

BAB III

METODE PENELITIAN

Metode penelitian yang digunakan dalam penulisan Tugas Akhir ini, sebagai berikut:

3.1 Instrumen Penelitian

Berdasarkan permasalahan yang telah diuraikan pada bab sebelumnya, diperlukan bahan dan peralatan yang dapat membantu dalam proses kelancaran pembuatan penelitian ini yang meliputi:

3.1.1 Bahan

Dalam penelitian ini bahan yang dibutuhkan dan digunakan adalah data nilai matematika semester I tahun ajaran 2015/2016 siswa SMA Kesatrian 1 Semarang yang beralamat di Jalan Pamularsih 116 Semarang.

3.1.2 Peralatan

Dalam penelitian ini diperlukan peralatan yang meliputi kebutuhan perangkat keras (*hardware*) dan kebutuhan perangkat lunak (*software*). Kebutuhan perangkat keras dan kebutuhan perangkat lunak terdiri sebagai berikut:

Kebutuhan perangkat keras:

- a. Processor Intel(R) Core (TM) i5-2410M CPU @ 2.30GHz (4 CPUs), ~2.3GHz
- b. RAM 2048MB
- c. Layar monitor 14"
- d. Flashdisk 16GB
- e. Mouse
- f. Printer

Kebutuhan perangkat lunak:

- a. Sistem operasi Windows 8 Pro 32-bit

Sistem operasi yang digunakan pada laptop penulis.

- b. Microsoft Office Word Professional Plus 2010
Software ini digunakan untuk proses penulisan laporan hasil penelitian.
- c. Microsoft Office Excel Professional Plus 2010
Software ini digunakan untuk penulisan dataset.
- d. RapidMiner Studio 5.3.013
Framework yang digunakan untuk menguji dan melihat hasil akurasi dari dataset yang diteliti.
- e. Matlab Versions 7.10.0.499 (R2010a)
Framework yang digunakan untuk melihat hasil kinerja dari algoritma yang digunakan terhadap dataset.

3.2 Prosedur Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data yang diperoleh menurut sifat dan cara memperolehnya, yaitu data kualitatif, data kuantitatif, dan data primer.

Menurut sifatnya:

- Data kualitatif
Data yang diperoleh bukan dalam bentuk bilangan, yaitu nama, jenis kelamin, agama, jurusan, dan kelas.
- Data kuantitatif
Data yang diperoleh berupa bilangan, yaitu NIS, nilai UHT 1, nilai UHT 2, dan nilai mid. Dengan jumlah data asli sebanyak 1009.

Menurut cara memperolehnya:

- Data primer
Data yang diperoleh secara langsung dari sumber yang menjadi objek penelitian. Data yang diperoleh dari SMA Kesatrian 1 Semarang tersebut berupa data nilai matematika dalam bentuk *softcopy* di Microsoft Excel.

Digunakan beberapa studi pustaka dalam membantu penyusunan tugas akhir ini yang merupakan data sekunder, seperti:

1. E-book mengenai data mining dan algoritmanya.
2. Jurnal mengenai kasus estimasi dan model regresi linear berganda.
3. Penelitian sebelumnya yang digunakan sebagai referensi dan tinjauan pustaka.

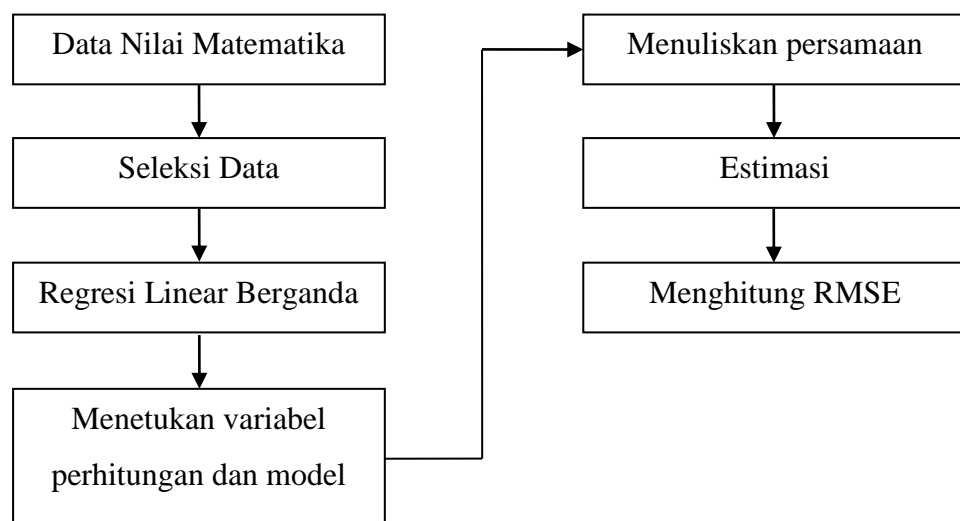
3.3 Teknik Analisis Data

Beberapa tahapan dilakukan dalam penelitian ini terhadap data yang akan diuji. Tahapan yang dilakukan antara lain:

1. Menyeleksi data asli yang diperoleh menjadi data yang siap digunakan dalam penelitian.
2. Mengelompokkan data berdasarkan jurusan.
3. Melakukan perhitungan estimasi menggunakan algoritma regresi linear berganda.
4. Melakukan perhitungan standar error dari hasil estimasi.

3.4 Metode Yang Diusulkan

Metode yang diusulkan untuk perhitungan estimasi menggunakan algoritma regresi linear, sebagai berikut:



Gambar 3.1 Metode yang diusulkan

Dari gambar 3.1 dapat dijelaskan langkah-langkah sebagai berikut:

1. Data yang akan diuji untuk penelitian adalah data nilai matematika yang terdiri dari nilai UHT 1, nilai UHT 2, dan nilai mid.
2. Data nilai matematika yang masih berupa data asli diseleksi dengan menghilangkan nilai atau angka 0 (nol) guna mempermudah proses perhitungan. Nilai 0 (nol) tersebut dikarenakan nilai belum masuk pada rekap data.
3. Mengelompokkan data nilai matematika sesuai dengan jurusan (IPA, IPS, dan Bahasa).
4. Menentukan variabel, perhitungan, dan model menggunakan tahapan regresi linear berganda.
5. Menuliskan persamaan regresi setelah mendapatkan model atau *rules*.
6. Dilakukan estimasi perhitungan.
7. Menghitung RMSE (*Root Mean Square Error*) untuk mengetahui seberapa besar error yang didapat.

3.5 Eksperimen dan Cara Pengujian Metode

Dilakukan eksperimen dan pengujian metode terhadap penelitian yang dilakukan dengan menggunakan metode yang diusulkan di atas.

3.5.1 Contoh Kasus

Estimasi hasil produktivitas tanaman tebu di Kabupaten Rembang dengan menggunakan 10 sampel data luas tanam bulan Maret 2013, produktivitas bulan Maret 2013, dan produksi Maret 2014.

Dengan menggunakan rumus:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

Y = variabel tidak bebas

a = konstanta

b_1, b_2 = bobot atau koefisien regresi

X_1, X_2 = variabel bebas

Tabel 3.1 dibawah ini merupakan data kasus yang akan digunakan sebagai sampel penelitian.

Tabel 3.1 Data Kasus

No	Luas (Ha)	Produksi (Ton)	Produktivitas (Kg/Ha)
1	120	1200	144
2	18	1190	22
3	21	900	25
4	18	825	22
5	60	975	72
6	10	1100	131
7	69	960	98
8	31	784	17
9	19	1260	257
10	23	883	44

Tabel 3.1 di atas terdiri dari sampel data yang masing-masing berjumlah 10. Setelah diketahui sampel data seperti di atas, kemudian dilakukan perhitungan sehingga data yang telah dihitung tersebut dapat digunakan sebagai data pembantu untuk proses perhitungan selanjutnya.

Tabel 3.2 Data Pembantu

No	X_1	X_2	Y	$X_1 Y$	$X_2 Y$	$X_1 X_2$	X_1^2	X_2^2	Y^2
1	120	1200	144	17280	172800	144000	14400	1440000	20736
2	18	1190	22	396	26180	21420	324	1416100	484
3	21	900	25	525	22500	18900	441	810000	625
4	18	825	22	396	18150	14850	324	680625	484
5	60	975	72	4320	70200	58500	3600	950625	5184
6	10	1100	131	1310	144100	11000	100	1210000	17161
7	69	960	98	6762	94080	66240	4761	921600	9604
8	31	784	17	527	13328	24304	961	614656	289

9	19	1260	257	4883	323820	23940	361	1587600	66049
10	23	883	44	1012	38852	20309	529	779689	1936
Σ	389	10077	832	37411	924010	403463	25801	10410895	122552

Dari tabel 3.2 di atas terdapat 3 variabel yang digunakan, yaitu X_1 , X_2 , dan Y yang masing-masing mempunyai nilai berbeda dan terdiri dari 2 jenis variabel yaitu variabel independen/bebas dan variabel dependen/tidak bebas. Dalam penelitian ini, luas dan produksi dilambangkan X_1 dan X_2 , yaitu sebagai variabel independen, sedangkan produktivitas dilambangkan Y , yaitu sebagai variabel dependen. Setelah menentukan jenis variabel kemudian dicari nilai bobot a , a_1 , dan a_2 dengan menggunakan algoritma linear regresi, yaitu regresi linear berganda sehingga didapat persamaan regresi.

Perhitungan regresi linear berganda:

$$\begin{aligned}\Sigma x_1^2 &= \Sigma X_1^2 - ((\Sigma X_1)^2 / n) &= 25801 - (151321 / 10) \\ & &= 10668.9\end{aligned}$$

$$\begin{aligned}\Sigma x_2^2 &= \Sigma X_2^2 - ((\Sigma X_2)^2 / n) &= 10410895 - (101545929 / 10) \\ & &= 256302.1\end{aligned}$$

$$\begin{aligned}\Sigma y^2 &= \Sigma Y^2 - ((\Sigma Y)^2 / n) &= 122552 - (692224 / 10) \\ & &= 53329.6\end{aligned}$$

$$\begin{aligned}\Sigma x_1 y &= \Sigma X_1 Y - ((\Sigma X_1 \cdot \Sigma Y) / n) &= 37411 - ((389 \cdot 832) / 10) \\ & &= 37411 - (323648 / 10) \\ & &= 5046.2\end{aligned}$$

$$\begin{aligned}\Sigma x_2 y &= \Sigma X_2 Y - ((\Sigma X_2 \cdot \Sigma Y) / n) &= 924010 - ((10077 \cdot 832) / 10) \\ & &= 924010 - (8384064 / 10) \\ & &= 85603.6\end{aligned}$$

$$\begin{aligned}\sum x_1 x_2 &= \sum X_1 X_2 - ((\sum X_1 \cdot \sum X_2) / n) = 403463 - ((389 \cdot 832) / 10) \\ &= 403463 - (323648 / 10) \\ &= 371098.2\end{aligned}$$

Setelah mendapatkan hasil dari perhitungan di atas, kemudian dilakukan perhitungan untuk mencari a , b_1 , dan b_2 .

$$\begin{aligned}b_1 &= \frac{[(\sum x_2^2 \cdot \sum x_1 y) - (\sum x_2 y \cdot \sum x_1 x_2)]}{[(\sum x_1^2 \cdot \sum x_2^2) - (\sum x_1 x_2)^2]} \\ &= \frac{[(256302.1 \times 5046.2) - (85603.6 \times 371098.2)]}{[(10668.9 \times 256302.1) - (371098.2)^2]} \\ &= \frac{1293351657 - 31767341874}{2734461475 - 137713874043.24} \\ &= \frac{-30473990217}{-134979412568.24} \\ &= 0.22577\end{aligned}$$

$$\begin{aligned}b_2 &= \frac{[(\sum x_1^2 \cdot \sum x_2 y) - (\sum x_1 y \cdot \sum x_1 x_2)]}{[(\sum x_1^2 \cdot \sum x_2^2) - (\sum x_1 x_2)^2]} \\ &= \frac{[(10668.9 \times 85603.6) - (5046.2 \times 371098.2)]}{[(10668.9 \times 256302.1) - (371098.2)^2]}\end{aligned}$$

$$\frac{913296248 - 1872635737}{2734461475 - 137713874043.24}$$

$$\frac{-959339488.8}{-134979412568.24}$$

$$= 0.00711$$

$$a = \frac{(\sum Y) - (b_1 \cdot \sum X_1) - (b_2 \cdot \sum X_2)}{n}$$

$$\frac{832 - (0.22577 \times 389) - (0.00711 \times 10077)}{10}$$

$$\frac{832 - 87.82453 - 71.64747}{10}$$

$$= 67.2528$$

Dari perhitungan di atas, didapatkan persamaan regresi linear sebagai berikut:

$$Y = 67.2528 + 0.22577 X_1 + 0.00711 X_2$$

Dengan menggunakan persamaan regresi linear di atas, maka estimasi produktivitas dapat dihitung dengan memasukkan X_1 (luas) dan X_2 (produksi), seperti tabel 3.3 di bawah ini:

Tabel 3.3 Data Hasil Estimasi

No	a	b_1	b_2	$b_1 X_1$	$b_2 X_2$	Hasil Estimasi
1	67.2528	0.22577	0.00711	27.0924	8.532	102.8772
2	67.2528	0.22577	0.00711	4.06386	8.4609	79.77756
3	67.2528	0.22577	0.00711	4.74117	6.399	78.39297
4	67.2528	0.22577	0.00711	4.06386	5.86575	77.18241
5	67.2528	0.22577	0.00711	13.5462	6.93225	87.73125
6	67.2528	0.22577	0.00711	2.2577	7.821	77.3315
7	67.2528	0.22577	0.00711	15.57813	6.8256	89.65653
8	67.2528	0.22577	0.00711	6.99887	5.57424	79.82591
9	67.2528	0.22577	0.00711	4.28963	8.9586	80.50103
10	67.2528	0.22577	0.00711	5.19271	6.27813	78.72364
Jumlah						832

3.5.2 Ukuran Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$$

Keterangan:

Y_i = data awal (data sebenarnya)

\hat{Y}_i = data akhir (data hasil estimasi)

n = jumlah data

\hat{Y}_i didapat dari persamaan regresi $Y = 67.2528 + 0.22577 X_1 + 0.00711 X_2$ dengan memasukkan X_1 (luas), X_2 (produksi) ke dalam persamaan tersebut.

Tabel 3.4 Perhitungan RMSE

No	Yi	Ŷi	Yi - Ŷi	(Yi - Ŷi) ²
1	144	102.8772	41.1228	1691.08468
2	22	79.77756	-57.77756	3338.24644
3	25	78.39297	-53.39297	2850.809245
4	22	77.18241	-55.18241	3045.098373
5	72	87.73125	-15.73125	247.4722266
6	131	77.3315	53.6685	2880.307892
7	98	89.65653	8.34347	69.61349164
8	17	79.82591	-62.82591	3947.094967
9	257	80.50103	176.49897	31151.88641
10	44	78.72364	-34.72364	1205.731175
Jumlah				50427.3449

$$\begin{aligned}
 RMSE &= \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} \\
 &= \sqrt{\frac{50427.3449}{10}} \\
 &= 71.0122
 \end{aligned}$$

Hasil dari perhitungan standar error tergolong besar yaitu 71.0122. Dikatakan besar karena standar error yang didapat masih jauh mendekati kisaran 0.0 – 1.0.

Tujuan dari perhitungan yang akan dilakukan adalah:

1. Membuktikan perhitungan hasil keputusan yang telah dilakukan apakah sudah akurat atau masih terdapat kesalahan dalam proses perhitungan.
2. Membuktikan perhitungan apakah sudah benar atau masih terdapat error.