007	2	2	4
385	22/11/2007	2	4	3
386	23/11/2007	2	3	2
387	20/01/2008	2	NULL	9
388	24/01/2008	2	6	NULL
389	25/01/2008	2	5	NULL
390	13/07/2008	3	13	13,2
391	14/07/2008	3	12,2	13
392	15/07/2008	3	12,2	11,2
393	16/07/2008	3	12	12,3
394	17/07/2008	3	12	13
395	18/07/2008	3	14	13,8
396	19/07/2008	3	11,8	13

Gambar 10. Replace missing value.

Misalnya pada baris ke 146-148 pemerahan pagi tidak diisi, maka untuk selanjutnya bisa diambilkan nilai dari record sebelumnya.

- e. Untuk membentuk data menjadi *univariat*, data pemerahan pagi dan pemerahan sore dijumlahkan. Sehingga menghasilkan tanggal perah dan jumlah laktasi per hari.

Selanjutnya pada tahap *pre-processing*, yaitu dengan melakukan reduksi data menggunakan algoritma *Haar wavelet discrete wavelet transform*. Seperti dikatakan Ian Kaplan dalam [23] bahwa *Haar wavelet transform* memiliki beberapa kelebihan diantaranya:

- a. Memiliki konsep yang sederhana.
- b. Proses cepat.
- c. Efisiensi memori, dimana proses penghitungan tidak memerlukan *array* temporary.
- d. Operasi *inverse* dilakukan dengan hasil yang akurat dan sempurna tanpa menghasilkan *edge effect* seperti pada transformasi *wavelet* lainnya.

Data produksi susu pada laktasi pertama yang berjumlah 251 baris selanjutnya dilakukan operasi wavelet menggunakan algoritma Haar Wavelet sebanyak satu level. Langkah-langkah transformasi wavelet menggunakan algoritma Haar Wavelet adalah sebagai berikut:

- a. Menghitung *high pass filter* dengan rumus  $(S_{n+1} + S_n)/2$  sehingga dihasilkan nilai *Wavelet approximation coefficient* (WAC). Nilai WAC inilah yang akan dilakukan pemrosesan prediksi menggunakan algoritma neural network.
- b. Menghitung *low pass filter* dengan rumus  $(S_{n+1} - S_n)/2$  sehingga dihasilkan nilai *Wavelet detail coefficient* (WDC). Dinama nilai ini nantinya berfungsi untuk melakukan inverse.

Pada tahap *post-processing* (tahap setelah memperoleh hasil prediksi dari data yang telah direduksi) akan dilakukan proses rekonstruksi, yaitu proses untuk mengembalikan nilai dari data yang telah direduksi. Misalnya data sebelum dilakukan reduksi berjumlah 251 baris, setelah dilakukan reduksi dengan *Haar*

*Wavelet Transform* menjadi 125 baris maka dikembalikan lagi menjadi 251 baris. Proses rekonstruksi yang dilakukan menggunakan rumus:

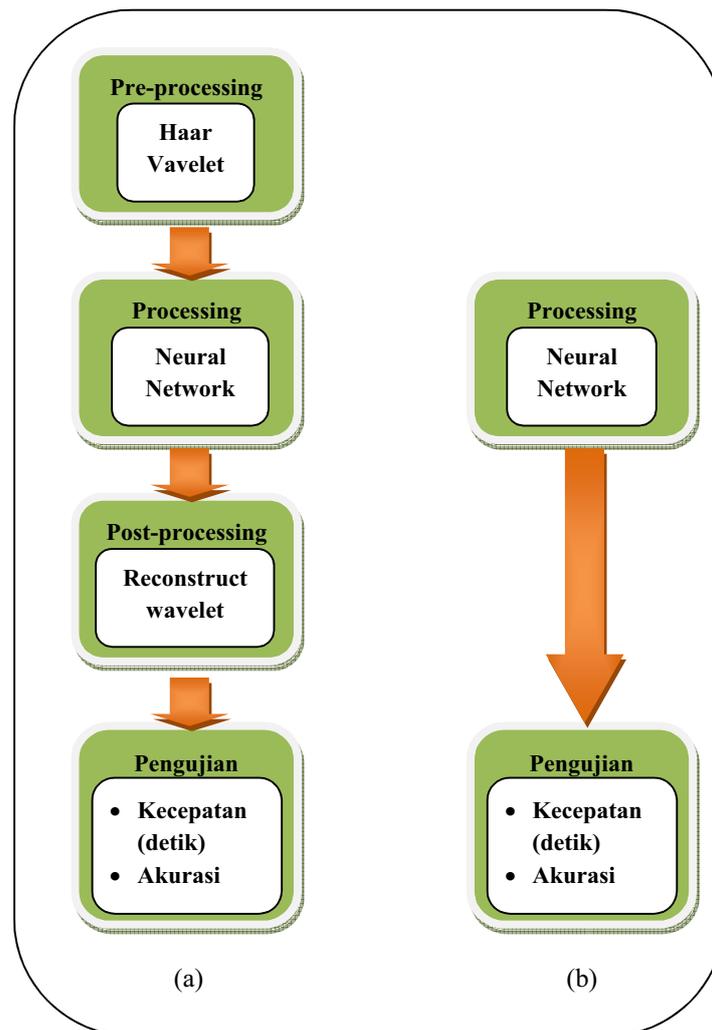
- a. Untuk  $S_{n+1}$  digunakan rumus  $WDC_n - WAC_n$
- b. sedangkan untuk  $S_n$  digunakan rumus  $WDC_n + WAC_n$ .

Nilai hasil rekonstruksi ini selanjutnya dilakukan evaluasi terhadap kecepatan dan akurasi hasil prediksi.

### 3.4 Eksperimen dan pengujian model

Eksperimen dilakukan menggunakan metode *trial and error* yang hasilnya diuji dengan membandingkan kecepatan dan akurasinya dengan tanpa implementasi model diatas. Seperti tercantum dalam kerangka pemikiran bahwa selain pada reduksi data, variabel-variabel lain yang diobservasi adalah *hidden layer*, jumlah *neuron*, dan fungsi aktivasi. Observasi pada variabel-variabel tersebut dimaksudkan untuk mencari arsitektur network yang paling efisien untuk menyelesaikan masalah ini. Untuk mengetahui arsitektur neural network yang tepat maka dilakukan pengukuran tingkat error menggunakan rumus (5) *Root Mean Square Error* seperti pada sub bab 2.4.

*Hidden layer* sebenarnya bisa ditentukan lebih dari satu lapis, namun sangat jarang peneliti yang mengimplementasikan arsitektur neural network yang terdiri lebih dari 1 lapis, apalagi lebih dari itu. Untuk penelitian ini, ditentukan *hidden layer* sebanyak satu lapis saja. Eksperimen berikutnya adalah dilakukan penentuan jumlah neuron pada hidden layer. Yang perlu diperhatikan pada tahap penentuan neuron ini adalah jika jumlahnya terlalu sedikit maka jaringan kurang dapat mendeteksi pola dalam data. Sedangkan jika neuron terlalu banyak maka jumlah informasi dalam training set menjadi terbatas, sehingga tidak cukup untuk melatih semua neuron dalam hidden layer.



**Gambar 11. Alur eksperimen**

Pada gambar 11 bagian (a) adalah alur eksperimen pada metode yang diusulkan, sedangkan bagian (b) adalah pemrosesan data tanpa proses reduksi data. Hasil kedua eksperimen tersebut selanjutnya dilakukan perbandingan metode mana yang memiliki keunggulan. Nilai atribut dan parameter pada tahap processing dibuat sama agar pengujian dan perbandingan lebih obyektif.

### **3.5 Perangkat pengujian model**

Salah satu komponen yang diuji adalah mengenai kecepatan, oleh karena itu perlu diinformasikan mengenai spesifikasi perangkat keras yang digunakan untuk pengujian ini. Perangkat tersebut penting diketahui karena jika diuji dengan spesifikasi yang berbeda tentu hasilnya juga akan berbeda.

Spesifikasi perangkat keras yang digunakan dalam pengujian ini adalah:

- a. Processor intel Core 2 Duo T5870 2GHz
- b. RAM 3 GB

Sedangkan perangkat lunak yang digunakan adalah:

- a. Windows 7 Professional Edition.
- b. Rapid Miner Versi 5.2.001
- c. Microsoft Excel 2010