BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Statistik

2.2.1 Analisis Deskriptif

Menurut (Sugiyono, 2004) Analisis deskriptif adalah statistik yang digunakan untuk menganalisis data dengan mendiskripsikan data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang berlaku untuk umum atau generalisasi. Menurut (Hasan, 2001) statistik deskriptif merupakan bagian dari statistika yang mempelajari cara pengumpulan data dan penyajian data sehingga data mudah dipahami. Statistik deskriptif berhubungan dengan hal menguraikan atau memberikan keterangan-keterangan mengenai data atau keadaan.

Statistika deskriptif berfungsi menerangkan keadaan, gejala, atau persoalan (Priyatno, 2016). (Bambang Suryoatmono, 2004) mengatakan bahwa statistika deskriptif adalah statistika yang menggunakan data untuk menjelaskan atau menarik kesimpulan mengenai suatu kelompok. Dari uraian-uraian diatas dapat disimpulkan bahwa analisis deskriptif atau statistika deskriptif adalah cara mengumpulkan dan menyajikan data dengan tujuan menerangkan keadaan atau suatu permasalahan agar mudah dipahami

2.1.1 *Bayesian Model Averaging* (BMA)

Bayesian Model Averaging (BMA) merupakan metode statistik yang mempertimbangkan model yang tidak pasti dalam pemilihan variabel dengan mengkombinasikan model-model yang terbentuk dari variabel prediktor. Prinsip dasar BMA adalah memprediksi model terbaik berdasarkan rata-rata terboboti dari seluruh model. Hasil dari estimasi mencakup semua model yang kemungkinan terbentuk sehingga bisa mendapatkan hasil estimasi yang lebih baik (Madigan dan Raftery, 1994).

Raftery *et al* (1997) menjelaskan BMA menggunakan pendekatan *Occam's window*. Model *averaging* ini memberikan kinerja prediksi yang lebih baik dibandingkan dengan model tunggal yang terpilih dalam ketidakpastian model. Pendekatan ini digunakan untuk menyelesaikan kasus ekstrim di mana ada banyak variabel prediktor tetapi tidak ada hubungan antara salah satu dari mereka dengan variabel respon, maka dilakukan prosedur pemilihan variabel standar dengan memilih beberapa subset variabel yang menghasilkan R^2 tinggi dan signifikan dalam uji keseluruhan nilai F. Dalam situasi ini, *Occam's Window* biasanya menunjukkan model null sebagai pertimbangan dalam memilih model signifikan. Penelitian ini menggunakan pendekatan *Occam's Window* dalam menyeleksi model.

Fungsi padat probabilitas (PDF) prediktif BMA merupakan kombinasi linier beberapa model prakiraan dimana tiap model memiliki kontribusi atau bobot (*weight*) yang berbeda terhadap pembentukan PDF prediktif. Dalam BMA, penentuan bobot dilakukan berdasarkan konsep rata-rata (model averaging).

Besaran bobot tergantung dari kemampuan model dalam melakukan prakiraan komponen cuaca dimana bobot akan semakin besar jika hasil prakiraan semakin mendekati nilai observasi riil. Dari sudut pandang Bayesian, bobot-bobot tersebut merupakan probabilitas posterior masing-masing model (Feldmann, 2012).

Misalkan $M = (M_1,...,M_k)$ adalah himpunan model yang dipertimbangkan. Sebuah model dapat didefinisikan dengan berbagai atribut seperti himpunan bagian dari variabel penjelas dalam model atau bentuk varian kesalahan. Maka distribusi posterior dari Δ data Z adalah

$$p(\Delta|Z) = \sum_{k=1}^{K} p(\Delta|Z, M_k)p(M_k|Z)$$

Dimana :

M_k : Model amatan

p : distribusi posterior

Δ : parameter model

Z : data amatan

k : jumlah model amatan

Ini adalah rata-rata distribusi prediktif posterior untuk masing-masing model yang dipertimbangkan, dibobot dengan probabilitas model posterior yang sesuai. Probabilitas posterior untuk model M_k sebagai berikut.

$$p(M_k|Z) = \frac{p(Z|M_k)p(M_k)}{\sum_{l=1}^{K} p(Z|M_l)p(M_l)}$$

Dimana

$$p(Z|M_k) = \int p(Z|\theta_1, M_1)p(\theta_1|M_1)d\theta_1 \int ... \int p(Z|\theta_k, M_k)p(\theta_k|M_k)d\theta_k$$

adalah kemungkinan terintegrasi model *Mk* dimana:

θk : vektor parameter model *Mk*

$p(\theta_k|M_k)$: *prior density* dari parameter model *Mk*

$p(Z|\theta_k, M_k)$: probabilitas atau kemungkinan

p(Mk) : prior probabilitas bahwa *Mk* adalah model sebenarnya

Semua probabilitas secara implisit bersyarat pada M, himpunan semua model yang dipertimbangkan. Estimasi parameter dan jumlah ketertarikan lainnya disediakan melalui aplikasi langsung dari prinsip-prinsip yang dijelaskan di atas. Misalnya, estimasi *Bayesian Model Averaging* (BMA) dari sebuah parameter θ adalah

$$\hat{\theta}BMA = \sum_{k=1}^{K} \hat{\theta}kp(M_k|Z)$$

Dimana $\hat{\theta}_k$ menunjukkan *mean posterior* untuk model k. Varians dari estimasi ini dan kuantitas lainnya juga tersedia (misalnya, Hoeting *et al*. 1999 dan Viallefont *et al.* 2001). Ada banyak tantangan dalam implementasi *Bayesian Model Averaging* (BMA), termasuk perhitungan untuk sejumlah besar model, evaluasi integral tersirat yang biasanya tidak ada dalam bentuk tertutup, dan probabilitas spesifikasi model prior *p(Mk)*.

2.1.1 Spasial *Bayesian Model Averaging*

Metode Spasial BMA merupakan salah satu metode *postprocessing* sebagai pencarian solusi prediksi terbaik. Spasial *Bayesian Model Averaging* (Spasial BMA) merupakan kombinasi antara metode *Bayesian Model Averaging* (BMA) dan *Geostatistical Output Perturbation* (GOP) (Berrocal *et.al.*, 2007). Metode Spasial

BMA diharapkan mampu menutupi kekurangan dari metode BMA dan GOP. Seperti pada BMA, *Probability Density Function* (PDF) prediktif Spatial BMA merupakan rata- rata terboboti dari PDF prakiraan yang terpusat pada koreksi bias. Pada Spasial BMA, PDF prakiraan merupakan densitas multivariat dengan struktur kovarians dibentuk untuk mempertimbangkan struktur spasial dari pengamatan.

2.1.2 Matriks Pembobot Spasial

Hubungan kedekatan (*neighbouring*) antar lokasi atau pengamatan dinyatakan dalam matriks pembobot spasial (Bekti dan Stikno, 2010). Pembobot dalam bentuk matriks adalah

$$W = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1n} \\ W_{21} & W_{22} & \cdots & W_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{n1} & W_{n2} & \cdots & W_{nn} \end{bmatrix}$$

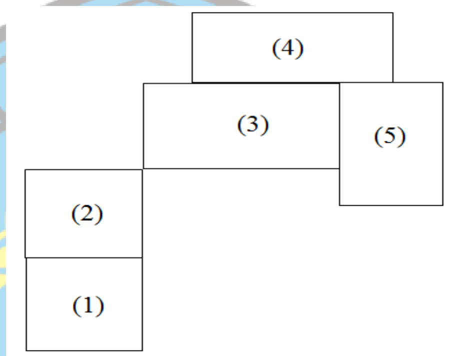
Matriks pembobot spasial W berdasarkan persinggungan batas wilayah (*contiguity*) menyatakan bahwa interaksi spasial terjadi antar wilayah yang bertetangga, yaitu interaksi yang memiliki persinggungan batas wilayah. Matriks pembobot spasial ini bersifat simetri denngan diagonal utamanya bernilai nol.

Menurut (Kosfeld, 2006), matriks pembobot spasial dapat didefinisikan dalam beberapa cara berikut

a Persinggungan *rook* adalah persinggungan sisi wilayah satu dengan sisi wilayah lain yang bertetangga

b Persinggungan *bishop* adalah persinggungan sudut wilayah satu dengan sudut wilayah lain yang bertetangga c Persinggungan *queen* adalah persinggungan sisi maupun sudut wilayah satu dengan wilayah lain yang bertetangga, atau dengan kata lain perpaduan antara persinggungan rook dan bishop.

Pada penelitian ini digunakan matriks pembobot spasial persinggungan *queen*.



Gambar 2. 1 Ilustrasi Persinggungan

Konstruksi matriks pembobot spasial dengan kode biner sebagai berikut

$$W_{ij} = \begin{cases} 1, untuk\ i\ dan\ j\ yang\ berdekatan \\ 0, \quad untuk\ lainnya \end{cases}$$

Dengan W_{ij} adalah elemen matriks pembobot W yang menyatakan ukuran pembobot spasial antara daerah ke -i dan j. Gambar (2.1) merupakan ilustrasi persinggungan mengenai perhitungan matriks pembobot spasial dengan lima daerah sebagai amatannya. Sesuai dengan persinggungan *queen* didapatkan elemen-elemen dari matriks pembobot spasial yaitu $W_{12} = W_{21} = 1$, $W_{23} = W_{32} = 1$, $W_{34} = W_{43} = 1$, $W_{35} = W_{53} = 1$, $W_{45} = W_{54} = 1$, dan yang lain sama dengan nol. Matriks pembobot spasial W ini memiliki ukuran 5 x 5. Matriks pembobot spasial dengan persinggungan *queen* yang dapat terbentuk dari Gambar (2.1) sebagai berikut

$$W = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Matriks W yang diperoleh dibentuk matriks terstandarisasi baris untuk mendapatkan jumlah setiap baris yang sama yaitu bernilai 1, sehingga perhitungan didasarkan pada jumlah tetangga pada satu baris matriks pembobot dengan rumus $W_{ij}^{*} = \frac{\widetilde{W_{IJ}}}{\sum_{j=1}^{n} W_{ij}}; i = 1,2, \ldots, n.$ Bentuk standarisasi matriksnya adalah

$$W = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Optimalisasi matriks W dalam penelitian ini menggunakan pendekatan K-NN dengan tujuan menentukan banyak tetangga yang paling optimal dengan fungsi tujuannya adalah memaksimumkan nilai morans I , R^2, dan meminimumkan AIC.

K-NN dilakukan dengan tahapan:

1. Menghitung jarak Euclidiean lokasi i ke j
2. Mengurutkan jarak yang diperoleh
3. Memilih k lokasi dengan jarak terdekat sebagai nilai optimum.
4. Penentuan nilai k pertama kali di dasarkan pada statistik moran I. Prosesnya dilakukan secara iterasi. Nilai k terpilih berdasarkan nilai

moran I terbesar. Selanjutnya nilai ini digunakan untuk menentukan matrik jarak yang optimum dalam pemodelan spasial Lag dependent.

2.1.3 Uji Efek Spasial

Pada pengujian efek spasial dilakukan berdasarkan uji dependensi spasial (Anselin, 1993) menyatakan bahwa uji untuk mengetahui dependensi spasial dengan menggunakan indeks Moran dan uji pengali Lagrange.

a) Uji pengali Lagrange

Uji pengali Lagrange digunakan untuk mengidentifikasi model spasial yang akan digunakan. Apabila pengali Lagrange lag signifikan maka model spasial BMA dapat digunakan. Pengujian ini dilakukan dengan hipotesis

H_0: tidak terdapat dependensi spasial lag

H_1: terdapat dependensi spasial lag

DK uji ini adalah $\{LM_{lag} | LM_{lag} > X^2_{(a;1)}\}$. H_0 ditolak apabila $LM_{lag} \epsilon DK$ dengan

$$LM_{lag} = \frac{(\varepsilon^T W^* y)^2}{S^2((W^* y\beta)^T M(W^* X\beta) + TS^2)}$$

Dimana $M = I - X(X^T X)^{-1} X^T$, $T = tr[(W^{*T} + W^*)W^*]$, $S^2 = \frac{\varepsilon^T \varepsilon}{n}$, ε adalah vektor error, W^* adalah matriks pembobot terstandarisasi, β adalah vektor parameter regresi dan X adalah matriks variabel independen.

b) Indeks Moran

Indeks Moran digunakan untuk melihat nilai autokorelasi spasial yang digunakan untuk melihat pola penyebaran data. Menurut (Lee dan Wong dalam (Arrowiyah dan Sutikno, 2001)) indeks moran dapat diperoleh menggunakan persamaan

$$I = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \sum_{j=1}^{n} W_{ij}^{*}(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - x)^2}$$

Atau dalam bentuk matriks dapat ditulis

$$I = \frac{(x - \bar{x})^t W^{*} (x - \bar{x})}{(x - \bar{x})^t (x - \bar{x})}$$

Dimana :

I : indeks Moran

n : jumlah pengamatan atau lokasi

x_i : data variabel lokasi ke -i

x_j : data variabel lokasi ke -j

$\bar{x}$: rata-rata variabel x,

w_{ij} : elemen dari matriks pembobot

W^{*} : matriks pembobot terstandarisasi.

Nilai dari indeks Moran berkisar antara -1 dan 1. Identifikasi pola menggunakan kriteria nilai indeks Moran, yaitu jika $I > I_0$ maka memiliki pola mengelompok (cluster), $I < I_0$ memiliki pola menyebar. Jika $I = I_0$ maka memiliki pola menyebar tidak merata (tidak ada autokorelasi) sedangkan $I \neq I_0$ berarti terjadi autokorelasi positif saat I positif dan sebaliknya terjadi autokorelasi negatif saat

I negatif. I_0 merupakan nilai ekspetasi dari I yang dirumuskan dengan $E(I) = I_0 = \frac{-1}{n-1}$

2.1.4 Uji Kebaikan Model

Root Mean Square Error (RMSE) merupakan salah satu indikator yang digunakan untuk mengevaluasi kebaikan model, khususnya untuk data testing. RMSE merupakan akar kuadrat dari MSE, yaitu rata-rata jumlah kuadrat dari selisih antara nilai prakiraan dan observasi. Formula untuk menghitung RMSE sebagai berikut.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Dengan *n* adalah banyak observasi. Kelebihan RMSE terletak pada satuan perhitungan yang sama dengan satuan variabel, sehingga mudah untuk dijadikan sebagai alat perbandingan daripada MSE atau MAE (Anggraeni, 2013).

2.2 Tinjauan Non Statistik

2.2.1 Pendidikan Sekolah Dasar

Pengertian pendidikan dasar yang tercantum dalam Undang-Undang Sistem Pendidikan Nasional No 20 Tahun 2003 di dalam Pasal 17 ialah pendidikan dasar merupakan jenjang pendidikan yang melandasi jenjang pendidikan menengah. Peraturan lain yang memuat tentang pendidikan dasar adalah Peraturan Pemerintah Republik Indonesia Nomor 66 tahun 2010 pasal 1 ayat 7 yang menyebutkan bahwa pendidikan dasar adalah jenis pendidikan formal sebagai landasan menuju jenjang pendidikan menengah, diselenggarakan oleh satuan pendidikan berbentuk Sekolah

Dasar dan Madrasah Ibtidaiyah atau bentuk lain yang sederajat serta menjadi satu kesatuan kelanjutan pada satuan pendidikan berbentuk Sekolah Menengah Pertama dan Madrasah Tsanawiyah atau bentuk lain yang sederajat.

Pada ayat selanjutnya yaitu ayat 8 dijelaskan mengenai Sekolah Dasar atau disingkat SD adalah salah satu bentuk satuan pendidikan formal yang menyelenggarakan pendidikan umum pada jenjang pendidikan dasar. Di sisi lain Peraturan Pemerintah Republik Indonesia Nomor 47 Tahun 2008 Tentang Wajib Belajar dijelaskan pengertian Sekolah Dasar dalam pasal 1 ayat 3 adalah Satuan pendidikan formal yang berperan menyelenggarakan pendidikan umum pada jenjang pendidikan dasar adalah sekolah dasar yang selanjutnya disebut SD.

Dari beberapa peraturan yang menjelaskan tentang pendidikan dasar dapat disimpulkan bahwasannya pendidikan dasar adalah jenis pendidikan formal sebagai landasan siswa didik untuk melanjutkan pendidikan ke jenjang yang lebih tinggi, yaitu jenjang SMP dan sederajatnya. Menurut (Itsnaini, 2015) Sekolah Dasar adalah satuan pendidikan dasar yang disediakan pemerintah dalam bentuk pendidikan formal dalam wujud sekolah dasar.

2.2.2 Angka Putus Sekolah

Menurut E.M. Sweeting dan Muchlisoh dalam laporan teknis Direktorat Jendral Pendidikan Dasar dan Menengah Direktorat Pendidikan Menengah Umum (1998: 14) mengemukakan bahwa siswa putus sekolah adalah siswa yang tidak menyelesaikan pendidikan 6 tahun sekolah dasar dan mereka yang oleh karena itu tidak memiliki ijazah SD. Merujuk pendapat Ali Imron (2004: 125) yang menyatakan bahwa siswa putus sekolah ialah siswa yang dinyatakan telah keluar dari sekolah yang bersangkutan sebelum waktu yang telah ditentukan atau sebelum dinyatakan lulus dan tidak mendapat ijazah dari sekolah.

Angka Putus Sekolah (APTS) merupakan perbandingan antara jumlah siswa putus sekolah pada jenjang pendidikan tertentu dalam hal ini merujuk pada jenjang SD, SMP, maupun SMA dengan jumlah murid pada jenjang pendidikan tersebut dinyatakan dalam persentase. Hasil dari perhitungan APTS biasa digunakan untuk mengetahui banyaknya siswa putus sekolah di suatu jenjang pendidikan tertentu pada suatu wilayah tertentu. Semakin tinggi APTS berarti semakin banyak siswa yang putus sekolah di suatu jenjang pendidikan pada suatu wilayah (Wakhinuddin, 2009). Menurut Badan Pusat Statistika Provinsi Sumatra Utara yang dimaksud dengan APTS adalah tingkat putus sekolah di suatu jenjang pendidikan, misalnya angka putus sekolah jenjang SD menunjukkan persentase anak yang berhenti sekolah sebelum tamat SD yang dinyatakan dalam persen.

Dilansir dari Sirusa BPS menyatakan bahwa semakin tinggi angka putus sekolah menggambarkan kondisi pendidikan yang tidak baik dan tidak merata. Begitu sebaliknya jika angka putus sekolah semakin kecil maka kondisi pendidikan di suatu wilayah semakin baik.

$$APTS = \frac{Jumlah\ penduduk\ yang\ tidak\ sekolah}{Jumlah\ penduduk\ yang\ sedang/pernah\ bersekolah}\, x\ 100\%$$

Dengan objek penelitian APTS jenjang SD sederajat di Provinsi Sumatra Utara dan unit pengamatan pada setiap kabupaten/kota di Provinsi Sumatra Utara yang berupa wilayah (spasial) maka faktor kedekatan wilayah perlu diperhitungkan dengan menggunakan matriks pembobot spasial.