BAB II

LANDASAN TEORI

Landasan teori merupakan sekumpulan teori atau bahan pustaka yang berhubungan dengan permasalahan yang diteliti sehingga dapat memperdalam pengetahuan mengenai permasalahan yang bersangkutan. Dengan adanya landasan teori diharapkan dapat digunakan untuk memecahkan permasalahan yang diangkat dalam penelitian ini yaitu : Penerapan Merode Regresi Linier Berganda untuk prediksi hasil panen jagung di kabupaten Karanganyar. Adapun landasan teori yang digunakan adalah :.

2.1. Metode Regresi Linier Berganda

2.1.1. Regresi Linier

Algoritma *Linear Regression* atau Regresi Linier termasuk jenis aturan *classification* and *regression* pada data mining selain *Linear Regression* yang termasuk pada golongan ini adalah *Support Vector Machine, Logistic Regression* dan lain-lain. Analisis Regresi Linier merupakan proses teknik data mining pada alat statistik yang menjelaskan pola hubungan variable independen merupakan variabel dimana dia mempengaruhi variabel lain yang bersifat berdiri sendiri. Variabel independen ini dinotasikan dengan “X”. Sedangkan variabel dependen merupakan variabel yang dimana keberadaannya dipengaruhi dengan variabel lain yang memiliki sifat tidak berdiri sendiri. Variabel dependen ini dinotasikan dengan “Y” (E. S. Tataming, 2014).

Pada Linear Regression sederhana memiliki variabel bebas hanya ada satu, sedangkan linier regression berganda memiliki variabel bebas lebih dari satu. Diagram yang menggambarkan nilai-nilai observasi peubah tak bebas dan peubah bebas.

Nilai peubah bebas ditulis pada sumbu X (sumbu horizontal)

Nilai peubah takbebas ditulis pada sumbu Y (sumbu vertikal)

Nilai peubah takbebas ditentukan oleh nilai peubah bebas.

Regresi Linier Sederhana mempunyai bentuk fungsi persamaan (Wedasari, N. L. N. M. 2015):

$$Y = a + bX \tag{2.1}$$

Dimana

$$b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2} \tag{2.2}$$

$$a = \frac{\Sigma Y - (b\Sigma x)}{n}$$

Keterangan

Y = Variabel tidak bebas (nilai yang diprediksikan)

X = Variabel bebas

a = Konstanta (nilai Y apabila X= 0)

2.1.2. Regresi Linier Multiple / Berganda

Regresi linier berganda adalah analisis regresi yang menjelaskan hubungan antara peubah respon (variabel dependen) dengan faktor-faktor yang mempengaruhi lebih dari satu prediktor (variabel independen) (Ngumar, 2008). Ketika suatu hasil/keluaran,atau kelas berupa numerik, dan semua atribut adalah

numerik, regresi linear adalah teknik yang tepat untuk menyelesaikan. Ini adalah metode pokok di dalam ilmu statistik. Gunanya adalah untuk mengekspresikan kelas sebagai kombinasi linear dari atribut, dengan bobot yang telah di tentukan:

$$Y = a + b_1X_1 + b_2X_2 + ... + b_nX_n. \quad (2.3)$$

Di mana Y adalah kelas;X1,X2,….,Xn adalah nilai atribut; dan a,b1,…,bn adalah bobot. Bobot dihitung dari data sampel. Disini notasi menjadi sedikit sulit, karena membutuhkan suatu cara untuk mengekpresikan nilai-nilai atribut untuk setiap contoh sampel.

Keterangan

Y = Variabel tidak bebas (nilai yang diprediksikan)

X1 dan *X2*.. = Variabel bebas

a = Konstanta (nilai *Y* apabila *X1, X2…..Xn = 0)*

b1, b2… bk = Koefisien regresi (nilai peningkatan ataupun penurunan)

Dimana nilai *a, b_1, b2… b_k* dapat dihitung dengan metode persamaan normal yaitu :

$$\Sigma Y = a_n + b_1\Sigma X_1 + b_2\Sigma X_2$$

$$\Sigma X_1 Y = a\Sigma X_1 + b_1\Sigma X_1X_2 + b_2\Sigma X_1X_2$$

$$\Sigma X_2 Y = a\Sigma X_2 + b_1\Sigma X_1X_2 + b_2\Sigma X_2$$

Atau nilai *a, b_1, b_2…* b_k dapat juga dihitung dengan metode kuadran terkecil yaitu :

$$b_1 = \frac{(\Sigma x_2{}^2)(\Sigma x_1 y) - (\Sigma x_2 y)(\Sigma x_1 x_2)}{(\Sigma x_1{}^2)(\Sigma x_2{}^2) - (\Sigma x_1 x_2)^2}$$

$$b_2 = \frac{(\Sigma x_1{}^2)(\Sigma x_2 y) - (\Sigma x_1 y)(\Sigma x_1 x_2)}{(\Sigma x_1{}^2)(\Sigma x_2{}^2) - (\Sigma x_1 x_2)^2}$$

$$a = \frac{\Sigma Y - (b_1 \Sigma x_1) - (b_2 \Sigma x_2)}{n}$$

Dimana :

$$\Sigma {x_1}^2 = \Sigma {x_1}^2 - \frac{(\Sigma x_1)^2}{n}$$

$$\Sigma {x_2}^2 = \Sigma {x_2}^2 - \frac{(\Sigma x_2)^2}{n}$$

$$\Sigma x_1 x_2 = \Sigma x_1 x_2 - \frac{(\Sigma x_1)(\Sigma x_2)}{n}$$

$$\Sigma x_1 Y = \Sigma x_1 Y - \frac{(\Sigma x_1)(\Sigma Y)}{n} \Sigma x^2 Y = \Sigma x^1 Y - \frac{(\Sigma x^2)(\Sigma Y)}{n}$$

$$\Sigma Y^2 = \Sigma Y^2 - \frac{(\Sigma Y)^2}{n}$$

Ketika variable bebas lebih dari 2, nilai konstanta dan koefisien regresi setiap variabel bebas dapat diperoleh dengan menggunakan matriks determinan (Supratno, 2000). Contohnya adalah ketika terdapat 3 persamaan dengan 3 variabel yang tidak diketahui nilainya, yaitu a, b1, b2 & b3, persamaan tersebut dapat dinyatakan dalam persamaan matriks sebagai berikut:

$$A = \begin{bmatrix} N & \Sigma X_1 & \Sigma X_2 & \Sigma X_3 \\ \Sigma X_1 & \Sigma(X_1.X_1) & \Sigma(X_1.X_2) & \Sigma(X_1.X_3) \\ \Sigma X_2 & \Sigma(X_2.X_1) & \Sigma(X_2.X_2) & \Sigma(X_2.X_3) \\ \Sigma X_3 & \Sigma(X_3.X_1) & \Sigma(X_3.X_2) & \Sigma(X_3.X_3) \end{bmatrix}$$

$$A0 = \begin{bmatrix} \Sigma(Y) & \Sigma X_1 & \Sigma X_2 & \Sigma X_3 \\ \Sigma(X_1.Y) & \Sigma(X_1.X_1) & \Sigma(X_1.X_2) & \Sigma(X_1.X_3) \\ \Sigma(X_2.Y) & \Sigma(X_2.X_1) & \Sigma(X2.X_2) & \Sigma(X_2.X_3) \\ \Sigma(X_3.Y) & \Sigma(X_3.X_1) & \Sigma(X_3.X_2) & \Sigma(X_3.X_3) \end{bmatrix}$$

$$\mathrm{A1} = \begin{bmatrix} \mathrm{N} & \Sigma(\mathrm{Y}) & \Sigma X_2 & \Sigma X_3 \\ \Sigma X_1 & \Sigma(X_1.\mathrm{Y}) & \Sigma(X_1.X_2) & \Sigma(X_1.X_3) \\ \Sigma X_2 & \Sigma(X_2.\mathrm{Y}) & \Sigma(X_2.X_2) & \Sigma(X_2.X_3) \\ \Sigma X_3 & \Sigma(X_3.\mathrm{Y}) & \Sigma(X_3.X_2) & \Sigma(X_3.X_3) \end{bmatrix}$$

$$A2 = \begin{bmatrix} N & \Sigma X_1 & \Sigma(Y) & \Sigma X_3 \\ \Sigma X_1 & \Sigma(X_1.X_1) & \Sigma(X_1.Y) & \Sigma(X_1.X_3) \\ \Sigma X_2 & \Sigma(X_2.X_1) & \Sigma(X_2.Y) & \Sigma(X_2.X_3) \\ \Sigma X_3 & \Sigma(X_3.X_1) & \Sigma(X_3.Y) & \Sigma(X_3.X_3) \end{bmatrix}$$

$$A3 = \begin{bmatrix} N & \Sigma X_1 & \Sigma X_2 & \Sigma(Y) \\ \Sigma X_1 & \Sigma(X_1.X_1) & \Sigma(X_1.X_2) & \Sigma(X_1.Y) \\ \Sigma X_2 & \Sigma(X_2.X_1) & \Sigma(X_2.X_2) & \Sigma(X_2.Y) \\ \Sigma X_3 & \Sigma(X_3.X_1) & \Sigma(X_3.X_2) & \Sigma(X_3.Y) \end{bmatrix}$$

Kemudian dapat dilakukan perhitungan untuk determinasi matriks A, A0, A1, A2 dan A3 sebagai berikut:

$$\begin{aligned} Det(A) = {} & N\,\Sigma(X_1X_1)\,\Sigma(X_2X_2)\,\Sigma(X_3X_3) + N\,\Sigma(X_1X_2)\,\Sigma(X_2X_3)\,\Sigma(X_3X_1) + N \\ & \Sigma(X_1X_3)\,\Sigma(X_2X_1)\,\Sigma(X_3X_2) + \Sigma X_1\,\Sigma X_1\,\Sigma(X_2X_3)\,\Sigma(X_3X_2) + \Sigma X_1 \\ & \Sigma(X_1X_2)\,\Sigma X_2\,\Sigma(X_3X_3) + \Sigma X_1\,\Sigma(X_1X_3)\,\Sigma(X_2X_2)\,\Sigma X_3 + \Sigma X_2\,\Sigma X_1 \\ & \Sigma(X_2X_1)\,\Sigma(X_3X_3) + \Sigma X_2\,\Sigma(X_1X_1)\,\Sigma(X_2X_3)\,\Sigma X_3 + \Sigma X_2\,\Sigma(X_1X_3)\,\Sigma X_2 \\ & \Sigma(X_3X_1) + \Sigma X_3\,\Sigma X_1\,\Sigma(X_2X_2)\,\Sigma(X_3X_1) + \Sigma X_3\,\Sigma(X_1X_1)\,\Sigma X_2\,\Sigma(X_3X_2) \\ & + \Sigma X_3\,\Sigma(X_1X_2)\,\Sigma(X_2X_1)\,\Sigma X_3 - N\,\Sigma(X_1X_1)\,\Sigma(X_2X_3)\,\Sigma(X_3X_2) - N \\ & \Sigma(X_1X_2)\,\Sigma(X_2X_1)\,\Sigma(X_3X_3) - N\,\Sigma(X_1X_3)\,\Sigma(X_2X_2)\,\Sigma(X_3X_1) - \Sigma X_1\,\Sigma X_1 \\ & \Sigma(X_2X_2)\,\Sigma(X_3X_3) - \Sigma X_1\,\Sigma(X_1X_2)\,\Sigma(X_2X_3)\,\Sigma X3 - \Sigma X_1\,\Sigma(X_1X_3)\,\Sigma X_2 \\ & \Sigma(X_3X_2) - \Sigma X_2\,\Sigma X_1\,\Sigma(X_2X_3)\,\Sigma(X_3X_1) - \Sigma X_2\,\Sigma(X_1X_1)\,\Sigma X_2\,\Sigma(X_3X_3) - \\ & \Sigma X_2\,\Sigma(X_1X_3)\,\Sigma(X_2X_1)\,\Sigma X_3 - \Sigma X_3\,\Sigma X_1\,\Sigma(X_2X_1)\,\Sigma(X_3X_2) - \Sigma X_3 \\ & \Sigma(X_1X_1)\,\Sigma(X_2X_2)\,\Sigma X_3 - \Sigma X_3\,\Sigma(X_1X_2)\,\Sigma X_2\,\Sigma(X_3X_1) \end{aligned} \qquad (2.4)$$

Dengan cara yang sama seperti menghitung Det(A), dapat diperoleh pula Det(A0), Det(A1), Det(A2) & Det(A3).

$$a = \frac{det(A0)}{det(A)} \qquad (2.5)$$

$$b1 = \frac{det(A1)}{det(A)} \qquad (2.6)$$

$$b2 = \frac{det(A2)}{det(A)} \qquad (2.7)$$

$$b3 = \frac{det(A3)}{det(A)} \qquad (2.8)$$

2.2. MAPE *(Mean Absolute Percentage Error)*

Dalam banyak situasi peramalan, ketepatan dipandang sebagai kriteria penolakan untuk memilih suatu metode peramalan. Jika *Xi* merupakan data aktual (data hasil pengamatan = observasi) untuk periode i dan *Fi* merupakan ramalan untuk periode yang sama (periode i), maka kesalahan didefinisikan sebagai (Darsono, 2008) :

$$E_i = x_i - f_i$$

Jika terdapat nilai pengamatan dan ramalan untuk n periode waktu, maka akan terdapat n buah galat (kesalahan = error). Nilai-nilai yang umum digunakan untuk mengukur ketepatan pemakaian suatu metode peramalan tertentu dalam suatu kumpulan data adalah : MAPE (*Mean Absolute Percentage Error)*, MAD *(Mean Absolute Deviation),* dan MSD (*Mean Square Deviation)* atau MSE *(Mean Square Error).* Untuk menentukan metode peramalan mana yang sesuai, pada penelitian ini menggunakan adalah MAPE. Rumus dari MAPE *(Mean Absolute Percentage Error)* adalah sebagai berikut :

$$MAPE = \frac{\sum_{i=1}^{n} |PE|}{n} \quad \text{dimana} \quad PE_i = \left(\frac{x_i - f_i}{x_i}\right)(100\%) \qquad (2.9)$$

Untuk mengevaluasi hasil dari peramalan pada penelitian ini digunakan MAPE. MAPE mengukur error mutlak sebagai persentase bukan dari tiap periodenya melainkan dari rata-rata error mutlak pada sejumlah periode data aktual. Hal tersebut dapat menghindari permasalahan dalam interpretasi pengukuran akurasi relatif terhadap besarnya nilai aktual dan nilai prediksi. Nilai yang dihasilkan malalui evaluasi ini, menunjukkan kemampuan peramalan seperti yang ditunjukkan dalam kriteria MAPE pada Tabel 2.1 Kriteria MAPE (Setiyoutami, 2012). Dimana nilai MAPE di bawah 20% sudah dapat dikatakan baik, dan nilai MAPE kurang dari 10% dinyatakan sangat baik.

Tabel 2. 1 Kriteria MAPE (Setiyoutami, 2012)

MAPE	**Pengertian**
< 10%	Sangat Baik
10 % - 20 %	Baik
20 % - 50 %	Cukup
> 50 %	Buruk

2.3. Prediksi (*Forecasting*)

Prediksi adalah suatu proses memprediksi secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prakiraan tidak harus memberikan jawaban secara pasti tentang kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. Prakiraan merupakan penggunaan data masa lalu dari sebuah variabel atau kumpulan variabel untuk mengestimasi nilainya di masa yang akan datang. Memprakirakan atau memprediksi hasil panen jagung untuk masa yang akan datang merupakan pekerjaan yang tidak mudah karena harus mempertimbangkan berbagai faktor yang sangat mempengaruhi hasil panen jagung. Prakiraan atau prediksi disebut juga peramalan *(forecast)* yang pada dasarnya merupakan dugaan atau prakiraan mengenai peristiwa di masa yang akan datang. Secara umum tahapan dalam membuat prakiraan untuk hasil panen jagung di bagi menjadi 3 tahap, yaitu:

1. Pengumpulan dan penyiapan data
2. Pengolahan dan analisa data
3. Penentuan Metode dan pembuatan model. (Rahman, 2015)

2.4. Jagung (*Zea Mays L*)

Tanaman jagung yang dalam bahasa ilmiahnya disebut *Zea mays L*, adalah salah satu jenis tanaman biji-bijian yang menurut sejarahnya berasal dari Amerika. Orang-orang Eropa yang dating ke Amerika membawa benih jagung ke negaranya. Melalui Eropa tanaman jagung terus menyebar ke Asia dan Afrika.

Baru sekitar abad ke-16 tanaman jagung ini dibawa oleh Portugis ke Pakistan, Tiongkok dan daerah-daerah lainnya di Asia termasuk Indonesia (Wirawan dan Wahab, 2007).

Jagung merupakan tanaman serealia yang paling produktif di dunia, sesuai ditanam di wilayah bersuhu tinggi, dan pematangan tongkol ditentukan oleh akumulasi panas yang diperoleh tanaman. Luas pertanaman jagung di seluruh dunia lebih dari 100 juta ha, menyebar di 70 negara, termasuk 53 negara berkembang. Penyebaran tanaman jagung sangat luas karena mampu beradaptasi dengan baik pada berbagai lingkungan. Jagung tumbuh baik di wilayah tropis hingga 50° LU dan 50° LS, dari dataran rendah sampai ketinggian 3.000 m di atas permukaan laut (dpl), dengan curah hujan tinggi, sedang, hingga rendah sekitar 500 mm per tahun (Dowswell et al, 1996). Pusat produksi jagung di dunia tersebar di negara tropis dan subtropis.

2.5. PHP *(Hypertext Processor)*

PHP kepanjangan dari *Hypertext Preprocessor* yaitu bahasa pemrograman *web server-side* yang bersifat *open source*. PHP merupakan scriptyang terintegrasi dengan HTML dan berada pada server *(server side HTML embedded scripting).*

Bahasa pemrograman PHP adalah salah satu bahasa pemrograman untuk aplikasi web yang terkenal. Bahasa ini dieksekusi oleh server web setiap kali ada permintaan. Banyak sekali keuntungan menggunakan PHP, antara lain : penggunaan *syntax* pemrograman bahasa C yang sudah terkenal, tidak adanya tipe-tipe data untuk variabel yang digunakan, serta banyaknya fungsi yang disediakan untuk ekspansi menggunakan sistem lain. Selain itu, PHP merupakan

bahasa pemrograman yang sifatnya *open source* artinya bahasa ini boleh dipakai oleh siapapun juga tanpa harus mengeluarkan biaya untuk lisensinya. PHP juga boleh dimodifikasi sesuai kebutuhan penggunanya. Bahasa pemrograman ini dieksekusi oleh *web server*, maka tidak dibutuhkan *compiler* khusus untuk editor PHP.

Kelebihan PHP dibandingkan bahasa pemrograman lain :

a. Bahasa pemrograman PHP adalah sebuah bahasa *script* yang tidak memerlukan sebuah kompilasi dalam penggunaanya.
b. Web server yang mendukung PHP dapat ditemukan dimana-mana dari mulai *Apache, IIS, Lighttpd,* hingga *Xitami* dengan konfigurasi yang relative mudah.
c. Sisi pengembangan lebih mudah karena banyakya milis-milis dan *developer* yang siap membantu dalam pengembangan.
d. Referensi untuk bahasa *scripting* PHP banyak tersedia secara gratis, sehingga hal ini dapat meningkatkan sisi pemahaman dari pengguna (Anhar, 2010).

2.6. *MySQL* (*My Structure Query Language*)

MySQL merupakan sebuah perangkat lunak sistem manajemen basis data *SQL / DBMS* (*Database Management System*) yang *multithread*, *multiuser* dan sekitar 6 juta instalasi diseluruh indonesia. Didistribusikan secara gratis dibawah lisensi GPL (*General Public License*). Dimana setiap orang bebas untuk menggunakan *MySQL*, namun tidak boleh dijadikan produk turunan yang bersifat komersial. Keistimewaan *MySQL* yakni :

a. Portabilitas. *MySQL* dapat berjalan stabil pada berbagai sistem operasi seperti *Windows, Linux, FreeBSD, Mac Os X Server, Solaris, Amiga,* dan masih banyak lagi.
b. *Open Source. MySQL* didistribusikan secara *open source*, dibawah lisensi GPL sehingga dapat digunakan secara cuma-cuma.
c. *Multiuser*. *MySQL* dapat digunakan oleh beberapa user dalam waktu yang bersamaan tanpa mengalami masalah atau konflik.
d. Performance tuning. *MySQL* memiliki kecepatan yang menakjubkan dalam menangani *query* sederhana, dengan kata lain dapat memproses lebih banyak SQL per satuan waktu.
e. Jenis Kolom. *MySQL* memiliki tipe kolom yang sangat kompleks, seperti *signed / unsigned integer, float, double, char, text, date, timestamp*, dan lain lain.
f. Perintah dan Fungsi. *MySQL* memiliki operator dan fungsi secara penuh yang mendukung perintah *Select* dan *Where* dalam perintah *(query).*
g. Keamanan. *MySQL* memiliki beberapa lapisan sekuritas seperti level *subnetmask*, nama *host*, dan izin akses *user* dengan sistem perizinan yang mendetail serta sandi terenkripsi.

Dukungan banyak komunitas, biasanya tergabung dalam sebuah forum untuk berdiskusi dan bertukar informasi tentang *MySQL*. (Zinner, 2010).