

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Konsumsi susu sapi segar di Indonesia sejak beberapa tahun terakhir mengalami peningkatan, hingga tahun 2011 tercatat 11 liter per tahun per kapita. Namun peningkatan tersebut masih relatif kecil jika dibanding dengan negara tetangga seperti Malaysia yang mencapai tiga kali lipatnya. Untuk lebih meningkatkan lagi konsumsi susu sapi segar di Indonesia, Kementerian Pertanian pada tahun 2011 mencanangkan Gerakan Nasional Minum Susu Segar yang dilakukan di Kantor Dinas Perkebunan Propinsi Jawa Tengah di Ungaran pada bulan Juni 2011. Gerakan tersebut juga sekaligus untuk merangsang kenaikan produksi susu segar dalam negeri.

Produksi susu sapi segar di Indonesia pada tahun 2011 tercatat 92.580.000 ton, dimana hasil tersebut baru memenuhi sekitar 30% kebutuhan susu nasional. Sehingga untuk mencukupi 70% kekurangannya masih mengandalkan import dari berbagai negara. Keadaan ini menjadi alasan kuat untuk meningkatkan kapasitas produksi susu sapi segar nasional untuk memenuhi kebutuhan domestik.

Salah satu langkah yang diambil pemerintah untuk meningkatkan kapasitas produksi tersebut adalah dengan penyediaan bibit unggul. Dimana hal tersebut menjadi wewenang dan tanggungjawab Balai Besar Pembibitan Ternak Unggul (BBPTU) Sapi Perah Baturraden untuk menyediakan bibit unggul sapi perah untuk masyarakat peternak. Untuk pemilihan jenis sapi perah unggul dengan produktifitas susu tinggi, membuat kegiatan perencanaan kegiatan produksi susu, pengelolaan ekspor dan impor susu, maka perlu dilakukan prediksi produksi susu sapi.

Untuk melakukan prediksi produksi susu dapat dilakukan dengan berbagai metode data mining khususnya prediksi waktu. Gandhi et.al. [1] melakukan penelitian dengan membandingkan metode *Multiple Regression Analysis* (MRA) dan *Neural Network* (NN), dimana dia menunjukkan hasil tingkat akurasi 30,04% untuk algoritma neural network dan 25,62% untuk tingkat akurasi *multiple linear*

regression (MLR). Sementara itu perbandingan yang dilakukan oleh Grzsiak [2] menunjukkan tingkat *Mean Error Prediction* (MEP) yang masih tinggi yaitu 6,85% untuk NN dan 7,36% untuk *Multiple Linear Regression* (MLR).

Ruhil [3] melakukan penelitian khusus prediksi dengan algoritma neural network berdasar data parsial dengan hasil akurasi sebesar 80%. Perbandingan NN dan MLR juga dilakukan oleh Sharma [4] dengan hasil akurasi NN sebesar 88,15% dan MLR sebesar 86,31%. Secara garis besar, algoritma neural network memiliki kinerja dengan akurasi yang lebih tinggi dibanding dengan algoritma *multiple linear regression*.

Jeong [5] melakukan penelitian mengenai reduksi data berbasis *wavelet* untuk mendeteksi kesalahan proses. Data diambil dari proses produksi antenna dimana hasil deteksi menunjukkan metode *discrete wavelet transform* lebih baik dari fourier transform.

Namey et.al. [6] mengaplikasikan teknik reduksi dataset kualitatif yang berukuran besar. Penelitian dilakukan pada perilaku sosial mengenai penggunaan kata dalam percakapan atau penulisan. Para peneliti menggunakan analisis konten pada kata atau frase sebagai metode awal untuk menyaring data dan mengurangi ukuran dataset, atau mereka dapat melakukan koding secara struktural pada semua data sebelum melanjutkan ke tahap analisis. Metode reduksi data yang digunakan antara lain *Hierarchical Clustering Technique*, *Multidimensional Scaling (MDS)*.

Hierarchical Clustering menghasilkan suatu struktur data berbentuk pohon yang disebut dendogram dimana data dikelompokkan secara bertingkat dari yang paling bawah dimana tiap *instance* data merupakan suatu cluster tersendiri, hingga tingkat paling atas dimana keseluruhan data membentuk satu cluster besar berisi cluster-cluster yang lebih kecil [7]. Kelemahan dari teknik *clustering* ini pada ketidakjelasan kondisi penghentian dan juga kebanyakan algoritma dalam kategori ini tidak meninjau ulang *cluster* yang sudah terbentuk untuk memperbaiki kualitasnya.

Tujuan dari MDS adalah untuk memberi gambaran visual dari pola kedekatan yang berupa kesamaan atau jarak diantara sekumpulan obyek-obyek. Penerapan MDS dijumpai pada visualisasi ilmiah dan data mining dalam ilmu kognitif,

informasi. Dikatakan oleh Panagis Magalinos [8] bahwa kekurangan dari MDS adalah komputasional yang kompleks dan membutuhkan memori yang besar. Hal ini bisa diatasi dengan kelebihan dari pre-processing Haar wavelet karena komputasional sederhana sehingga waktu preprocessing lebih cepat.

Sangwoo Moon [9] dalam disertasinya mengajukan metode reduksi data menggunakan *Support Vector Machines* (SVM) dan *Independent Component Analysis* (ICA). Sangwoo menggabungkan kedua metode tersebut untuk SVM linier dan non linier plus ICA menunjukkan hasil yang lebih baik dibanding dengan metode konvensional lainnya baik pada pendekatan *supervised* maupun *unsupervised* dan *hybrid*.

Namun metode SVM memiliki kekurangan jika diimplementasikan dalam problem yang berskala besar, dalam hal ini adalah ukuran data yang diolah berukuran besar [10]. Karena metode SVM secara teoritik dikembangkan untuk masalah klasifikasi sehingga masih perlu pengkajian dan penelitian lebih lanjut lagi dalam hal implementasi pada masalah prediksi. Dalam kesimpulannya, Nugroho [10] juga mengatakan bahwa secara umum SVM menunjukkan kinerja yang lebih baik dari perceptron, walaupun demikian dikarenakan jumlah sampel yang relatif sedikit, hasil eksperimen yang dilakukannya belum dapat memberikan kesimpulan bahwa SVM superior atas perceptron. Dengan kata lain, Nugroho mengatakan jika perceptron –dalam hal ini *neural network*—kemungkinan besar memiliki kemampuan yang lebih baik dibanding SVM jika data yang diolah berjumlah banyak.

Seperti halnya SVM, kelemahan ICA juga pada komputasional yang kompleks sehingga membutuhkan waktu yang lama untuk menyelesaikan analisis [11]. Selain itu ternyata ICA memiliki dua ambiguitas yaitu ambiguitas terhadap penskalaan dan ambiguitas terhadap operasi permutasi [11].

Pada proses data mining, sebelum data dilakukan pemrosesan maka perlu dilakukan data *pre-processing*, yaitu menyiapkan data agar dapat dilakukan prediksi maksimal. Jika data yang akan diproses berukuran semakin besar, maka waktu yang diperlukan juga semakin lama. Pada penelitian ini, akan dilakukan penelitian

mengenai kecepatan kerja algoritma dan bagaimana caranya untuk memperkecil ukuran data (reduksi data) namun hasil prediksi tetap maksimal. Survey yang dilakukan oleh Chaovalit [12] mengenai data mining runtun waktu (*time-series prediction*) berbasis *discrete wavelet transform* (DWT), bahwa metode tersebut bisa diterapkan untuk mereduksi data *time series* pada tahap pre-processing data mining. Hal tersebut dimaksudkan untuk memperkecil ukuran data dan mempercepat proses. Selanjutnya dilakukan pengolahan dan analisis data yang telah siap menggunakan metode yang neural network. Menurut kaidah CRISP-DM dimana disana terdapat lima langkah data mining, pada langkah pre-processing diusulkan menggunakan metode *discrete wavelet transform* untuk melakukan reduksi data.

Penelitian-penelitian yang dilakukan oleh Gandhi[1], Gzesiak [2], Ruhil [3], dan Sharma [4] serta peneliti lainnya baru sampai pada tahap seleksi atribut yang paling berpengaruh terhadap produksi susu. Namun belum ada yang mengimplementasikan *pre-processing* dalam rangka untuk reduksi volume data. Langkah *pre-processing* ini cukup penting dilakukan mengingat volume data yang diolah dalam data mining yang sebenarnya (bukan dalam rangka penelitian) biasanya berukuran sangat besar karena terdiri dari rekaman beberapa tahun sebelumnya.

1.2 Rumusan Masalah

Rumusan masalah penelitian ini adalah:

1. Masih relatif sulit menentukan jenis sapi perah yang bisa dijadikan bibit unggul.
2. Masih diperlukannya tahap *pre-processing* lanjutan untuk mereduksi volume data agar proses prediksi bisa dilakukan dengan lebih cepat.

1.3 Tujuan penelitian

Meningkatkan kecepatan prediksi *time series* dan akurasi berbasis algoritma *neural network* melalui data *pre-processing* dengan metode *Discrete Wavelet Transform*.

1.4 Manfaat penelitian

- a. Mempercepat proses prediksi produksi susu di BBPTU Sapi Perah Baturraden.
- b. Untuk pemilihan jenis sapi perah unggul dengan produktifitas susu tinggi, membuat kegiatan perencanaan kegiatan produksi susu, pengelolaan ekspor dan impor susu, maka perlu dilakukan prediksi produksi susu sapi
- c. Menambah khasanah penelitian prediksi waktu terutama pada sisi kecepatan proses.
- d. Adanya peningkatan kecepatan proses prediksi berbasis neural network dengan mengimpelentasikan metode discrete wavelet transform pada tahap pre-processing.

1.5 Sistematika Penulisan

Sistematika penulisan pada penelitian ini adalah sebagai berikut:

Bab I Pendahuluan

Bab I berisi tentang latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

Bab II Tinjauan Pustaka

Bab II berisi penjelasan tentang tinjauan pustaka yang berkaitan dengan penelitian ini. Diuraikan pula mengenai tinjauan pustaka yang merupakan hasil dari penelitian-penelitian lainnya yang berkaitan dengan penelitian yang dilakukan. Pada bab memberikan penjelasan mengenai sumber acuan dari pustaka primer seperti jurnal, prosiding, buku, dan tulisan asli lainnya untuk mengetahui perkembangan penelitian yang relevan dengan judul atau tema penelitian yang dilakukan dan juga sebagai arahan dalam memecahkan masalah yang diteliti. Bab II ini juga menguraikan tentang kerangka pemikiran yang merupakan penjelasan tentang kerangka berpikir untuk memecahkan masalah yang sedang diteliti, termasuk menguraikan obyek penelitian.

Selain itu juga membahas tentang institusi tempat pengambilan sample data dilakukan. Uraian mengenai institusi mencakup latar belakang institusi, visi dan misi,

serta tugas pokok dan fungsi institusi Balai Besar Pembibitan Ternak Unggul (BBPTU) Sapi Perah Baturraden.

Bab III Metode Penelitian

Bab III berisi tentang penjelasan mengenai metode penelitian, penentuan masalah, penentuan *computing approach*, juga penjelasan bagaimana penerapan *computing approach* pada obyek penelitian, diuraikan pula cara evaluasi dan validasi hasil penerapan, metode pengumpulan data, metode analisis data, metode pengukuran penelitian.

Bab IV Hasil dan Pembahasan

Pada bab IV ini menguraikan tentang hasil penelitian lalu membahasnya mengenai implikasi dari penelitian dan eksperimen yang dilakukan. Hasil merupakan suatu penjelasan tentang pengolahan data dan interpretasinya, baik dalam bentuk deskriptif maupun inferensinya. Implikasi penelitian menjelaskan tentang tindak lanjut penelitian yang terkait dengan aspek manajerial, aspek sistem, maupun aspek penelitian lanjutan.

Bab V Kesimpulan dan Saran

Bab V ini berisi tentang ringkasan temuan, rangkuman kesimpulan dan saran. Kesimpulan merupakan pernyataan secara umum atau spesifik yang berisi hal-hal penting dan menjadi temuan penelitian yang bersumber pada hasil dan pembahasan. Saran merupakan pernyataan atau rekomendasi peneliti yang berisi hal-hal penting sebagaimana yang telah disampaikan yaitu implikasi penelitian.