

SPASIAL DATA MINING MENGGUNAKAN MODEL SAR-KRIGING

A. Setiawan Abdullah dan R. Rosadi

Abstract. *The region of Indonesia is very sparse and it has a variation condition in social, economic and culture, so the problem in education quality at many locations is an interesting topic to be studied. Database used in this research is Base Survey of National Education 2003, while a spatial data is presented by district coordinate as a least analysis unit. The aim of this research is to study and to apply spatial data mining to predict education quality at elementary and junior high schools using SAR-Kriging method which combines an expansion SAR and Kriging method. Spatial data mining process has three stages. preprocessing, process of data mining, and post processing.*

For processing data and checking model, we built software application of Spatial Data Mining using SAR-Kriging method. An application is used to predict education quality at unsample locations at some cities at DIY Province. The result shows that SAR-Kriging method for some cities at DIY for elementary school has an average percentage error 6.43%. We can conclude that for elementary school, SAR-Kriging method can be used as a fitted model.

Keywords— Expansion SAR, SAR-Kriging, quality education

1. PENDAHULUAN

Spasial data mining merupakan penambangan pengetahuan dari data spasial dalam jumlah besar, merupakan bidang yang menarik untuk diteliti. Pengetahuan yang ditemukan dalam data spasial mempunyai berbagai bentuk, dapat didasarkan pada aturan pengelompokan atau diskriminan, ekstraksi dan deskripsi dari struktur cluster, asosiasi spasial, dan lain-lain. Metode spasial data mining dapat diterapkan dalam pemetaan mutu pendidikan, mengingat Indonesia memiliki sebaran lokasi yang luas, terdiri dari: provinsi, kabupaten/kota, kecamatan, dan desa.

Atje Setiawan, staf pengajar Jurusan Matematika, FMIPA Unpad, Jl. Raya Bandung-Sumedang Km. 21 Jatinangor (e-mail: atje@bdg.centrin.net.id).

Rudi Rosadi, staf pengajar Jurusan Matematika, FMIPA Unpad, Jl. Raya Bandung-Sumedang Km. 21 Jatinangor (e-mail: rudirosadi@gmail.com).

Pembangunan pendidikan nasional merujuk pada empat aspek sebagai tema pokok, yaitu: pemerataan kesempatan belajar, peningkatan mutu, peningkatan efisiensi pelaksanaan pendidikan, dan relevansi pendidikan dengan pembangunan. Salah satu upaya yang dilakukan pemerintah dalam mendukung pembangunan pendidikan nasional melalui Balitbang

Diknas dalam membangun database pendidikan adalah melakukan Survei Dasar Pendidikan Nasional tahun 2003 (SDPN2003). Hasil SDPN 2003, khususnya untuk persekolahan secara nasional terjaring 203.590 record dengan 569 indikator, meliputi SD 158.590 record dengan 122 indikator, SMP 28.949 record dengan 138 indikator, SMA 10.810 record dengan 142 indikator, serta SMK 5.156 record dengan 167 indikator [3]. Pada survei tersebut, tidak semua provinsi tersampel, dan terjadinya pemekaran wilayah berupa provinsi/ kabupaten/kota baru, memerlukan suatu metode untuk prediksi mutu pendidikan di lokasi yang tidak tersampel tersebut.

Penelitian ini bertujuan untuk pengembangan dan penerapan spasial data mining untuk prediksi mutu pendidikan dengan aplikasi Spatial Autoregressive-Kriging (SAR-Kriging), meliputi analisis, disain dan implementasi sistem. Model SAR-Kriging merupakan penggabungan model ekspansi SAR dengan metode Kriging, digunakan untuk prediksi mutu pendidikan pada lokasi yang tidak tersampel pada berbagai jenjang persekolahan di kabupaten/kota di DIY. Dengan demikian metode SAR-Kriging dapat melengkapi model ekspansi SAR untuk prediksi mutu pendidikan di berbagai lokasi yang tidak tersampel [1].

2. METODE PENELITIAN

2.1 Data Mining

Sejalan dengan pertumbuhan teknologi informasi, Data mining merupakan salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database skala

besar yang makin banyak terakumulasi. Beberapa definisi umum data mining adalah sebagai berikut :

Data mining merupakan serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data [5].

Proses KDD terdiri dari 3 tahap: data *preprocessing*, data mining dan *postprocessing*. Beberapa tantangan spesifik mendorong pengembangan data mining adalah *scalability, high dimensionality, heterogeneous and complex data, Data ownership and distribution, non-traditional analysis* [11]. Proses data mining dan *Knowledge Discovery in Databases (KDD)* secara garis besar dijelaskan sebagai berikut: *data selection, pre-processing/ cleaning, transformation, data mining, interpretation/ evaluation*. Tugas data mining secara umum terdiri dari dua kategori prediksi (*predictive tasks*) dan deskripsi (*descriptive tasks*) [5].

Proses data mining terdiri dari tujuh langkah, yaitu: (a) definisi tujuan analisis, (b) seleksi dan organisasi data, (c) analisis eksplorasi dan transformasi data, (d) spesifikasi metode statistika yang akan digunakan tahap analisis, (e) analisis data berdasarkan metode yang dipilih, (f) evaluasi dan perbandingan metode, dan (g) interpretasi model untuk pengambilan keputusan [4].

Metodologi data mining secara ringkas dapat dibagi menjadi beberapa tahap: *data cleaning, integrasi data, transformasi data, aplikasi teknik data mining, evaluasi pola yang ditemukan, dan presentasi pengetahuan* [5].

2.2 Spasial Data Mining

Database spasial menyimpan objek-objek spasial yang direpresentasikan oleh tipe data spasial dan hubungan spasial di antara objek-objek [5]. Data spasial membawa topologi atau informasi jarak dan seringkali disusun oleh struktur indeks spasial yang dapat diakses oleh metode-metode spasial. Pengembangan database spasial seperti struktur data spasial dan komputasional geometri, memberikan jalan untuk mengkaji spasial data mining. Spasial data mining adalah suatu proses mengekstrak pengetahuan yang diinginkan dari relasi spasial atau berbagai pola yang secara eksplisit tidak ditemukan dalam database [8].

Metode spasial data mining dapat digunakan untuk mengekstrak pengetahuan regular dan menarik dari database spasial yang besar [6].

Masing-masing bagian, dapat digunakan untuk memahami data spasial, menemukan hubungan antara data spasial dan non spasial, membangun base pengetahuan spasial, optimisasi query, reorganisasi database spasial, menjelaskan karakteristik umum melalui cara yang sederhana, dan sebagainya.

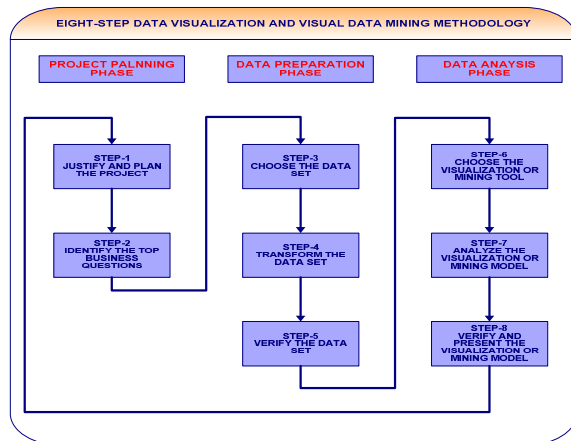
Salah satu teknik klasifikasi adalah pengembangan teknik penambangan data klasik untuk menggabungkan autokorelasi spasial, yang menjadi kunci untuk membedakan properti data spasial. Dengan menggunakan regresi linier sebagai *prototype*, metode-metode klasifikasi dapat dikembangkan untuk model autokorelasi spasial. Salah satu teknik regresi yang menjelaskan autokorelasi spasial paling sederhana adalah dengan memodifikasi model regresi melalui bantuan matriks bobot spasial W berupa model SAR [5].

Pengolahan data berupa penaksiran parameter model SAR serta output lainnya dapat dilakukan secara komputerisasi, misalnya menggunakan perangkat lunak MATLAB. MATLAB merupakan suatu perangkat lunak yang terkenal untuk kalkulasi data dan representasi grafik. Struktur dalam MATLAB memungkinkan user memanggil fungsi-fungsi yang berkaitan dengan analisis data spasial, melakukan komputasi dengan matriks sparse (unsur-unsur bernilai nol yang banyak) seperti matriks bobot spasial, penaksiran parameter model spasial dan sebagainya. Desain dan implementasi spasial data mining dapat menggunakan MATLAB untuk disain fungsi *brain image* dalam pemetaan parameter-parameter secara statistika (*Statistical Parameters Mapping/SPM*). SPM adalah salah satu contoh keberhasilan MATLAB dalam pengolahan data dan penggunaan GUI (*Graphical User Interface*). Kombinasi MATLAB, GIS dan spasial data mining membangun perangkat lunak kerangka sistem spasial data mining dalam platform MATLAB yang diintegrasikan dalam algoritma-algoritma seperti: aturan asosiasi spasial, analisis pengelompokan spasial, analisis keputusan pohon (*decision tree*) dan penggunaan sistem untuk database spasial dalam penggunaan daratan, pemrosesan data vektor spasial dan berbagai aspek lainnya [7].

2.3 Proses Data Mining

Tiga fase metodologi visualisasi data mining, terdiri dari, fase perencanaan proyek, fase penyiapan data dan fase analisis data. Ketiga fase

dirinci menjadi delapan langkah metodologi visualisasi data mining, meliputi: penilaian dan perencanaan proyek, identifikasi permasalahan bisnis, pemilihan kumpulan data, transformasi kumpulan data, verifikasi kumpulan data, pemilihan alat visualisasi atau alat penambangan, analisis model visualisasi dan model penambangan, serta verifikasi dan presentasi visualisasi dan model penambangan [10] seperti disajikan pada Gambar 2.1.



Gambar 2.1. Delapan Langkah Metodologi Visualisasi Data Mining

2.4 Model Input Output Pendidikan

2.4.1 Pengertian Mutu Pendidikan

Pendidikan pada tingkat persekolahan adalah proses pembelajaran melalui sekolah yang dikenakan kepada siswa, agar mempunyai tingkat mutu/kompetensi tertentu dalam kemampuan kognitif, psikomotorik, dan afektif, sesuai dengan yang ditetapkan dalam kurikulum pendidikan dasar dan menengah. Terdapat 2 aspek, “proses pembelajaran melalui sekolah” dan “tingkat mutu/kompetensi tertentu”. Mutu pendidikan didefinisikan sebagai prestasi yang dicapai oleh siswa dan diukur berdasarkan nilai UAN, merupakan prestasi tertinggi siswa setelah mengikuti proses pembelajaran yang diselenggarakan sekolah [9]. Analisis mutu pendidikan dalam penelitian ini mencakup juga mutu dari proses pembelajaran dengan berbagai inputnya.

2.4.2 Variabel yang Digunakan dalam Analisis

Variabel yang digunakan, yaitu *variabel dasar* dan *variabel indikator*. Variabel dasar adalah semua variabel yang terdapat didalam “raw data individual sekolah”. Sedangkan variabel indikator adalah variabel yang diperoleh berdasarkan variabel -variabel dasar [1]. Variabel dasar dan variabel indikator yang digunakan dalam analisis berhubungan dengan; identitas

sekolah, siswa, guru, sarana/prasarana dan UAN. Dari variabel dan indikator di atas, kemudian dibangun model input output pendidikan. Dalam model input output pendidikan, input terdiri dari indikator siswa, proses terdiri indikator sarana/prasarana dan indikator guru, sedangkan output terdiri dari indikator mutu pendidikan meliputi indikator jumlah nilai UAN dan rata-rata tingkat kelulusan siswa. Seleksi indikator dilakukan menggunakan analisis faktor dan analisis Struktural Equation Model (SEM).

2.4.3 Model SAR-Kriging

Model SAR-Kriging merupakan gabungan model ekspansi SAR dengan metode Kriging. Model ekspansi SAR dikembangkan oleh Casetti (1972) dalam [6] adalah model kausal yang hanya dapat digunakan untuk prediksi di lokasi-lokasi tersampel. Model ekspansi SAR dinyatakan:

$$y = \alpha + \mathbf{X}\boldsymbol{\beta} + \mathbf{XZ}_x\boldsymbol{\beta}_x + \mathbf{XZ}_y\boldsymbol{\beta}_y + \varepsilon \quad (1)$$

dengan error berdistribusi normal, mean serta variansi konstan. Jika persamaan tersebut melibatkan jarak, persamaan tersebut menjadi:

$$y = \alpha + \mathbf{X}\boldsymbol{\beta} + \mathbf{XD}\boldsymbol{\beta}_0 + \varepsilon \quad (2)$$

Di sisi lain, metode Kriging adalah salah satu metode dalam analisis spasial yang dapat digunakan untuk prediksi di lokasi-lokasi tidak tersampel [2]. Melalui metode Kriging diperoleh bobot untuk lokasi-lokasi tidak tersampel, dinyatakan dalam persamaan:

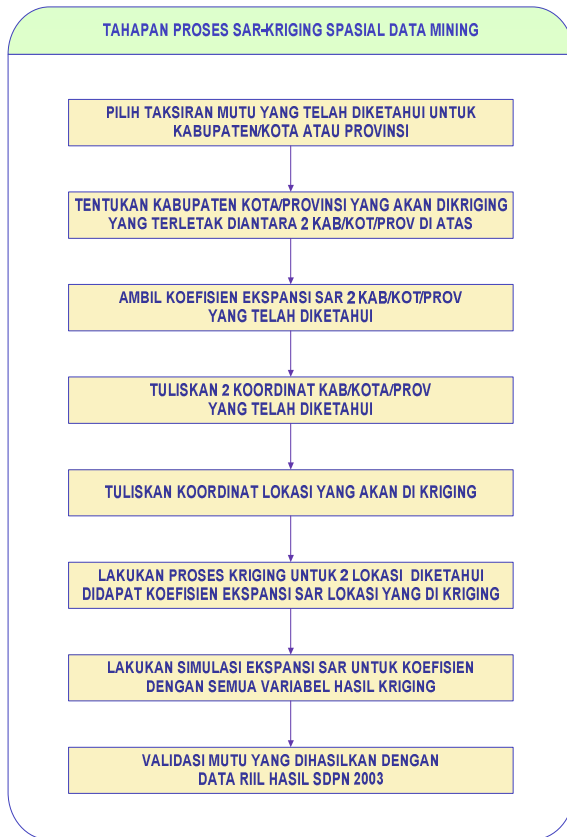
$$\hat{Y}(s_0) = \sum_{i=1}^n w_i Y(s_i) \quad (3)$$

$\hat{Y}(s_0)$ adalah nilai taksiran pada lokasi yang tidak tersampel sedangkan w_i adalah faktor bobot di lokasi i (1, 2, ..., n).

Dalam model SAR-Kriging, taksiran parameter model ekspansi SAR pada persamaan (2) menjadi input dalam metode Kriging pada persamaan (3). Dengan hasil tersebut selanjutnya taksiran model ekspansi SAR di lokasi tidak tersampel dapat diperoleh dan digunakan sebagai prediksi mutu pendidikan di lokasi tidak tersampel dengan memasukkan input variabel-variabel indikator di lokasi tidak tersampel tersebut[1].

Secara operasional, tahapan model SAR-Kriging untuk prediksi mutu pendidikan di lokasi-

lokasi tidak tersampel yang berada di antara dua lokasi sampel digambarkan pada Gambar 2.2.



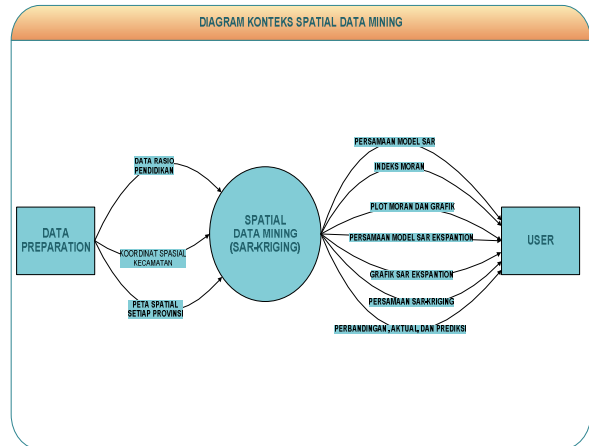
Gambar 2.2 Tahapan Proses SAR-Kriging Spasial Data Mining

2.5 Desain SAR-Kriging

2.5.1 Desain Aplikasi SAR-Kriging

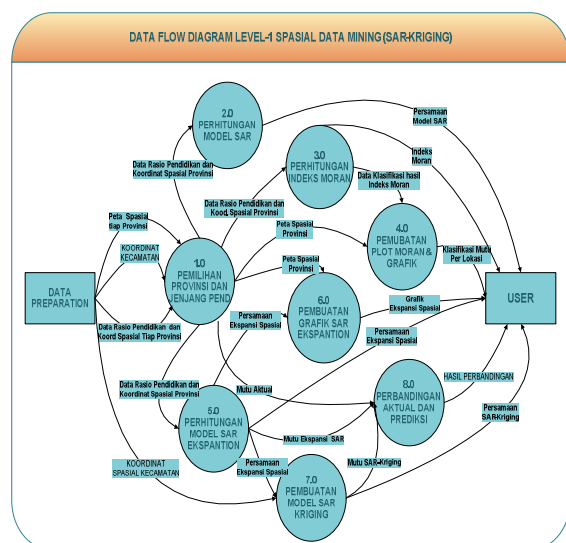
Aplikasi SAR-Kriging, mengikuti metodologi penambangan data, terdiri dari 3 tahap. Pertama *preprocessing*, meliputi seleksi data, data cleaning, dan transformasi data. Hasil dari *preprocessing* berupa data agregat dalam bentuk rasio merupakan gabungan dari indikator penelitian, diturunkan dari indikator siswa, sarana, guru dan mutu pendidikan melalui analisis faktor dan *structural equation model*, serta data koordinat kecamatan-kecamatan sebagai unit analisis terkecil. Kedua, proses data mining, meliputi pemilihan dan penggunaan algoritma model-model untuk pengolahan data hasil *preprocessing*, algoritma model-model yang digunakan antara lain model SAR, model ekspansi SAR dan model SAR-Kriging. Ketiga, *Postprocessing* merupakan interpretasi hasil proses data mining. Output yang dihasilkan berbentuk persamaan SAR untuk prediksi, indeks Moran dan plot Moran dan grafiknya untuk klasifikasi, persamaan model ekspansi SAR, untuk melihat pengaruh data spasial dan non spasial terhadap mutu pendidikan dengan

melibatkan unsur jarak setiap lokasi, serta persamaan hasil SAR-Kriging untuk prediksi mutu pendidikan di lokasi tidak tersampel. Diagram konteks dari aplikasi SAR-Kriging dapat dilihat pada Gambar 2.3.



Gambar 2.3. Diagram Konteks Aplikasi SAR-Kriging

Pada diagram konteks sebagai input adalah data *preprocessing* atau penyiapan data, yang menghasilkan data rasio indikator mutu pendidikan yang diperoleh sebagai hasil seleksi variable menggunakan model analisis faktor dan structural equation model, serta data koordinat setiap lokasi kecamatan dan peta spasial sebagai input untuk proses data mining. Pada proses, dilakukan pengolahan data menggunakan model SAR-Kriging. Output dari aplikasi spasial data mining, adalah taksiran model SAR-Kriging untuk lokasi-lokasi tidak tersampel. Output menghasilkan *knowledge*, yang dapat digunakan user sebagai bahan pendukung keputusan. Data Flow Diagram aplikasi Spasial Data Mining, dapat dilihat pada Gambar 2.4.



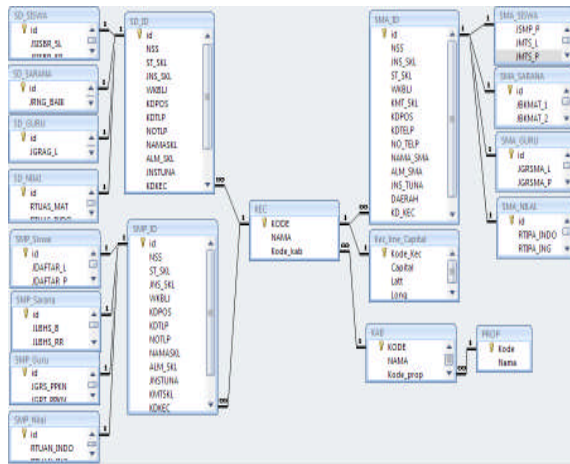
Gambar 2.4. Data Flow Diagram Level-1 Aplikasi SAR-Kriging

Data Flow diagram SAR-Kriging, terdiri dari 8 sub proses, meliputi :

1. Sub proses pemilihan provinsi/kabupaten/kota dan jenjang pendidikan (1.0), untuk memilih provinsi/kab/kota. Input data rasio mutu pendidikan, peta spasial, dan data koordinat tiap kecamatan sebagai unit analisis terkecil, sedangkan output, merupakan perhitungan Ekspansi-SAR dan pembuatan grafik Ekspansi-SAR.
2. Sub proses perhitungan model SAR (2.0), menaksir parameter persamaan model SAR digunakan untuk melihat pengaruh spasial dan non spasial terhadap variabel tertentu dan menghitung koefisien determinasi model spasial variabel tertentu dan unsur error. Input data rasio pendidikan, output persamaan model SAR untuk lokasi tertentu.
3. Sub proses perhitungan indeks Moran (3.0), untuk klasifikasi variabel dalam suatu peta, melalui otokorelasi spasial, jika koefisien MC mendekati 1 artinya nilai pengamatan similar mengelompok dalam suatu peta dan hubungan kuat, jika MC mendekati -1 artinya hubungan lemah. Input data rasio mutu pendidikan, outputnya adalah indeks Moran masing-masing indikator.
4. Sub proses pembuatan plot Moran dan grafiknya (4.0), untuk mengklasifikasikan pengamatan dalam empat kuadran. Koefisien kemiringan ρ , mendekati 1, tingkat otokorelasi spasial positif, plot Moran terdapat pengamatan dominan nilai tinggi berada di kuadran I, kuadran III pengamatan dominan nilai rendah. Scatter plot pola acak, tidak ada ketergantungan antara pengamatan, Nilai negatif ρ di kuadran II, dan IV. Input peta spasial dan data klasifikasi hasil indeks Moran. Outputnya klasifikasi mutu dengan plot Moran dan peta spasial.
5. Sub proses penaksiran parameter model ekspansi-SAR (5.0), menghitung model spasial linier secara lokal dalam mengukur heterogenitas di dasarkan pada koordinat lokasi spasial. Persamaan dihasilkan menunjukkan koefisien variabel individual, menggambarkan pengaruh total terhadap variabel dependen karena variabel bebas. Input data rasio mutu dan koordinat lokasi spasial. Outputnya persamaan taksiran model ekspansi-SAR.
6. Sub proses pembuatan grafik Ekspansi-SAR (6.0), memberikan informasi pengaruh total variabel bebas ke- i terhadap variabel dependen. Jika grafik kecenderungan turun, diinterpretasikan semakin jauh dari koordinat pusat, semakin kecil pengaruh suatu variabel independen terhadap variabel dependen. Jika grafik kecenderungan naik, diinterpretasikan semakin jauh dari pusat, semakin besar pengaruh suatu variabel independen terhadap variabel dependen. Input adalah persamaan ekspansi spasial dan output adalah grafik hasil ekspansi spasial.
7. Sub proses pembuatan model SAR-Kriging (7.0), untuk prediksi di lokasi-lokasi pengamatan yang tidak tersampel. Input persamaan Ekspansi-SAR kab/kota dan koordinat lokasi kecamatan pada lokasi kab/kota. Output adalah persamaan SAR-Kriging untuk masing-masing Kabupaten/Kota.
8. Sub proses perbandingan hasil aktual mutu pendidikan SDPN 2003 dengan mutu hasil prediksi hasil SAR_Kriging, untuk melihat kecocokan model, dengan cara menghitung eror dalam prosentase absolut, selisih antara data aktual dengan hasil prediksi, jika prosentase eror di bawah 10 %, model cocok dengan aktual, jika melebihi 10% model tidak cocok. Input data mutu pendidikan hasil SDPN 2003, hasil ekspansi-SAR dan hasil SAR-Kriging. Outputnya nilai prosentase eror hasil perbandingan antara data aktual dengan data hasil prediksi.

2.5.2 Relasi Antar Tabel

Relasi antar tabel dalam suatu database menunjukkan jaringan data yang ada untuk suatu sistem yang diberikan, dibutuhkan pada proses *preparation* dalam data mining. Relasi antar tabel dapat dilihat pada Gambar 2.5.



Gambar 2.5 Relasi Antar Tabel

Aturan bisnis dari relasi antar tabel di atas adalah sebagai berikut:

1. Tabel SD_SISWA, SD_SARANA, SD_GURU, SD_NILAI, SD_ID, mempunyai nomor identitas sekolah (id) yang unik dan dijadikan sebagai *primary key*.
2. Tabel KEC, KAB, PROP, KEC_LINE_CAPITAL, mempunyai masing Kode yang unik dan dijadikan sebagai *primary key*.
3. Satu identitas sekolah pada tabel SD_ID hanya ada satu identitas sekolah pada tabel SD_SISWA, SD_SARANA, SD_GURU, SD_NILAI.
4. Satu kecamatan terdiri dari beberapa sekolah SD.
5. Satu kecamatan mempunyai satu koordinat spasial.
6. Satu Kabupaten terdiri dari banyak Kecamatan
7. Satu Provinsi terdiri dari banyak Kabupaten

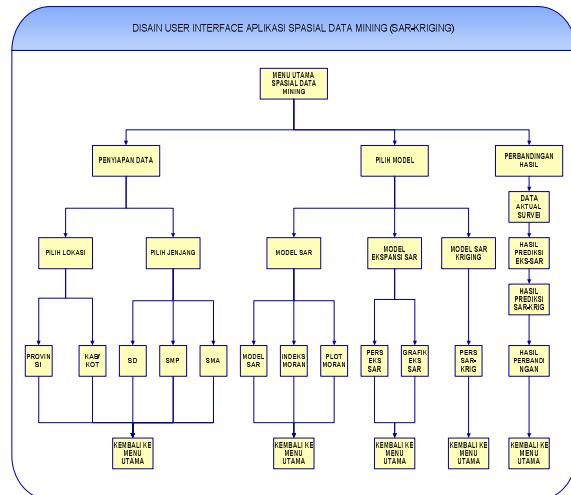
Contoh teks file hasil transformasi untuk jenjang SD ditunjukkan pada Gambar 2. 6.

31.34	0.18	0.48	0.41	0.62	23.64	106.555001	-6.648066
28.59	0.09	0.52	0.31	0.67	23.55	106.611129	-6.615522
36.25	0.09	0.51	0.36	0.62	22.78	106.672554	-6.625359
35.21	0.15	0.40	0.27	0.67	24.31	106.706143	-6.612787
38.13	0.13	0.34	0.38	0.75	25.29	106.732287	-6.57554
35.46	0.23	0.37	0.50	0.74	26.13	106.752269	-6.637841
35.88	0.15	0.50	0.41	0.63	23.32	106.843241	-6.709286
36.96	0.30	0.37	0.61	0.72	22.49	106.878958	-6.698758
38.23	0.16	0.49	0.67	0.73	21.35	106.942236	-6.693755
35.80	0.06	0.46	0.40	0.62	24.00	106.911308	-6.678063
29.60	0.01	0.48	0.21	0.78	25.17	107.131221	-6.560116
29.26	0.06	0.46	0.35	0.76	23.27	107.027467	-6.529019
39.24	0.38	0.36	0.60	0.81	21.98	106.971841	-6.437206
35.66	0.44	0.24	0.52	0.74	25.23	106.926054	-6.377265
37.32	0.41	0.36	0.53	0.79	22.97	106.898905	-6.544311
36.33	0.40	0.35	0.44	0.70	23.73	106.838799	-6.472039
37.86	0.40	0.40	0.43	0.56	21.70	106.781239	-6.460851
33.95	0.15	0.35	0.47	0.69	22.56	106.70066	-6.455514
33.50	0.10	0.36	0.24	0.76	25.91	106.696311	-6.375653
33.57	0.00	0.48	0.34	0.54	24.22	106.65271	-6.442126
34.03	0.01	0.60	0.34	0.70	24.73	106.524596	-6.556658
35.07	0.09	0.51	0.23	0.84	25.06	106.458087	-6.49347
34.16	0.02	0.54	0.30	0.60	25.46	106.495198	-6.358155

Gambar 2.6 Teks File Hasil Preparation Jenjang SD

2.5.3 Desain User Interface SAR-Kriging

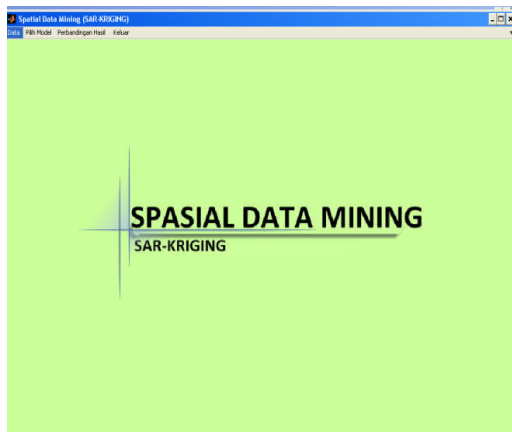
Disain user interfase aplikasi spasial data mining (SAR-Kriging), meliputi menu utama spasial data mining, terdiri dari tiga pilihan, menu penyiapan data, menu pemilihan model dan menu perbandingan model. Data yang digunakan merupakan data hasil *preprocessing*, mulai dari seleksi variabel, *cleaning* data, dan transformasi data ke dalam bentuk rasio, serta penggabungan dengan data spasial. Sedangkan model-model yang digunakan merupakan hasil pemilihan model spasial yang akan digunakan pada proses pengolahan data. Model-model yang digunakan terdiri dari model SAR, model Ekspansi-SAR, model SAR-Kriging. Kemudian dilakukan perbandingan data aktual hasil SDPN 2003 dengan model hasil prediksi SAR-Kriging. Gambar disain user interface yang akan digunakan dapat dilihat pada Gambar 2.7.



Gambar 2.7. Disain User Interface Aplikasi SAR-Kriging

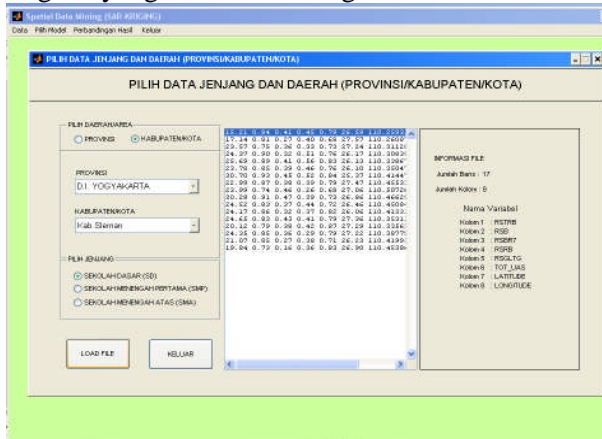
2.5.4 Aplikasi SAR-Kriging

Implementasi aplikasi SAR-Kriging dibuat menggunakan Matlab 7, database dalam microsoft access, sedangkan untuk pembuatan peta digunakan ArtView. Menu utama hasil implementasi aplikasi spasial data mining menggunakan SAR-Kriging, diperlihatkan pada Gambar 2.8.



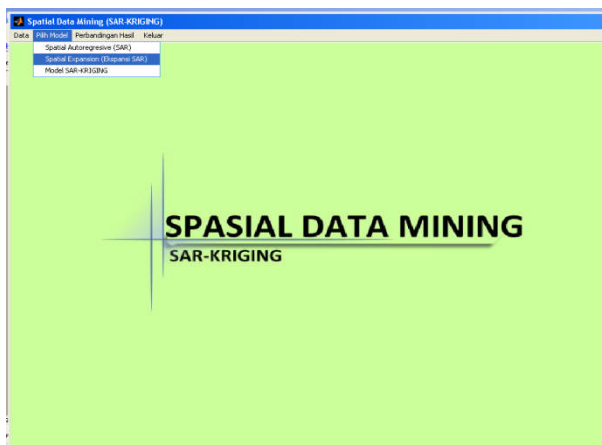
Gambar 2.8. Menu Utama SAR-Kriging

Misalkan diperoleh data mutu pendidikan jenjang SD kabupaten Sleman dan kabupaten Gunungkidul, dan akan mengkriging mutu pendidikan jenjang SD kabupaten Kulonprogo, kabupaten Bantul dan kota Yogyakarta. Langkah-langkah yang dilakukan sebagai berikut.



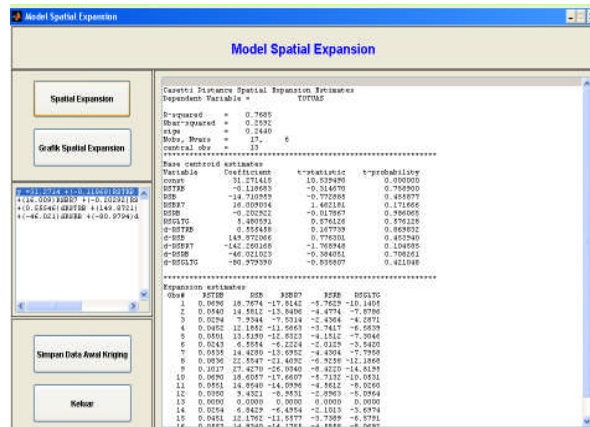
Gambar 2.9. Pemilihan Provinsi DIY dan Kabupaten Sleman

1. Gambar 2.9, menunjukkan langkah pemilihan kabupaten Sleman, sebelumnya didahului memilih provinsi DIY, maka muncul data mutu pendidikan kabupaten Sleman, berikut variabelnya.



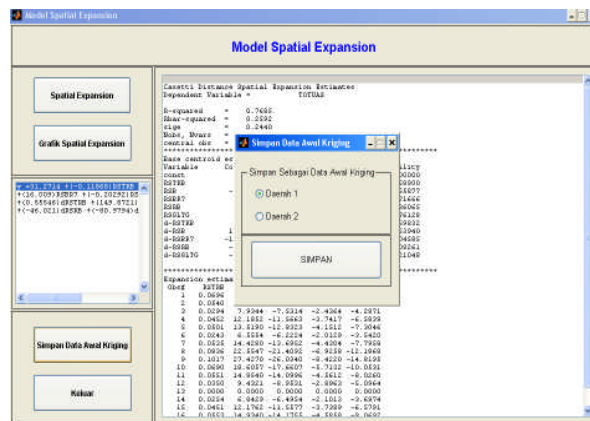
Gambar 2.10. Pemilihan Modul Ekspansi SAR

2. Gambar 2.10, menunjukkan langkah pemilihan modul Ekspansi SAR.



Gambar 2.11. Hasil Eksekusi model Ekspansi SAR kabupaten Sleman

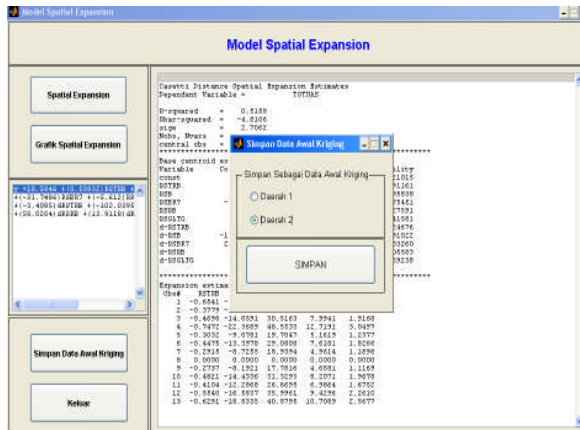
3. Gambar 2.11, hasil eksekusi modul Ekspansi SAR kabupaten Sleman, pada gambar muncul model prediksi mutu pendidikan kabupaten Sleman, berikut koefisien determinasi.



Gambar 2.12. Dialog Penyimpanan Data Awal Kriging daerah 1, kabupaten Sleman

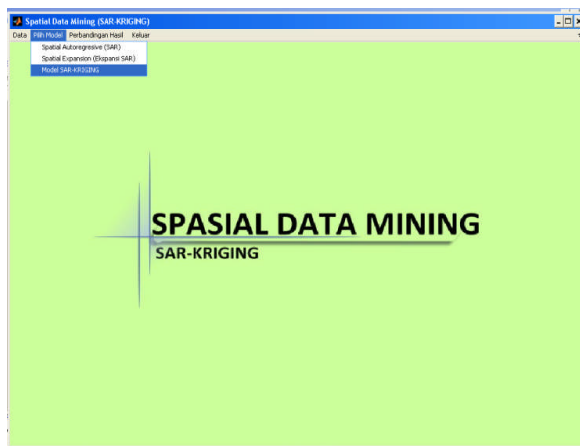
4. Gambar 2.12, menunjukkan langkah penyimpanan untuk kabupaten Sleman, dan disimpan sebagai data awal Kriging daerah 1.

5. Ulangi langkah 1 sampai dengan 4, untuk kabupaten Gunungkidul dan simpan sebagai data awal Kriging daerah 2. Hasilnya dapat dilihat pada Gambar 2.13.



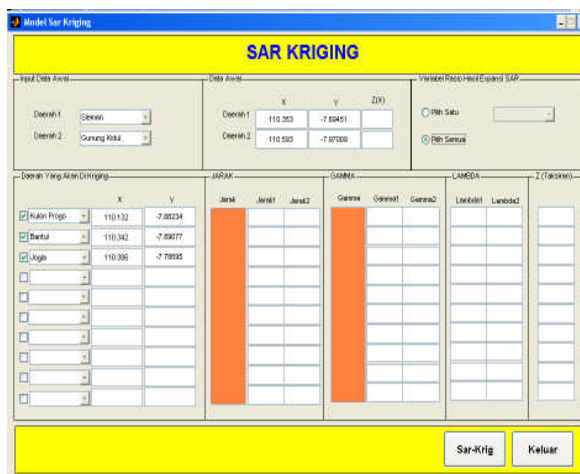
Gambar 2.13. Dialog Penyimpanan Data Awal Kriging Daerah 2, Kabupaten Gunungkidul

6. Langkah berikutnya pilih modul SAR-Kriging, seperti pada Gambar 2.14.



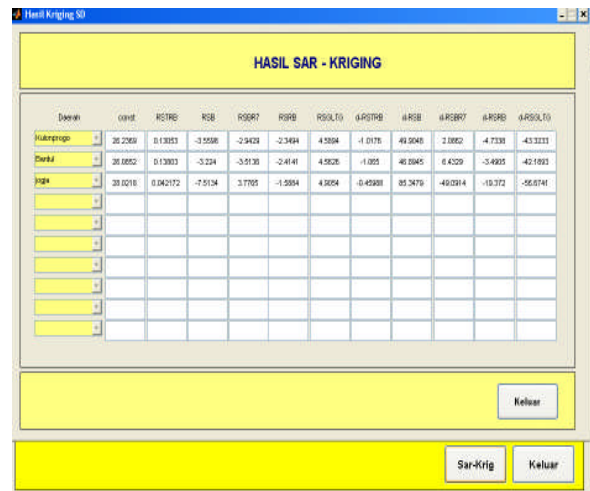
Gambar 2.14. Pemilihan Modul SAR-Kriging diantara kabupaten Selman dan kabupaten Gunungkidul

7. Maka akan muncul seperti pada Gambar 2.15, pada data awal telah dipilih daerah 1 Kabupaten Bantul, daerah 2 Kabupaten Kulonprogo dan daerah 3 kota Yogyakarta.



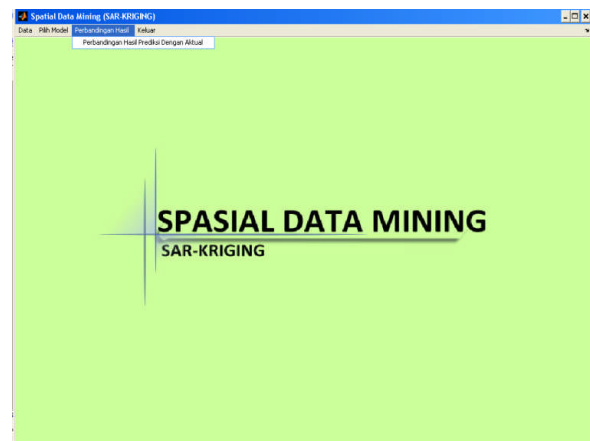
Gambar 2.15. Lokasi kabupaten yang akan di kriging

8. Pilih semua parameter yang ada pada Ekspansi-SAR secara sekaligus, dengan memilih “pilih semua”, hasilnya diperlihatkan pada Gambar 2.16.



Gambar 2.16. Hasil eksekusi model SAR-Kriging

9. Selanjutnya pilih modul perbandingan, yang akan membandingkan data hasil SDPN 2003 dan hasil model taksiran ekspansi-SAR, seperti Gambar 2.17.



Gambar 2.17. Pemilihan modul perbandingan actual dan prediksi.

6. Akan muncul tampilan model untuk hasil model ekspansi-SAR, SAR-Kriging, dan selisih mutu hasil aktual dan prediksi yang ditunjukkan pada Gambar 2.18, Gambar 2.19 dan Gambar 2.20.



NO	KAB/KOTA	MUTU AKTUAL	MUTU PREDIKSI	ERROR	APE
1	BANTUL	28.13	25.81	2.32	8.25%
2	KULONPROGO	25.14	23.23	1.91	7.60%
3	YOGYA	26.47	27.38	0.91	3.44%
MAPE					6.43%

Gambar 2.18. Selisih data actual dan SAR-Kriging kabupaten Kulonprogo



Gambar 2.19. Selisih Data Actual dan SAR-Kriging Kabupaten Bantul



Gambar 2.20. Selisih data actual dan SAR-Kriging kota Yogyakarta

Dari ketiga selisih tersebut diperoleh error model kabupaten Kulonprogo, kabupaten Bantul dan kota Yogyakarta, semuanya di bawah 10 %.

3. HASIL DAN PEMBAHASAN

Data aktual adalah data mutu pendidikan persekolahan merupakan data hasil SDPN 2003. Sedangkan data prediksi merupakan hasil prediksi menggunakan SAR-Kriging untuk lokasi yang tidak tersampel. Hasil perbandingan kemudian dilihat erornya dengan cara melihat kecocokan antara data aktual dan data prediksi. Jika erornya di bawah 10%, model prediksi dinyatakan cocok dengan aktualnya. Jika nilai eror di atas 10 % maka model tidak cocok dengan data aktualnya.

Tabel 3.1 Perbandingan Data Aktual Dengan Data Hasil Prediksi Mutu Pendidikan Jenjang SD

Dari Tabel 3.1 diperoleh rata-rata prosentase eror absolut (MAPE) jenjang SD di Kabupaten Sleman, kabupaten Kulonprogo dan Kota Yogyakarta adalah 6.43%, tingkat rata-rata eror di bawah 10 %. Artinya tingkat kecocokan model akurasi tinggi. Menunjukkan model SAR-Kriging cocok digunakan untuk prediksi mutu pendidikan jenjang SD di Provinsi Yogyakarta.

4. KESIMPULAN

Hasil pengolahan data menggunakan SAR-Kriging, untuk prediksi mutu pendidikan di lokasi tidak tersampel pada berbagai jenjang persekolahan di lokasi kabupaten/kota di DIY menghasilkan *knowledge*, berikut :

- Hasil prediksi mutu pendidikan menggunakan model SAR-Kriging menunjukkan, Rata-rata prosentase eror absolut jenjang SD di 3 kabupaten/kota di DIY adalah 6.43%. menunjukkan tingkat rata-rata eror di bawah 10 %. Artinya tingkat kecocokan model akurasi tinggi. Hal ini menunjukkan model SAR-Kriging cocok digunakan untuk prediksi mutu pendidikan jenjang SD di lokasi yang tidak tersampel.
- Model prediksi mutu pendidikan menggunakan model SAR-Kriging untuk Kabupaten Sleman adalah:

$$Y = 19.76 + (-0.06)RSTRB + 10.89RSB + (-24.59)RSBR7 + 9.18RSRB + 8.23RSGLYK + (0.28)dRSTRB + (-6.40)dRSB + (289.07)dRSBR7 + (-126.39)dRSRB + (-106.19)dRSGLYK$$
 sedangkan model SAR-Kriging untuk kota Yogyakarta adalah:

$$Y = 20.06 + (-0.21)RSTRB + 10.25RSB + (-22.78)RSBR7 + 12.94RSRB + 9.50RSGLYK + (1.21)dRSTRB + (17.90)dRSB + (306.24)dRSBR7 + (-173.23)dRSRB + (-136.70)dRSGLYK$$

5. SARAN

Perlu pengembangan lebih lanjut dari sisi model, aplikasi, dan pemutakhiran data agar model dapat digunakan sesuai perkembangan pendidikan di Indonesia.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Prof. Drs. Subanar, Ph.D, Drs. Retantyo Wardhoyo, M.Sc, Ph.D, Dra. Sri Hartati, M.Sc, Ph.D, Drs. Agus Harjoko, M.Sc, Ph.D yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1]. Abdullah, A. S. 2009. Implementasi Spasial Data Mining menggunakan Model Spasial Autoregressive-Kriging (SAR-Kriging). Disertasi tidak dipublikasikan. Program Doktor Ilmu Komputer, PPs UGM. Yogyakarta.
- [2]. Armstrong, M. 1998. *Basic Linear Geostatistics*. Springer Verlag, New York.
- [3]. Balitbang Depdiknas. 2003. *Survei Dasar Pendidikan Nasional Tahun 2003*. Jakarta.
- [4]. Giudici, P. 2003. *Applied Data Mining Statistical Methods for Business and Industry*. England: John Wiley & Sons Ltd.
- [5]. Han, J. and Kamber, M. 2006. *Data Mining, Concept and Techniques*. Academic Press, USA.
- [6]. LeSage P. J. 1999. *The Theory and Practice of Spatial Econometrics*. Department of Economics, University Toledo.
- [7]. Lu, Z., Xinqi, Z., and Shuqing, W. 2008. Design and Implementation of Spatial Data Mining System (M-SDM) based on MATLAB. *Journal of Computer*. Vol. 3. No. 10. China University of Geosciences, Beijing.
- [8]. Koperski, K., Han, J., and Adhikary, J. 1997. *Spatial Data Mining: Progress and Challenges*. School of CS, Canada.
- [9]. Nababan, H., 2003, *Indikator Mutu Pendidikan di Lingkungan Pendidikan*. Balitbang Depdiknas, Jakarta.
- [10]. Soukup, T. and Davidson, I. 2002, *Visual Data Mining, Techniques and Tools for Data Visualization and Mining*. John Willey & Sons, Inc., Canada.
- [11]. Tan P., Steinbach M. and Kumar V. 2006. *Introduction to Data Mining*. Pearson Education, Inc. Boston.