

# IMPLEMENTASI DATA MINING MENGGUNAKAN CRISP-DM PADA SISTEM INFORMASI EKSEKUTIF DINAS KELAUTAN DAN PERIKANAN PROVINSI JAWA TENGAH

Indra Purnama, Ragil Saputra, Adi Wibowo

Universitas Diponegoro Semarang

ip.indrapurnama@googlemail.com, ragil.saputra@undip.ac.id, bowo.adi@gmail.com

## ABSTRAK

Dinas Kelautan dan Perikanan (DKP) Provinsi Jawa Tengah sebagai pengelola data produksi perikanan budidaya tingkat provinsi, memiliki tugas untuk membuat suatu laporan untuk keperluan Kepala Bidang Perikanan Budidaya di DKP dengan data yang bersumber dari hasil input data produksi pada Sistem Informasi Statistik Perikanan Bidang Budidaya (SIPBB) dari setiap kota/kabupaten di Provinsi Jawa Tengah. Data produksi perikanan budidaya dapat dimanfaatkan untuk proses penggalian pengetahuan tambahan mengenai potensi perikanan budidaya di Provinsi Jawa Tengah. Pengolahan *Data Mining* dengan menggunakan *Cross-Industry Standard Process for Data Mining (CRISP-DM)* terdiri dari fase pemahaman bisnis, pemahaman data, pengolahan data, permodelan, evaluasi, serta penyebaran agar didapatkan suatu hasil maksimal dari hasil pemrosesan data. Fase penyebaran pada *CRISP-DM* merupakan implementasi *data mining* yang diterapkan pada Sistem Informasi Eksekutif DKP untuk menggali pengetahuan mengenai potensi perikanan budidaya. Sistem informasi Eksekutif DKP dikembangkan dengan menggunakan bahasa pemrograman PHP dan *database management system* MySQL. Hasil yang diperoleh dari implementasi *data mining* pada Sistem Informasi Eksekutif DKP adalah berupa laporan mengenai potensi perikanan budidaya setiap kota/kabupaten di Provinsi Jawa Tengah setiap tahun.

**Kata kunci:** *Data Mining*, *CRISP-DM*, sistem informasi, Produksi perikanan budidaya.

## Pendahuluan

Dinas Kelautan dan Perikanan (DKP) Provinsi Jawa Tengah, merupakan lembaga negara yang bertugas melaksanakan semua pengelolaan kegiatan perikanan dan kelautan di wilayah Provinsi Jawa Tengah. Di dalam instansi DKP terdapat berbagai bidang yang terfokus pada tugas dan fungsi masing-masing.

Untuk menunjang pencapaian tugas pokok dan fungsinya, Dinas Kelautan dan Perikanan Provinsi Jawa Tengah ditopang oleh struktur/susunan organisasi sebagai berikut[1]:

- 1) Kepala Dinas
- 2) Sekertariat
- 3) Bidang Kelautan dan Pesisir dan Pulau-pulau Kecil
- 4) Bidang Perikanan Tangkap
- 5) Bidang Perikanan Budidaya
- 6) Bidang Usaha dan Pemasaran Ikan
- 7) Unit Pelaksana Teknis.

Dalam setiap bidang di DKP Provinsi Jateng ini menjalankan fungsi dan peranan masing-masing termasuk bidang perikanan budidaya. Sistem Informasi Perikanan Bidang Budidaya (SIPBB) merupakan sistem pengolah data di bidang perikanan budidaya di wilayah Provinsi Jawa Tengah dimana didalamnya mengolah data

perikanan bidang budidaya yang nantinya akan dihasilkan laporan-laporan detail mengenai hasil budidaya perikanan di Provinsi Jawa Tengah. [2]

Informasi yang dihasilkan dari SIPBB dibutuhkan oleh DKP Provinsi Jawa Tengah dijadikan sebagai laporan ke Kementerian Kelautan dan Perikanan (KKP) Republik Indonesia setiap tahun. Dalam proses evaluasi perlu adanya hasil analisa secara keseluruhan mengenai data yang diproses, sehingga diperlukan sistem yang mengolah data tersebut. Maka pada sistem tersebut dibutuhkan untuk dikembangkannya Sistem Informasi Eksekutif sebagai media pengolah data

Data perikanan budidaya memiliki jumlah yang banyak dan kompleks menyebabkan perlu adanya pemrosesan agar dapat mengolah data yang ada menjadi sebuah pengetahuan. *Data Mining* merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar[3]. Contoh implementasi data mining seperti Analisa Pasar dan Manajemen (contoh: menembak target pasar, identifikasi kebutuhan *customer*), Analisa Perusahaan dan Manajemen Resiko (contoh: *Business Intelligence*, *Resource Planning*,

Perencanaan Keuangan dan Evaluasi Aset), Telekomunikasi, Keuangan dan lain-lain[4]. Diantara berbagai contoh implementasi pada data mining tadi ada yang digunakan pula dalam Sistem Informasi Eksekutif seperti dalam *Resource Planning* yang dimiliki beberapa perusahaan besar di Indonesia.

Implementasi data mining pada sistem ini, diharapkan dapat menghasilkan pengetahuan lebih mengenai data perikanan di Provinsi Jawa Tengah, dan bisa menghasilkan suatu acuan yang nantinya dapat dimanfaatkan oleh penggunanya. Data yang akan diambil di khususkannya pada data produksi dan nilai produksi ikan. Pendekatan yang akan dilakukan yakni menggunakan *CRISP-DM* atau *Cross Industry Standard Process for Data Mining* satu standar pemrosesan *Data Mining* yang sudah dikembangkan dimana data yang ada akan melewati fase-fase yang ada yakni fase pemahaman bisnis, pemahaman data, pengolahan data, permodelan, pengujian, serta penyebaran sehingga bisa diperoleh pengetahuan tambahan dari data produksi dan nilai produksi yang ada. Sesuai dengan langkah-langkah yang ada tadi diharapkan dapat mendapatkan permodelan yang paling sesuai dalam pemrosesan data mining melalui pemahaman data yang ada, sehingga informasi yang dihasilkan lebih efisien[5]. Contoh dari implementasi *CRISP-DM* itu sendiri diantaranya *Identifying Bank Frauds Using CRISP-DM and Decision Trees*[6] yang menjelaskan mengenai bagaimana proses identifikasi tingkat kecurangan atau penipuan dalam dunia perbankan, yang didalamnya menggunakan teknik klasifikasi yang dilakukan pengujian berdasar data nyata dan melihat skema pemrosesan serta durasi eksekusi algoritma. Selain itu persentasi kesalahan pada proses *training* yang proses pengujiannya dapat dijadikan acuan dalam pemrosesan *data mining* pada penelitian ini. Contoh penelitian lain yakni *Developing Cyberspace Data Understanding Using CRISP-DM for Host IDS Feature Mining*[7] yang membahas mengenai pemahaman data *cyberspace* yang nantinya digunakan untuk keperluan militer, yang di dalamnya terdapat permodelan menggunakan teknik peramalan yang dapat dijadikan acuan pada penelitian ini.

### **Sistem Informasi**

Sistem informasi merupakan suatu sistem yang dibuat oleh manusia yang terdiri dari komponen-komponen dalam organisasi untuk mencapai suatu tujuan yaitu menyajikan informasi. Berikut akan dijelaskan mengenai komponen dan kegiatan yang ada dalam sistem informasi[8].

Komponen sistem informasi terdiri dari:

- a) *Hardware*, terdiri dari computer, peripheral dan jaringan.
- b) *Software*, merupakan kumpulan dari perintah/fungsi yang ditulis dengan aturan tertentu untuk memerintahkan komputer melaksanakan tugas tertentu. Contoh : sistem operasi dan utilitas.
- c) *Data*, merupakan komponen dasar dari informasi yang akan diproses lebih lanjut untuk menghasilkan informasi.
- d) Manusia, yang terlibat dalam komponen manusia seperti operator dan pemimpin sistem informasi.
- e) Prosedur, seperti dokumentasi prosedur/proses sistem, buku penuntun operasional dan teknis.

Kegiatan di sistem informasi mencakup :

- a) *Input*, menggambarkan suatu kegiatan untuk menyediakan data untuk diproses.
- b) Proses, menggambarkan bagaimana suatu data di proses untuk menghasilkan suatu informasi yang bernilai tambah.
- c) *Output*, suatu kegiatan untuk menghasilkan laporan dari proses dan input.
- d) Penyimpanan, suatu kegiatan untuk memelihara dan menyimpan data.
- e) *Control*, suatu aktivitas untuk menjamin bahwa sistem informasi tersebut berjalan sesuai dengan yang diharapkan.

Kemudian dari sistem informasi ini berkembang menjadi beberapa sistem pengembangan. Salah satunya Sistem Informasi Eksekutif. Sistem Informasi Eksekutif ini memberikan akses yang ditingkatkan ke suatu informasi dalam berbagai format[3]. Sistem Informasi Eksekutif merupakan sistem yang bertujuan dapat membantu para eksekutif atau bisa dikatakan para *top manager levels*, sehingga dapat mengetahui bagaimana hasil dari pengolahan informasi secara maksimal demi kebutuhan para eksekutif.

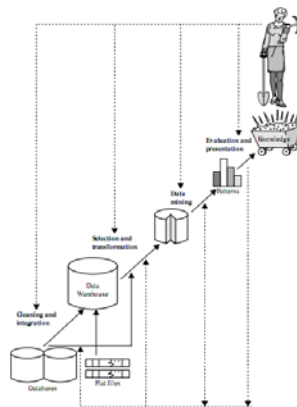
Pemrosesan yang ada dalam sistem informasi eksekutif ini berhubungan dengan transaksi perusahaan atau proses yang lebih bersifat manajerial. Berdasar teori fungsi manajemen menurut Fayol bahwa semua manajer melakukan fungsi manajemen yang sama, yakni perencanaan, pengorganisasian, pengaturan staff, pengarahan dan pengontrolan. Sedangkan teori peranan manajerial menurut Mintzberg bahwa semua manajer menjalankan semua peranan, namun orientasinya berbeda-beda menurut tingkatannya[9]. Dari dua teori tadi dapat dilihat bahwa peranan dari Sistem Informasi Eksekutif akan membantu proses tadi

disesuaikan dengan apa tujuan utama dari pembuatan atau penggunaan Sistem Informasi Eksekutif ini sesuai dengan kebutuhan.

### Data Mining

*Data Mining* merupakan suatu proses menemukan hubungan yang berarti kecenderungan dengan memeriksa sekumpulan besar data yang tersimpan dalam penyimpanan dengan teknik pengenalan pola seperti teknik statistik dan matematika[10]. Ketersediaan data yang melimpah, kebutuhan akan informasi atau pengetahuan sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis, dan dukungan infrastruktur di bidang teknologi informasi merupakan pendukung dari pemrosesan *data mining* ini. Berbagai penerapan *data mining* yakni Analisa pasar dan manajemen, Analisa perusahaan dan manajemen resiko, Telekomunikasi, Keuangan, Asuransi, Olahraga dan Astronomi.

Sebagai suatu teknologi yang dapat menghasilkan *knowledge*, *data mining* melakukan pemrosesan dalam beberapa tahap. Berikut tahapan-tahapan yang di lakukan dalam proses *data mining*.



Gb1. Tahap-tahap Data Mining[11]

- 1) *Data Cleaning*  
Proses *Data Cleaning* ini merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data yang tidak relevan. Data-data yang harus dibersihkan seperti data yang memiliki isian yang tidak sempurna seperti data yang hilang atau data yang tidak valid. Pembersihan data ini juga akan mempengaruhi hasil dari proses *data mining* yang dihasilkan karena data yang akan ditangani berkurang secara kuantitas dan kompleksitas pemrosesannya.
- 2) *Data Integration*

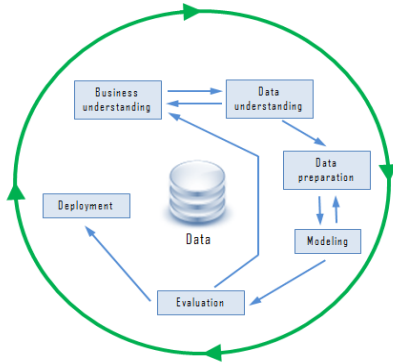
Proses ini merupakan proses penggabungan data kedalam satu *database* dimana terkadang proses *data mining* yang akan dilakukan memerlukan data lebih dari satu *database* atau memerlukan pemrosesan menggunakan *database* lain. Integrasi data perlu dilakukan secara cermat, karena dapat menghasilkan hasil yang menyimpang jika terjadi kesalahan pada proses ini.

- 3) *Data Selection*  
Proses penyeleksian data yang dilakukan karena tidak semua data yang ada di dalam *database* akan digunakan, hanya data yang sesuai untuk di analisis yang akan digunakan.
- 4) *Data Transformation*  
*Data Transformation* merupakan pengubahan data atau penggabungan data yang dilakukan sebelum akan dilakukan proses *data mining*. Karena dalam proses *data mining* dalam beberapa metode membutuhkan format data yang khusus sebelum data diaplikasikan.
- 5) Proses *Data Mining*  
Proses utama yang dilakukan berdasarkan metode yang sudah dipilih untuk menghasilkan suatu pengetahuan.
- 6) *Pattern Evaluation*  
Pada proses ini diidentifikasi pola-pola yang ada kedalam *knowledge based* dimana hasil dari teknik *data mining* dievaluasi untuk dilakukan proses penilaian apakah pola hipotesa yang ada benar-benar tercapai, bila hasil pemrosesan tidak sesuai maka dapat dilakukan perbaikan pada proses *data mining*, mencoba metode pemrosesan yang lain atau menerima hasil ini sebagai bahan acuan untuk penelitian *data mining* dimasa akan datang.
- 7) *Knowledge Persentation*  
Merupakan penyajian hasil dari pemrosesan yang ada. Pada tahap ini terfokus pada bagaimana penyajian atau *visualisasi* dari hasil pengetahuan yang sudah didapat sehingga dapat dipahami oleh pengguna atau yang memanfaatkan hasil dari pemrosesan tersebut.

### CRISP-DM

*Cross-Industry Standard Process for Data Mining (CRISP-DM)* merupakan suatu standar yang telah dikembangkan pada tahun 1996 yang ditujukan untuk melakukan proses analisis dari suatu industri sebagai strategi pemecahan masalah dari bisnis atau unit penelitian[5]. Untuk data yang dapat diproses dengan *CRISP-DM* ini, tidak ada ketentuan atau karakteristik tertentu, karena data tersebut akan diproses kembali pada fase-fase di

daalamnya. Terdapat enam fase dalam *CRISP-DM* ini yakni dijelaskan pada gambar berikut[13] :



Gb. 2 Fase *CRISP-DM* [14]

- a) Fase Pemahaman Bisnis (*Business Understanding Phase*)
  - 1) Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.
  - 2) Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan data mining.
  - 3) Menyiapkan strategi awal untuk mencapai tujuan.
- b) Fase Pemahaman Data (*Data Understanding Phase*)
  - 1) Mengumpulkan data, jika data berasal dari lebih dari satu *database* maka dilakukan proses integrasi data atau *Data Integration*.
  - 2) Mengembangkan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
  - 3) Mengevaluasi kualitas data, memeriksa data dan membersihkan data yang tidak *valid* atau proses *Data Cleaning*.
  - 4) Jika diinginkan, pilih sebagian kecil grup data yang mungkin mengandung pola dari permasalahan.
- c) Fase Pengolahan Data (*Data Preparation Phase*)
  - 1) Siapkan data awal, kumpulan data yang akan digunakan untuk keseluruhan fase berikutnya atau proses *Data Selection*.
  - 2) Pilih kasus dan variabel yang akan dianalisis, sesuai dengan analisis yang akan dilakukan.
  - 3) Lakukan perubahan pada variabel jika diperlukan.
  - 4) Siapkan data awal sehingga siap untuk perangkat permodelan atau *Data Transformation*

- d) Fase Permodelan (*Modelling Phase*)
  - 1) Pilih dan aplikasikan teknik permodelan yang sesuai.
  - 2) Kalibrasi aturan model untuk mengoptimalkan hasil.
  - 3) Dapat menggunakan beberapa teknik yang sama untuk permasalahan yang sama
  - 4) Dapat kembali ke fase pengolahan data jika diperlukan untuk menjadikan data ke dalam bentuk kebutuhan tertentu
- e) Fase Evaluasi (*Evaluation Phase*)
  - 1) Mengevaluasi satu atau lebih model yang digunakan dalam fase permodelan atau proses *Evaluation Pattern*.
  - 2) Menetapkan apakah model tadi sudah sesuai dengan tujuan pada fase awal.
  - 3) Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
  - 4) Mengambil keputusan berkaitan dengan penggunaan hasil dari data mining.
- f) Fase Penyebaran (*Deployment Phase*)  
Menggunakan model yang dihasilkan dan di persentasikan atau proses *knowledge persentation*.

#### Algoritma Neirest Neighbour

*Neirest Neighbour* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada[15]. Ini digunakan dalam pengklasifikasian suatu kasus yang terjadi berdasar kasus yang sudah terjadi atau karakteristik suatu kasus yang bias dijadikan persamaan dengan kasus lainnya yang belum diidentifikasi atau diklasifikasikan lebih lanjut.

Misal dalam sebuah kasus mencari solusi terhadap seorang pasien baru dengan menggunakan solusi dari pasien terdahulu. Maka dihitung kedekatan kasus pasien baru dengan semua kasus pasien lama. Kasus pasien lama dengan kedekatan terbesar akan diambil untuk digunakan pada kasus pasien baru.

Proses perhitungan nilai kedekatan tersebut dengan menggunakan rumus :

$$stmlarity(T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) x w_i}{w_i}$$

Keterangan :

T : Kasus baru

S : Kamus yang ada dalam penyimpanan

n : Jumlah atribut dalam setiap kasus

i : Atribut individu antara 1 sampai dengan n

f : Fungsi similarity atribut 1 sampai dengan n  
w : Bobot yang diberikan pada atribut ke-i

Dimana setiap kasus lama dihitung tiap kedekatannya dengan kasus baru, dan nantinya nilai terbesar dari setiap perhitungan tersebut akan diambil sebagai kesimpulan dari kasus baru yang telah dihitung kedekatannya. Kedekatan biasanya berada pada nilai 0 sampai dengan 1. Nilai 0 artinya kedua kasus mutlak tidak mirip, dan nilai 1 kasus mutlak mirip. Misalnya untuk memprediksi apakah nasabah bank yang baru bermasalah atau tidak berdasarkan data-data yang dimiliki.

### Implementasi *CRISP-DM*

Implementasi *CRISP-DM* terdiri dari 6 fase yakni fase pemahaman bisnis, pemahaman data, pengolahan data, permodelan, evaluasi dan penyebaran.

#### i. Pemahaman Bisnis

Seperti yang sudah dijelaskan bahwa pengimplementasian *data mining* pada aplikasi ini adalah berhubungan dengan data produksi ikan. Faktor-faktor apa saja yang mempengaruhi produksi ikan sesuai data yang sudah ada. Memeriksa keterkaitan antara tiap atribut yang mempengaruhi produksi ikan yang nantinya dapat dijadikan acuan untuk memprediksi hasil produksi ikan selanjutnya sehingga memaksimalkan pemahaman data yang ada.

Tujuan dari pengimplementasian *data mining* pada aplikasi ini adalah agar didapatkan suatu pengetahuan tambahan mengenai produksi ikan sehingga dapat menjadi tolak ukur mengenai potensi produksi ikan pada data yang ada sebagai *review* bagi pengguna sistem dengan melihat faktor – faktor yang mempengaruhi produksi tersebut.

Batasan-batasan dalam implementasi *data mining* pada aplikasi ini adalah diantaranya sebagai berikut :

- a. Tidak ada faktor luar diluar sumber data yang ikut menjadi atribut yang mempengaruhi proses implementasi *data mining* ini. Seperti cuaca, musim, angin, dan lain-lain.
- b. Jumlah komoditi ikan yang cukup banyak dan tidak semuanya pada komoditi tersebut memiliki hasil produksi di ruang lingkup perairan di Provinsi Jawa Tengah, sehingga akan dipilih beberapa komoditi unggulan agar memaksimalkan proses peminingan data sehingga pemrosesan lebih terfokus.

Dengan batasan diatas diharapkan dapat memaksimalkan pemrosesan data yang ada

sehingga bisa mencapai mengenai potensi produksi perikanan budidaya.

#### ii. Pemahaman Data

Pada tahap ini pengumpulan data terlebih dahulu dilakukan, semua data yang diperlukan nanti akan dilakukan pengolahan secara menyeluruh. Pengumpulan data berupa dengan mengambil data yang ada lewat buku statistik dan data yang ada di Dinas Kelautan dan Perikanan Provinsi Jawa Tengah.

Data akan diintegrasikan secara langsung, tanpa harus adanya proses *data warehouse* seperti pada tahapan *data mining* pada gambar 1 dan tanpa adanya migrasi data terlebih dahulu. Sistem ini tidak melakukan pemrosesan pemeriksaan data yang diinput pada SIPBB. Dalam SIPBB terdapat suatu standar penginputan dimana kesalahan proses *input* dari tiap kota/kabupaten mengenai data produksi perikanan budidaya telah dilakukan pengecekan melalui proses verifikasi. Format hasil proses *input* sudah sesuai dengan prosedur pemrosesan data yang akan dilakukan pada sistem ini. Analisis data yang akan dilakukan mengenai penyelidikan data produksi beserta faktor-faktor yang mempengaruhi akan hasil produksi, bagaimana faktor-faktor seperti jenis komoditi, kota/kabupaten, dan jenis budidaya yang nantinya berkaitan dengan produksi ikan. Bagaimana faktor-faktor tadi dapat dijadikan atribut acuan dalam mendapatkan informasi atau pengetahuan sesuai dengan pada tahap pemahaman bisnis sebelumnya.

Kemudian dilakukan pengevaluasian dari kualitas data yang ada, data yang banyak dan tidak beraturan dan di dalamnya bisa terjadi adanya kesalahan sehingga perlu adanya pemeriksaan dan memilah bagian atribut yang akan digunakan dalam pencarian pengetahuan, sehingga dapat dihasilkan pelaporan yang nantinya akan benar-benar digunakan sebagai berikut :

- a. Bagaimana pelaporan berdasar pada hubungan antara kota/kabupaten, komoditi, dan hasil produksi
- b. Bagaimana pelaporan berdasar pada hubungan antara komoditi, jenis budidaya dan hasil produksi
- c. Bagaimana pelaporan berdasar pada hubungan antara kota/kabupaten, komoditi, dan nilai produksi

d. Bagaimana pelaporan berdasar pada hubungan antara komoditi, jenis budidaya dan nilai produksi

iii. Pengolahan Data

Pada fase ini akan dilakukan proses pemilihan dan pengolahan data yang nanti akan diperlukan dalam tahap permodelan sehingga nanti permodelan yang dilakukan dapat memberikan hasil maksimal sesuai dengan target yang diinginkan.

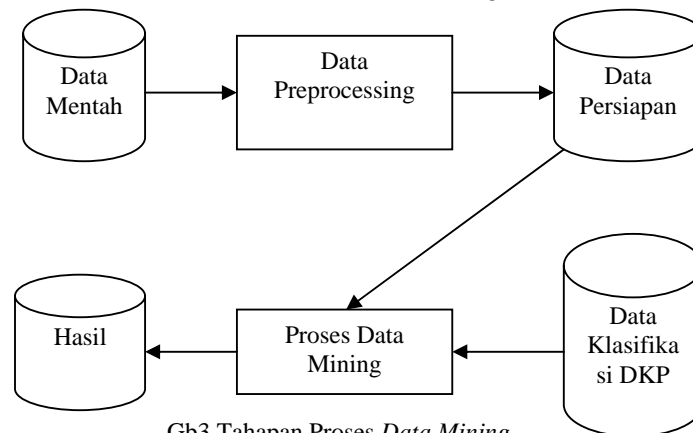
Dengan menggunakan tools *weka* untuk melakukan proses komparasi data, bagaimana data dapat diolah, dan dianalisa sehingga dapat ditentukan suatu permodelan yang sesuai dengan data sehingga dapat menghasilkan suatu hasil yang diinginkan mengenai potensi produksi yang sesuai dengan kondisi data yang ada.

Data yang akan dipilah adalah sebagai berikut :

1. Produksi dan Nilai berdasar jenis budidaya perkuartal dan pertahun
2. Produksi dan Nilai berdasar komoditi perkuartal dan pertahun
3. Produksi dan Nilai berdasar jenis budidaya dan kota/kabupaten perkuartal dan pertahun.

iv. Permodelan

Pemilihan Data merupakan dalam fase Pengolahan data dimana menyesuaikan dengan proses mining yang akan dilakukan. Proses pengklasifikasian yang dilakukan yakni melihat tinggi rendahnya hasil produksi setiap kuartal pada setiap kota/kabupaten berdasarkan jenis budidaya yang ada. Sebagai pembanding dari kasus pada data sebenarnya, dibuat suatu klasifikasi data yang di dasarkan pada data klasifikasi yang ada di Dinas Kelautan dan Perikanan Jateng mengenai potensi produksi perikanan budidaya, berikut tahapan pengolahan data hingga implementasi proses mining :



Gb3 Tahapan Proses Data Mining

Teknik permodelan data yang dilakukan antara lain :

1. Teknik Peramalan (*forecasting*)

Pada kasus data produksi ikan, peramalan dapat dilakukan untuk tujuan bagaimana memperkirakan produksi ikan di masa depan dengan melihat pola data yang ada yang berbasis *time series*. Data produksi perikanan budidaya terdapat dari tahun 2006 hingga 2011. Data selama 6 tahun tersebut di masukkan setiap kuartal, sehingga terdapat 4 data kolektif setiap tahun perkomoditi, perkota/kabupaten dan sesuai dengan jenis pembudidayaannya.

2. Klasifikasi

Proses klasifikasi dilakukan berdasar kriteria tertentu sehingga proses pengklasifikasian dapat dilakukan. Pada data produksi perikanan

budidaya dapat dilakukan proses klasifikasi yang disesuaikan dengan jenis budidaya, daerah dan waktu. Namun untuk implementasinya diperlukan pemilihan algoritma yang tepat agar tidak memberatkan sistem yang dibangun

a. Data Klasifikasi

Berdasarkan data klasifikasi DKP, proses klasifikasi yang akan dilakukan yakni menggunakan *weka* dengan menggunakan 2 algoritma yakni *Decision Trees* dan *Nearest Neighbour*. Menggunakan *Decision Trees* dikarenakan data potensi yang ada dapat dilakukan pendekatan sehingga proses klasifikasi data dapat terlihat pengklasifikasian berdasar potensi pada data klasifikasi dan untuk penggunaan *Nearest Neighbour* dicari kedekatan atau *similarity* agar dapat

diklasifikasikan potensi perikanan dari hasil produksi perikanan budidaya yang ada pada data LB3 dengan data klasifikasi.

b. *Preprocessing*

Data LB3 tidak dapat diproses secara langsung pada permodelan ini. Berdasar pada data klasifikasi, maka data LB3 akan dilakukan proses *preprocessing* agar permodelan dapat dilakukan. Pada proses ini data LB3 yang ada akan dilakukan proses perhitungan agar didapatkan data produksi dalam bentuk skala tinggi, sedang dan rendah.

Proses perhitungan skala tadi, didapatkan berdasar dari ketentuan yang ada di Dinas Kelautan dan Perikanan di Provinsi Jawa Tengah. Proses perhitungannya adalah sebagai berikut.

Produksi Suatu Kota/Kabupaten pada suatu budidaya dikategorikan <b>Tinggi</b> jika $Produksi > \frac{3}{2}R$
Produksi Suatu Kota/Kabupaten pada suatu budidaya dikategorikan <b>Sedang</b> jika $\frac{1}{2}R \leq Produksi \leq \frac{3}{2}R$
Produksi Suatu Kota/Kabupaten pada suatu budidaya dikategorikan <b>Rendah</b> jika $Produksi < \frac{1}{2}R$

Keterangan.

R : Rata-rata produksi suatu budidaya tahunan di Provinsi Jawa Tengah

Setelah proses tadi, data hasil *preprocessing* disimpan dalam data persiapan.

Data persiapan tersebut sudah siap diproses oleh *weka* untuk dilakukan proses perbandingan antara hasil menggunakan *Decision Tree* dengan *Neirest Neighbour*.

c. Hasil Perbandingan

Data hasil *preprocessing* tadi diproses oleh *weka* dan akan dilakukan perbandingan dan hasilnya algoritma *Neirest Neighbour* lebih cepat dalam melakukan pemrosesan sehingga permodelan yang akan diimplementasikan adalah menggunakan algoritma tersebut.

Data Hasil merupakan hasil proses *data mining* dari Data Persiapan dibandingkan dengan Data Klasifikasi dari Dinas Kelautan dan Perikanan agar dicari di tiap data persiapan dikategorikan ke potensi manakah dengan menghitung nilai kedekatan yang paling dekat dengan data klasifikasi yang ada.

Hasil akhir yang didapat berupa *review* potensi hasil produksi perikanan secara menyeluruh pada setiap jenis budidaya di tiap daerah didasarkan pada tinggi rendahnya hasil produksi yang dilakukan pada fase pengolahan data sebelumnya.

Dengan pengimplementasian algoritma *Neirest Neighbour* dibutuhkan sebuah data training. Data training tersebut diambil dari seluruh data klasifikasi DKP sebagai syarat pengklasifikasian potensi perikanan budidaya di Jawa Tengah. Kemudian penentuan bobot pada tiap atribut dan nilai kedekatan di setiap atribut yang akan digunakan untuk menghitung nilai kedekatan yang ada di masing-masing kasus dan data klasifikasi yang ada.

Besaran nilai bobot yang ada ditentukan berdasar pada asumsi bahwa tidak ada atribut yang lebih dominan daripada atribut lain, sehingga diberi nilai sama untuk semua atribut, yakni 1. Berikut tabel nilai bobot tiap atribut pada tabel 3.9 dan asumsi kedekatan.

Tabel 3.11 Tabel Nilai Bobot Tiap Atribut

Atribut	Bobot
Tambak	1
Kolam	1
Karamba	1
Karambaja	1
Sawah	1
Laut	1

Tabel 3.14 Asumsi Nilai Kedekatan

Atribut	Atribut	Kedekat
Tinggi	Rendah	0.3
Rendah	Tinggi	0.3
Sedang	Rendah	0.6
rendah	Sedang	0.6
Tinggi	Sedang	0.5
Sedang	Tinggi	0.5
Tinggi	Tinggi	1
Sedang	Sedang	1
Rendah	Rendah	1

v. Evaluasi

Dilakukan 2 proses evaluasi, yakni pengujian dan perbandingan. Dengan pengujian diambil 30 data secara acak dinilai secara manual dan dibandingkan dengan hasil permodelan yang dilakukan. Kemudian dengan perbandingan dengan menggunakan Fungsi kondisional kemudian dibandingkan dengan permodelan yang sudah dimodelkan tadi. Hasil dari

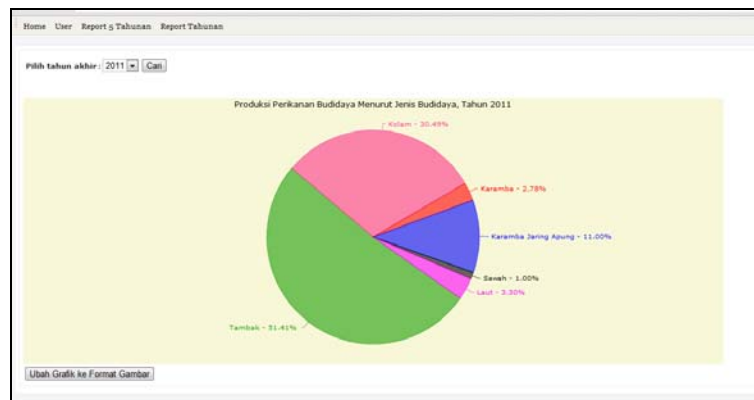
evaluasi ini, permodelan yang dilakukan telah sesuai dengan tujuan pengimplementasian *data mining* pada sistem ini.

vi. Penyebaran

Setelah tahap evaluasi dimana menilai secara detail hasil dari permodelan dilakukan pengimplementasian dari keseluruhan model yang telah dirancang. Selain itu juga dilakukan penyesuaian dari model dengan sistem yang akan dibangun sehingga dapat menghasilkan suatu hasil yang sesuai dengan target diawal tahap *CRISP-DM* ini

### Implementasi Sistem

Pada Pelaporan grafik, terdapat beberapa report 5 tahunan dan tahunan produksi perikanan budidaya. Pada proses peminingan, terdapat 3 form utama yakni form penyiapan data, hasil, dan *report mining*. Pada form penyiapan data dan hasil dapat dilihat data pada tahun mana saja yang sudah diproses dalam penyiapan data dan hasil tersebut. Report yang dihasilkan berupa grafik bar dengan spesifikasi potensi perikanan budidaya setiap tahunnya



Gambar 4 Report Produksi Tahunan



Gambar 5 Report 5 Tahunan



Home User Report 5 Tahunan Report Tahunan Logout

## Proses Data Mining

Data Persiapan Data Hasil Report Potensi

Pemrosesan Data Persiapan (Tahun) :

### Data Persiapan

Data Persiapan yang sudah siap

Tahun	Action
2006	<a href="#">Delete</a>
2007	<a href="#">Delete</a>
2008	<a href="#">Delete</a>
2009	<a href="#">Delete</a>
2010	<a href="#">Delete</a>
2011	<a href="#">Delete</a>

Gambar 6 Form Penyiapan Data

Home User Report 5 Tahunan Report Tahunan Logout

## Proses Data Mining

Data Persiapan Data Hasil Report Potensi

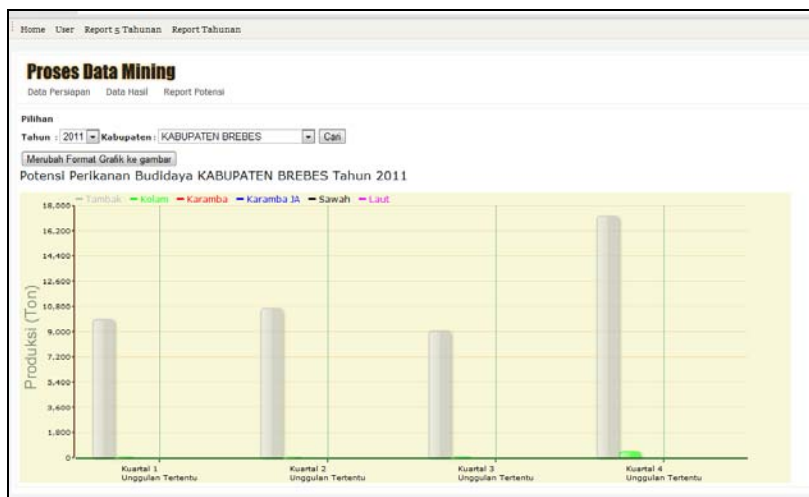
Pemrosesan Data Hasil (Tahun) :

### Data Hasil

Data Hasil Yang sudah diminingkan

Tahun	Action
2006	<a href="#">Delete</a>
2007	<a href="#">Delete</a>
2008	<a href="#">Delete</a>
2009	<a href="#">Delete</a>
2010	<a href="#">Delete</a>
2011	<a href="#">Delete</a>

Gambar 7 Form Hasil



Gambar 8 Report Potensi

## Kesimpulan dan Saran

Kesimpulan yang dapat diambil dalam pengerjaan penelitian ini adalah sebagai berikut :

- 1) Sistem Informasi Eksekutif yang dikembangkan menyajikan report grafik hasil produksi dan nilai produksi 5 tahunan perikanan budidaya serta persentasi produksi untuk setiap jenis budidaya di wilayah Provinsi Jawa Tengah untuk keperluan intern DKP Jateng.
- 2) Implementasi *data mining* pada Sistem ini dengan memanfaatkan data produksi perikanan budidaya yang ada sehingga dihasilkan suatu penilaian potensi perikanan budidaya pada setiap kuartal dari tahun terkait.

Saran untuk pengembangan SIE DKP Jateng ialah penggalian pengetahuan dari hasil produksi perikanan budidaya di Provinsi Jawa Tengah yakni dengan melihat pola produksi setiap jenis budidaya perikanan pada tiap kota/kabupaten di Provinsi Jawa Tengah.

## Referensi

- [1] \_\_\_\_\_,2012, “Dinas Kelautan dan Perikanan Provinsi Jawa Tengah” diakses dari <http://www.diskanlut-jateng.go.id/> diakses pada 1 januari 2012.
- [2] Purnama, Indra. 2012. *Laporan Praktik Kerja Lapangan : Pengembangan Sistem Informasi Pengolahan Statistik Perikanan Budidaya Dinas Kelautan dan Perikanan Provinsi Jawa Tengah*. Semarang: Fakultas MIPA Universitas Diponegoro.
- [3] Turban, E. Dkk. 2005. *Decision Support System and Intelligent Systems*. Yogyakarta : Andi.
- [4] \_\_\_\_\_,2012, “Learning Community”, diakses dari <http://andyku.wordpress.com/2008/04/17/pen-erapan-data-mining/> pada 25 Februari 2012 pukul 16.00 WIB.
- [5] Chapman, Peter. 2000. *CRISP-DM 1.0 Step by step Data Mining Guide*. SPSS Inc, diakses dari [http://www.spss.ch/upload/1107356429\\_CrispDM1.0.pdf](http://www.spss.ch/upload/1107356429_CrispDM1.0.pdf) pada 8 Januari 2012, pukul 18.05 WIB.
- [6] Da Rocha, Bruno Carneiro. 2010. *Identifying Bank Frauds Using CRISP-DM and Decision Trees*. International Journal. Brazillia : Bank of Brazil <http://www.scribd.com/doc/40397894/Identifying-Bank-Frauds-Using-CRISP-DM-and-Decision-Trees> diakses pada 25 Februari 2011 pukul 16.00 WIB
- [7] Joseph R. Erskine, Captain, USAF. 2010. *Developing Cyberspace Data Understanding Using CRISP-DM for Host IDS Feature Mining (Thesis)*. Ohio : Air Force Institute of Technology.
- [8] Leman. 1998. *Metodologi Pengembangan Sistem Informasi*. Jakarta : Gramedia.
- [9] \_\_\_\_\_,2012, “Materi Sistem Informasi Eksekutif.” ITTelkom, diakses dari <http://www.ittelkom.ac.id/staf/mhd/MateriKuliah/SPPK/Bab-207-20Executive-20Information-20System.pdf> pada 21 Juni 2012, pukul 13.00 WIB
- [10] Larose, Daniel T. 2005. *Discovering knowledge in Data : An Introduction to Data Mining*. John Willey dan Sons Inc.
- [11] Han, J. and Kamber, M. 2011. *Data Mining Concepts and Techniques Third Edition*. Elsevier.