

Perbandingan Metode Best Subset Dan Stepwise Untuk Mengetahui Pengaruh Tingkat Pendidikan Terhadap Pengangguran Di Jawa Timur

Puji Subekti
STMIK Asia Malang
pujisubekti88@gmail.com

ABSTRAK. Pengangguran adalah salah satu permasalahan ekonomi yang banyak terjadi di Jawa Timur. Dewasa ini tidak sedikit dari penduduk di Jawa Timur dalam usia angkatan kerja yang tidak bekerja seperti pada umumnya. Jumlah penduduk kian hari semakin meningkat namun lapangan pekerjaan yang tersedia semakin sedikit. Data yang digunakan dalam laporan PKL ini adalah data persentase pengangguran serta persentase tingkat pendidikan terakhir di Propinsi Jawa Timur yang dirinci Menurut Kabupaten/Kotamadya. Berdasarkan hasil Analisis Regresi Linier Berganda menggunakan minitab menghasilkan hubungan yang dinyatakan dalam persamaan :

$$Y = 11,1 - 0,0538 X_1 - 0,224 X_2 - 0,0003 X_3 - 0,17 X_4 + 0,167 X_5 + 0,006 X_6 \text{ dengan } S = 1,73963 \quad R\text{-Sq} = 71,1 \% \\ R\text{-Sq}(\text{adj}) = 65,5 \%$$

Dari hasil tersebut ternyata terjadi multikolinieritas. Selanjutnya akan digunakan metode Regresi Stepwise dan Best Subset untuk mengatasi kasus multikolinieritas tersebut. Dari hasil Best Subset kita peroleh persamaan baru terpilih yaitu :

$$Y = 6,99 - 0,198 X_2 + 0,164 X_5, \text{ dengan } R\text{-Sq} = 69,2\%, R\text{-Sq}(\text{adj}) = 67,5\%, \\ C\text{-p} = 1,0 \text{ dan } S = 1,6892. \text{ Sedangkan dari hasil Stepwise kita peroleh persamaan baru terpilih yaitu :}$$

$$Y = 6,991 + 0,164 X_5 - 0,198 X_2, \text{ dengan } R\text{-Sq} = 69,25\%, R\text{-Sq}(\text{adj}) = 67,49\%, S = 1,69 \text{ dan } C_p = 1,0.$$

Pada hasil analisis proses menggunakan kedua metode diatas dapat diketahui bahwa Tingkat pendidikan SMA dan SD lebih mempengaruhi pengangguran. Sehingga diharapkan tiap-tiap sekolah khususnya di tingkat SMA/Sederajat serta SD lebih dapat meningkatkan kualitas pembelajarannya.

Kata Kunci: *Tingkat Pendidikan, Pengangguran, Best Subset, Stepwise*

1. PENDAHULUAN

Dewasa ini tidak sedikit dari penduduk di Jawa Timur dalam usia angkatan kerja yang tidak bekerja seperti pada umumnya. Jumlah penduduk kian hari semakin meningkat namun lapangan pekerjaan yang tersedia semakin sedikit.

Tambahan angkatan kerja baru tersebut mencapai 3 juta pekerja per tahun dan angkatan kerja baru tersebut tidak semua diserap oleh pasar tenaga kerja. Dengan naiknya harga - harga barang dan jasa maka daya beli masyarakat akan mengalami penurunan sehingga permintaan masyarakat akan barang dan jasa rendah. Hal ini dapat mengendorkan aktivitas produksi yang dialami perusahaan dapat menyebabkan semakin terbatasnya peluang kerja dan memaksa perusahaan untuk mengurangi jumlah pekerjanya dengan cara melakukan pemutusan hubungan kerja

Adapun untuk mengatasi hal tersebut melalui penulisan ini maka peneliti akan mengkaji “Perbandingan Metode Best Subset Dan Stepwise Untuk Mengetahui Pengaruh Tingkat Pendidikan Terhadap Pengangguran Di Jawa Timur”

2. LANDASAN TEORI

2.1 Pengertian Tenaga kerja

Menurut Dumairy (1996 : 74) pengertian tenaga kerja sebagai berikut : “tenaga kerja adalah penduduk yang berumur di dalam batas usia kerja “. Batasan usia kerja berbeda-beda antara negara yang satu dengan negara yang lain. Batas usia kerja yang dianut oleh Indonesia adalah minimum 10 tahun tanpa batas umur maksimum. Jadi setiap orang atau semua penduduk yang berusia 10 tahun tergolong sebagai tenaga kerja.

2.2 Pengertian pengangguran

Menurut Simanjuntak (1998 : 36) Pengangguran adalah orang yang tidak bekerja sama sekali atau bekerja kurang dari 2 hari selama satu minggu sebelum mencari dan berusaha memperoleh pekerjaan.

Tingkat pengangguran adalah perbandingan jumlah pengangguran dengan jumlah angkatan kerja dan dinyatakan dalam persen. Rumus tingkat pengangguran dapat dituliskan sebagai berikut :

$$\text{Tingkat pengangguran} = \frac{\text{Jumlah pengangguran}}{\text{angkatan kerja}} \times 100\%$$

Dalam mengelompokkan masing- masing pengangguran menurut simanjuntak (1998 : 43) perlu diperhatikan :

- a. Waktu (banyak diantara mereka yang ingin bekerja lebih lama, misalnya jam kerja per hari, per minggu per tahun).
- b. Intensitas pekerjaan (yang berkaitan dengan gizi makanan).
- c. Produktivitas (kurangnya sumber daya untuk melakukan pekerjaan).

2.3 Pendidikan

Pendidikan merupakan sebagian dari kehidupan masyarakat dan juga sebagai dinamostator masyarakat itu sendiri. Memang kita semua mengetahui betapa sektor pendidikan selalu berkembang dalam berbagai sector pembangunan lainnya bukan saja karena sektor konsumtif, juga karena pendidikan adalah penjaaga keamanan masyarakat itu sendiri yang bisa digunakan sebagai dasar mencari pekerjaan. Karena pekerjaan merupakan hal yang penting untuk mempertahankan kelangsungan hidup manusia.

Tujuan pendidikan merupakan salah satu komponen utama pada sistem pendidikan. Dengan tujuan pendidikan diharapkan proses pendidikan dapat mencapai hasil secara efektif dan efisien Apalagi kehidupan manusia dewasa ini telah mengglobal sehingga tidak bisa mengelak dari perubahan- perubahan yang selalu berubah.

2.4 Analisis Linear Berganda

Analisis regresi linear berganda adalah suatu metode yang digunakan untuk menganalisis antara satu variabel terikat dengan beberapa variabel bebas (Drapper, Smith, 1992). Dalam regresi berganda terdapat satu peubah tak bebas yang akan dilihat hubungannya dengan dua atau lebih peubah bebasnya, misalkan peubah Y dapat dinyatakan dengan fungsi-fungsi linier dari peubah $X_1, X_2, X_3, \dots, X_n$ yang diketahui dan terdapat faktor sisa. Model regresinya sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \epsilon_i \dots \dots \dots (2.1)$$

$i = 1, 2, 3 \dots n$

$Y_i =$ Nilai pengamatan yang ke- i

$X_1, \dots, X_n =$ Peubah bebas yang menentukan nilai pengamatan ke- i

$\beta_1 \dots \beta_p =$ Koefisien – koefisien regresi sebagian (parsial) untuk peubah X_1, \dots, X_p secara berturut-turut.

$\epsilon_i =$ faktor sisaan yang ke-i

$n =$ banyaknya pengamatan

2.5 Multikolinieritas

Dalam mendapatkan model regresi berganda terbaik maka persamaan regresi yang telah diperoleh terlebih dahuluharus diuji ada tidaknya multikolinieritas. Permasalahan yang sering dihadapi dalam menggunakan analisis regresi berganda adalah adanya multikolinieritas (ketergantungan antar variabel bebas) sehingga terdapat kesulitan untuk mengetahui pengaruh masing-masing variabel tidak bebas.

Melihat adanya multikolinieritas ada beberapa cara yakni:

- a. Dengan melihat elemen matrik korelasi. Jika korelasi antarvariabel bebas lebih besar daripada korelasi antara variabel bebas dengan variabel terikatnya maka terjadi multikolinieritas.
- b. Metode lain untuk mengetahui multikolinieritas adalah menghitung besarnya multikolinieritas tiap variabel bebas dengan vaktor keragaman inflasi (varian Inflation Faktor = VIF) didefinisikan sebagai berikut:

$$VIF = \frac{1}{1-R_j^2} \dots \dots \dots (2.2)$$

R_j^2 adalah koefisien determinasi berganda dari peubah X_j dengan semua peubah bebas yang lain. Apabila $VIF > 10$ maka korelasi diantara variabel bebas sangat tinggi, dengan kata lain terjadi multikolinieritas . Jika terjadi multikolinieritas maka cara terbaik untuk mengatasinya adalah dengan menambah data pengamatan. Jika tidak mungkin untuk menambah data maka salah satu cara mengatasinya adalah dengan menghilangkan variabel yang menyebabkan multikolinieritas.

2.6 Metode Best Subset

Regresi terbaik (best subset Regression) digunakan untuk meregresikan satu peubah respon pada semua kemungkinan kombinasi subset peubah-peubah prediktor dan kemudian memilih subset terbaik untuk setiap ukuran (size) informasi model terbaik ini dipilih berdasarkan nilai R-square terbesar. Pada setiap regresi subset terbaik ditampilkan statistik, yaitu : R-sq, adj R-sq, S dan C-p. Jika model difit tanpa konstanta, R-sq dan adj R-sq tidak ditampilkan (Draper dan Smith (1992)).

Metode regresi Best Subset ini merupakan salah satu cara untuk mengatasi multikolinieritas yaitu dengan memilih peubah-peubah bebas yang dapat digunakan dalam model agar diperoleh persamaan regresi terbaik yang mengandung sebagian atau seluruh peubah bebas.

Kriteria-kriteria yang dapat digunakan untuk menentukan variabel / peubah-peubah bebas yang digunakan sehingga diperoleh model terbaik untuk melakukan regresi dengan jumlah variabel yang dipilih yaitu :

1. Nilai R-Sq maximal
2. Nilai R-Sq (Adj) maximal
3. Nilai C-p terkecil
4. S terkecil

2.7 Regresi Stepwise

Regresi Stepwise merupakan sebuah metode yang dapat digunakan dalam menentukan model regresi berganda. Prosedur yang dilakukan yaitu dengan menyusupkan satu demi satu variabel bebas sampai diperoleh persamaan regresi yang memuaskan. Urutan penyisipan ditentukan dari koefisien korelasi parsial (r) (Draper & Smith, 1992). Jadi, variabel bebas yang mempunyai korelasi terbesar dengan variabel terikat akan masuk pertama kali dalam model.

Tahap selanjutnya adalah memasukkan variabel bebas yang memiliki korelasi terbesar kedua dengan variabel terikat. Tahap ini terus dilakukan sampai semua variabel masuk dalam model. Namun setelah diperoleh satu variabel bebas yang masuk dalam model perlu dilakukan pengujian keberartian variabel dalam model.

Hipotesis :

$H_0 : \beta_i = 0$, dimana variabel yang masuk dalam model tidak berarti

$H_1 : \beta_i \neq 0$, dimana variabel yang masuk dalam model berarti.

Untuk suatu $i = 1, 2, 3, \dots$

Pengujian dilakukan dengan *uji-t* atau melihat *p-value* dengan daerah penolakan

$$T_{hitung} > t_{tabel} (t_{(\alpha, n-k-1)}) \text{ atau } p\text{-value} < \alpha$$

Dimana α merupakan taraf nyata yang digunakan ($\alpha = 0.05$).

Apabila hasil pengujian ternyata menolak H_0 maka variabel tersebut tetap berada dalam model dan apabila sebaliknya maka variabel harus keluar dari model. Dalam regresi Stepwise, apabila ada dua variabel bebas yang saling berkorelasi, maka hanya ada satu variabel yang masuk dalam model. Pemilihan variabel yang masuk didasarkan pada koefisien korelasi terbesar terhadap variabel terikat (Iriawan, 2006).

3. METODE PENELITIAN

3.1 Metode Pengumpulan Data

Metode-metode yang kami gunakan untuk penyusunan penelitian ini adalah sebagai berikut:

1. Arsip BPS Kota Malang
Data tersebut berisi tentang informasi-informasi yang dibutuhkan dalam penyusunan penelitian ini. Data tersebut meliputi arsip Program Jawa Timur yang berisi tentang data-data statistik Jawa Timur.
2. Situs Badan Pusat Statistik
Penganblian informasi yang lainnya adalah melalui situs Badan Pusat Statistik. Dalam situs tersebut kita dapat memperoleh banyak informasi tentang BPS.
3. Literatur
Metode literatur yang mengacu pada studi kepustakaan sebagai bahan referensi yang berkaitan erat dengan penelitian ini.

3.2 Jenis Data

Data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder merupakan data yang diperoleh dari instansi yang bersangkutan, contoh dari data tersebut adalah dokumen-dokumen atau literatur yang ada di Badan Pusat Statistik Kota Malang. Data sekunder yang penulis gunakan adalah data jumlah penduduk dan tingkat pendidikan dalam satu tahun dan jumlah pengangguran dalam satu tahun. Presentase pengangguran di Provinsi Jawa Timur ditunjukkan pada Tabel 1. Sedangkan presentase jumlah penduduk berdasarkan tingkat pendidikan yang ditamatkan ditunjukkan pada Tabel 2.

Tabel 1. Presentase Pengangguran di Provinsi Jawa Timur

KOTA/KABUPATEN	PERSENTASE TINGKAT PENGANGGURAN PERKABUPATEN/KOTA
Pacitan	3,1
Ponorogo	3,73
Trenggalek	3,73
Tulungagung	3,99
Blitar	4,26
Kediri	7,79
Malang	6,22
Lumajang	4,33
Jember	4,48
Banyuwangi	5,62
Bondowoso	2,93
Situbondo	3,89
Probolinggo	3,46
Pasuruan	6,72
Sidoarjo	11,91
Mojokerto	7,04
Jombang	5,78
Nganjuk	6,06
Madiun	8,44
Magetan	4,37
Ngawi	8,18
Bojonegoro	5,93
Tuban	5,74
Lamongan	6,3
Gresik	7,5
Bangkalan	7,26
Sampang	3,46
Pamekasan	3,42
Sumenep	3,92
Kediri	11,27
Blitar	6,97
Malang	11,14
Probolinggo	9,93
Pasuruan	10,72
Mojokerto	12,12
Madiun	12,72
Surabaya	11,84
Batu	8,95

Tabel 2. Presentase Jumlah Penduduk Berdasarkan Tingkat Pendidikan Yang Ditamatkan

KOTA/KAB UPATEN	Tidak pernah sekolah (X1)	Tidak Tamat SD (X2)	Tamat SD/ sederajat (X3)	Tamat SLTP sederajat (X4)	Tamat SLTA/ sederajat (X5)	Tamat PT (X6)
Pacitan	12,44	17,56	32,77	18,96	14,34	3,92
Ponorogo	15,06	18,53	29,56	19,09	14,02	3,74
Trenggalek	5,84	18,36	36,71	22,29	13,76	3,04
Tulungagung	6,43	17,93	30,94	22,72	17,47	4,52
Blitar	7,77	20,76	31,30	20,99	15,41	3,76
Kediri	9,7	18,05	30,46	18,72	19,18	3,89

Malang	11,4	21,29	31,53	18,05	14,72	3,02
Lumajang	16,82	20,02	34,61	13,72	11,44	3,39
Jember	18,07	23,52	28,62	13,67	12,74	3,38
Banyuwangi	12,52	22,42	27,01	18,63	14,78	4,65
Bondowoso	21	26,13	28,43	12,52	8,98	2,94
Situbondo	24,46	23,18	24,36	12,85	11,01	4,13
Probolinggo	18,15	29,93	29,94	12,31	8,14	1,53
Pasuruan	12,09	21,74	31,05	17,89	14,53	2,71
Sidoarjo	3	8,45	22,21	22,33	35,84	8,18
Mojokerto	7,79	20,84	23,72	23,6	20,55	3,51
Jombang	8,46	17,41	27,17	23,32	18,7	4,94
Nganjuk	8,14	19,62	30,96	18,36	18,53	4,38
Madiun	11,03	16,77	27,92	20,43	19,13	4,71
Magetan	8,52	14,99	30,95	19,34	19,91	6,29
Ngawi	17,07	17,56	34,33	15,89	11,37	3,78
Bojonegoro	16,41	17,64	32,3	19,36	11,42	2,87
Tuban	17,83	14,43	36,47	16,96	11,3	3
Lamongan	12,23	18,65	25,33	21,09	18,39	4,31
Gresik	8	11,87	25,96	21,58	26,61	5,98
Bangkalan	32,5	22,13	26,04	7,49	9,22	2,63
Sampang	38,97	23,69	24,54	7,07	4,11	1,62
Pamekasan	20,25	19,65	29,87	14,38	12,47	3,39
Sumenep	29,58	17,42	26,78	13,56	10,07	2,60
Kediri	3,05	6,1	21,98	22,71	34,46	11,7
Blitar	2,31	10,82	18,63	22,76	31,12	14,37
Malang	3,82	8,32	19,26	16,32	33,42	18,86
Probolinggo	11,66	10,55	26,56	17,56	25,57	8,1
Pasuruan	5,1	12,79	25,27	18,23	31,2	7,4
Mojokerto	3,17	8,26	19,78	22,59	35,28	10,93
Madiun	4,3	5,52	18,58	21,65	34,84	15,11
Surabaya	3,83	6,39	19,47	21,05	36,28	12,99
Batu	3,45	13,3	28,9	22,96	24,24	1,52

Dari data-data tersebut dibuat suatu hubungan linier antara peubah respon yaitu peubah tak bebas dengan peubah-peubah bebas secara bersama-sama dengan analisis regresi linier berganda. Dalam hal ini presentase pengangguran kita tempatkan sebagai peubah respon Y (variabel terikat), sedangkan sebagai variabel X_i dengan $i = 1,2,3...6$ (variabel bebas) adalah tamatan sekolah yang terakhir yang terdiri dari :

X_1 : Tidak pernah sekolah

X_2 : Tidak tamat SD

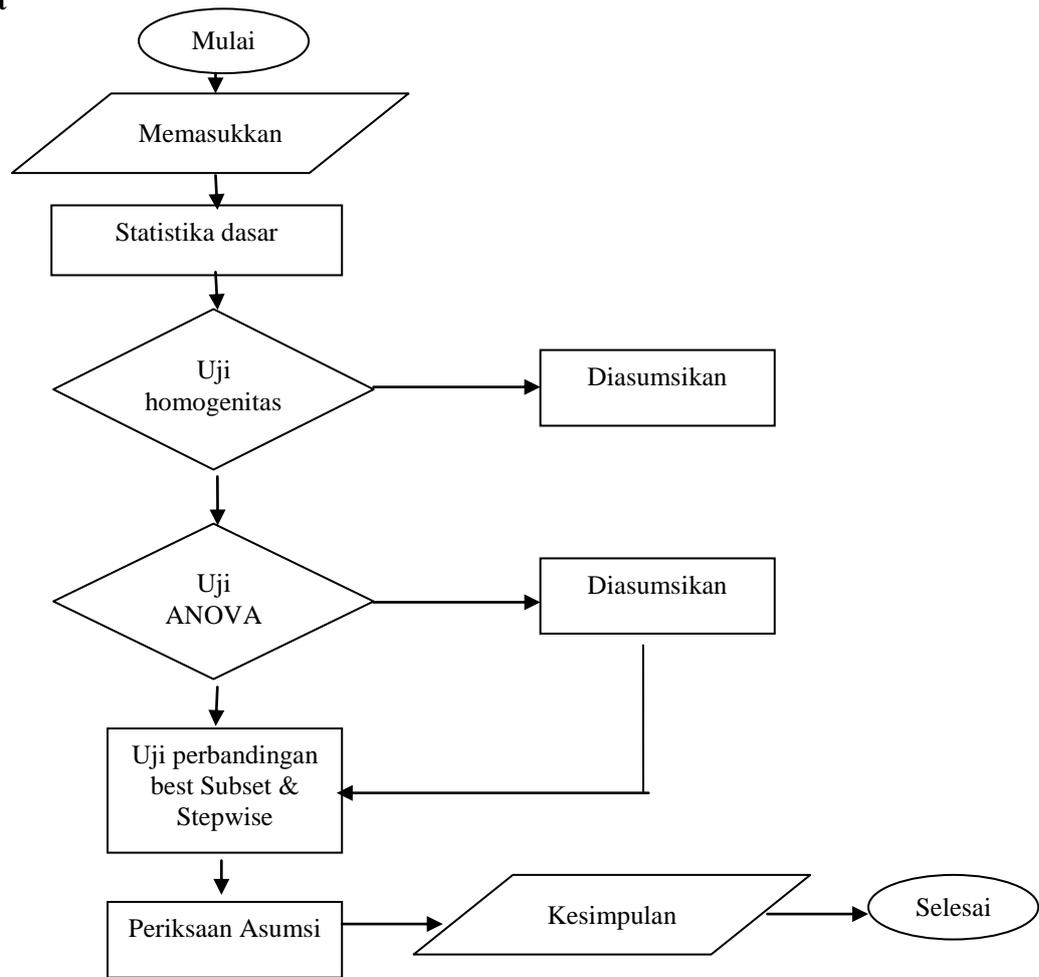
X_3 : Tamat SD

X_4 : Tamat SLTP/ MTS sederajat

X_5 : Tamat SLTA / MA sederajat

X_6 : Tamat PT

3.3 Flowchart



4. HASIL DAN PEMBAHASAN

4.1 Regresi Linier Berganda

Dari data yang diperoleh akan diregresikan menggunakan Minitab yaitu dengan melakukan analisis regresi berganda, dan menghasilkan model regresi sebagai berikut:

$$Y = 11,1 - 0,0538 X_1 - 0,224 X_2 - 0,0003X_3 - 0,17 X_4 + 0,167 X_5 + 0,006 X_6.$$

dengan $S = 1,73963$ $R-Sq = 71,1 \%$ $R-Sq (adj) = 65,5 \%$

Berdasarkan model regresi yang diperoleh dapat dikatakan bahwa pengaruh X_1 terhadap Y sebesar $-0,0538$, yang mempunyai arti bahwa orang yang tidak pernah bersekolah berpengaruh negatif atau berbanding terbalik dengan jumlah pengangguran sebesar $0,0538$. Semakin sedikit warga yang tidak bersekolah maka semakin banyak pengangguran di Jawa Timur. Demikian juga untuk warga yang mempunyai tamatan sedangkan pengaruh X_2 terhadap Y sebesar $-0,224$ dan pengaruh X_3 terhadap Y sebesar $-0,0003$.

Sebelum analisis ini dilakukan lebih lanjut maka perlu dilakukan pengujian terhadap asumsi klasik yang melandasi analisis regresi linier berganda. Pertama adalah pengujian kenormalan galat, hal ini bisa dilihat pada lampiran 4, dari grafik normal plot residual membentuk garis lurus dan sudutnya kurang dari 45° berarti ini sudah terpenuhi.

Dari hasil analisis regresi linier berganda terlihat bahwa peubah-peubah $X_1, X_2, X_3, X_4, X_5, X_6$ mempengaruhi jumlah pengangguran, hal ini dapat dilihat dari signifikansi $P = 0.000$ yang lebih kecil dari 0.05 dan nilai $F_{hitung} = 12,72$ yang lebih besar dari $F_{tabel} (F_{(0,05;6;31)})=24,094$. Dengan demikian dapat diambil kesimpulan bahwa tingkat pendidikan mempengaruhi tingkat pengangguran.

4.2 Metode Regresi Best Subset

Hasil Regresi best subset dari data diatas dapat dilihat pada Gambar 1 di bawah ini.

Vars	R-Sq	R-Sq(adj)	Mallows		X X X X X X						
			Cp	S	1	2	3	4	5	6	
1	63,9	62,9	4,7	1,8045							X
1	60,4	59,3	8,5	1,8903		X					
2	69,2	67,5	1,0	1,6892		X			X		
2	66,9	65,1	3,5	1,7514					X	X	
3	70,6	68,0	1,5	1,6751		X		X	X		
3	69,7	67,1	2,5	1,7003		X		X	X		
4	71,1	67,6	3,0	1,6862	X	X		X	X		
4	70,7	67,2	3,4	1,6972		X		X	X	X	
5	71,1	66,6	5,0	1,7122	X	X		X	X	X	
5	71,1	66,6	5,0	1,7123	X	X	X	X	X	X	
6	71,1	65,5	7,0	1,7396	X	X	X	X	X	X	X

Gambar 1. Hasil Regresi Best Subset

Dari hasil tersebut didapatkan 11 model regresi yang akan dipilih yang terbaik. Sesuai dengan kriteria-kriteria yang telah ditentukan maka model regresi terbaik yang akan gunakan adalah model regresi nomor 3 dan 5 , yaitu pada model regresi no 3 yang menggunakan 2 variabel yaitu X_2 dan X_5 dengan $R-Sq=69,2\%$, $R-Sq(adj) = 67,5\%$, $C-p = 1,0$ dan $S=1,6892$. Disamping itu juga perhatikan pada model regresi no 5 yang menggunakan variabel X_2 , X_4 dan X_5 dengan $R-Sq=70,6\%$, $R-Sq(adj)=68,0\%$, $C-p=1,5$ dan $S=1,6751$. Kemudian regresikan kembali dengan menggunakan variabel-variabel terpilih sehingga pada model regresi nomor 3 X_1, X_3, X_4 , X_6 dihilangkan. Sehingga diperoleh hasil analisis regresi linear berganda yang baru sebagai berikut : $Pengangguran (Y) = 6,99 - 0,198 X_2 + 0,164 X_5$.

Melalui hasil regresi yang didapatkan bahwa variabel X_2 mempengaruhi Y sebesar $-0,198$ dan variabel X_5 mempengaruhi Y sebesar $0,164$. Hasil yang didapatkan dari persamaan linier di atas menunjukkan bahwa semakin besar jumlah penduduk yang berpendidikan tidak tamat SD jumlah pengangguran semakin kecil dan sebaliknya. Sebaliknya ini bertentangan dengan jumlah penduduk dengan tamatan SLTA hal ini diperkirakan karena penduduk yang berpendidikan lebih rendah bersedia bekerja apapun tanpa memandang berat atau tidaknya suatu pekerjaan karena dalam pikiran mereka adalah mencari pekerjaan dengan tidak mempunyai ijazah yang tinggi adalah sulit. Sebaliknya bagi penduduk yang mempunyai ijazah yang lebih tinggi misalnya penduduk dengan tamatan SLTA merasa dirinya harus mendapatkan pekerjaan yang sesuai dengan tingkat pendidikannya, sehingga bukanlah hal yang mustahil didapatkan hasil persamaan regresi seperti yang tampak di atas yang tidak sesuai dengan harapan.

Pada model regresi yang no 5 X_1 , X_3 , X_6 dihilangkan. Sehingga diperoleh hasil analisis regresi linear berganda yang baru yaitu $Y = 8,67 - 0,214 X_2 - 0,104 X_4 + 0,190 X_5$. Dari hasil analisis regresi linier berganda yang baru terlihat jelas lebih baik dari yang pertama . Karena selain $R-Sq (Adj)$ meningkat variabel yang signifikan lebih bannyak dan tidak menunjukkan tanda-tanda multikolinearitas. Pengaruh variabel X_4 sangatlah kecil .

4.3 Metode Regresi Stepwise

Hasil analisis menggunakan stepwise dapat dilihat pada Gambar 2.

Vars	R-Sq	R-Sq(adj)	Mallows		X X X X X X						
			Cp	S	1	2	3	4	5	6	
1	63,9	62,9	4,7	1,8045							X
1	60,4	59,3	8,5	1,8903		X					
2	69,2	67,5	1,0	1,6892		X			X		
2	66,9	65,1	3,5	1,7514					X	X	
3	70,6	68,0	1,5	1,6751		X		X	X		
3	69,7	67,1	2,5	1,7003		X		X	X		
4	71,1	67,6	3,0	1,6862	X	X		X	X		
4	70,7	67,2	3,4	1,6972		X		X	X	X	
5	71,1	66,6	5,0	1,7122	X	X		X	X	X	
5	71,1	66,6	5,0	1,7123	X	X	X	X	X	X	
6	71,1	65,5	7,0	1,7396	X	X	X	X	X	X	X

Gambar 2. Hasil Regresi Stepwise

Pada tahap satu diperoleh variabel yaitu X_5 . Pada tahap ke dua variabel bebas yang memiliki korelasi terbesar dengan variabel terikat adalah X_5 , sedangkan variabel bebas yang memiliki korelasi terbesar kedua dengan variabel respon adalah X_2 . Karena variabel bebas X_5 dan X_2 saling berkorelasi, maka variabel yang masuk dalam model adalah variabel X_2 dan X_5 . Sehingga diperoleh persamaan regresi yang baru

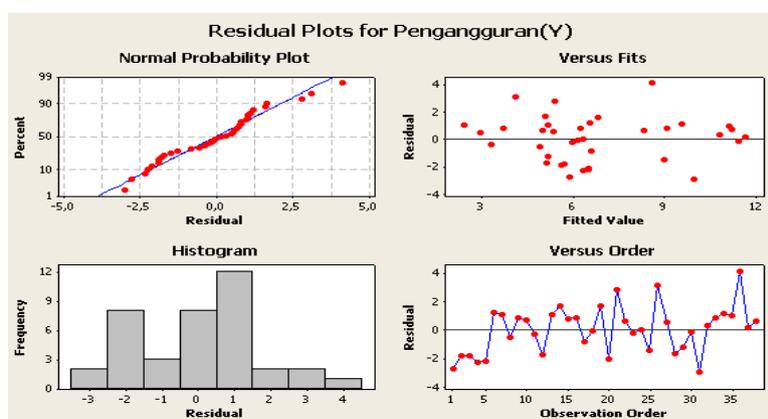
$Y = 6,991 + 0,164X_5 - 0,198X_2$ dengan $R\text{-Sq} = 69,25\%$, $R\text{-Sq}(\text{adj}) = 67,49\%$, $S = 1,69$ dan $C_p = 1,0$.

Dari analisis data diatas dapat disimpulkan bahwa pengaruh peubah X_5 dan X_2 secara berturut-turut terhadap peubah Y sebesar 0,164 dan -0,198, dengan kata lain yang paling berpengaruh terhadap kemiskinan adalah penduduk yang memiliki pendidikan SLTA/ sederajat dan tidak tamat SD. Karena nilai $R\text{-Sq}$ cukup besar, maka perlu mencoba dengan metode lain yang bisa mengatasi kasus multikolinieritas yaitu dengan metode Best Subset seperti yang telah dilakukan sebelumnya

4.4 Model

Dari hasil analisis yang didapat maka dapat diperoleh persamaan model regresi yang baru pada hasil regresi stepwise dan Best Subset yaitu $Y = 6,991 + 0,164X_5 - 0,198X_2 + \varepsilon$ dimana dapat dilihat bahwa pengaruh penduduk usia 15 tahun keatas tamatan SLTA sederajat sebesar 0,164 dan pengaruh penduduk usia 15 tahun keatas tidak tamat SD sebesar -0,198.

Karena korelasi antar variabel bebas tidak ada yang lebih tinggi dari korelasi variabel bebas dengan peubah respon, maka persamaan yang dihasilkan sama dengan menggunakan regresi linier berganda, dengan F_{hitung} sebesar 39,41 sehingga F_{hitung} jelas lebih besar dari F_{tabel} ($F_{(0.05;2;35)} = 3,267$) dan T_{hitung} sebesar 2,47 dan $3,18 < T_{\text{hitung}}$ sebesar 2,0261 atau nilai $P < \alpha$ yaitu $0,000 < 0,05$, ini berarti bahwa data sudah signifikan. Maka menolak H_0 dan dapat dikatakan bahwa pengaruh variabel bebas berarti terhadap model. Hasil kenormalan, kehomogenan dan independen data ditunjukkan pada Gambar 3.



Gambar 3. Hasil Kenormalan, Kehomogenan dan Independen Data

Dari hasil yang telah didapat dari uji asumsi residual yang terdiri atas kenormalan, kehomogenan dan independen seperti yang terlihat pada lampiran 8. Hasil telah menunjukkan adanya kenormalan data yang dalam hal ini dapat diketahui dari grafik Normal Probability Plot dengan hasil plot mendekati garis lurus dengan gradien 45° P-Value sebesar $0,505 > 0,05$ yang mengikuti sebaran normal. Pada uji independen hasil melalui Fitted value dan observation order pada lampiran 8 terlihat bahwa data telah bersifat menyebar dan tidak ada yang melampaui dari batas.

Dari beberapa uji yang telah dilakukan diatas maka hasil interpretasi model memenuhi syarat sehingga dapat digunakan.

4.5 Interpretasi Hasil terbaik

Dari kedua model prediksi tersebut model terbaik yang dihasilkan melalui metode regresi Best Subset dan Stepwise mampu menjelaskan keragaman data yang sama yaitu sebesar 69,25%. Kedua metode Best Subset dan Stepwise dapat digunakan untuk menjawab masalah pada laporan ini dan menghasilkan model yang sama yaitu $Y = 6,99 - 0,198 X_2 + 0,164 X_5$. Dimana melalui hasil regresi yang didapatkan bahwa variabel X_2 mempengaruhi Y sebesar -0,198 dan variabel X_5 mempengaruhi Y sebesar 0,164. Hasil yang didapatkan dari persamaan linier di atas menunjukkan bahwa semakin besar jumlah penduduk yang berpendidikan tidak tamat SD jumlah pengangguran semakin kecil dan sebaliknya. Sebaliknya bertentangan dengan jumlah penduduk dengan tamatan SLTA hal ini diperkirakan karena penduduk yang berpendidikan lebih rendah bersedia bekerja apapun tanpa memandang berat atau tidaknya suatu pekerjaan karena dalam pikiran mereka adalah mencari pekerjaan dengan tidak mempunyai ijazah yang tinggi adalah sulit. Sebaliknya bagi penduduk yang mempunyai ijazah yang lebih tinggi misal SLTA merasa dirinya harus mendapatkan pekerjaan yang sesuai dengan tingkat pendidikannya, sehingga bukanlah hal yang mustahil didapatkan hasil persamaan regresi seperti yang tampak di atas yang tidak sesuai dengan harapan. Sehingga

diharapkan tiap-tiap sekolah khususnya di tingkat SMA/Sederajat lebih dapat meningkatkan mutu kualitas pembelajarannya dan melaksanakan wajib belajar 9 tahun.

5. KESIMPULAN

Berdasarkan hasil dari pembahasan, maka kesimpulan yang dapat kita ambil adalah sebagai berikut :

1. Dari data hubungan antara presentase pengangguran dengan tingkat pendidikan terakhir ternyata terjadi multikolinieritas, oleh karena itu untuk mengatasinya dapat diterapkan Metode Best Subset dan Regresi Stepwise sehingga didapatkan model regresi terbaik yaitu
$$Y = 6,991 + 0,164X_5 - 0,198X_2$$
 dengan $R\text{-Sq} = 69,25\%$, $R\text{-Sq}(\text{adj}) = 67,49\%$, $S = 1,69$ dan $C_p = 1,0$.
2. Dari model regresi yang terpilih diketahui bahwa tingkat pendidikan terakhir yang sangat mempengaruhi presentase tingkat pengangguran adalah tamatan SLTA / sederajat dan penduduk tidak tamat SD karena nilai $R\text{-Sq}$ yang sangat besar.

DAFTAR PUSTAKA

- [1] Draper N. and Smith. H. 1992. *Analisis Regresi Terapan*. Jakarta : Gramedia Pustaka.
- [2] Dumairy, 1996, *Perekonomian Indonesia*, Erlangga Jakarta.
- [3] Iriawan, Nur. 2006. *Mengolah Data Statistik dengan Mudah Menggunakan Minitab 14*. Yogyakarta : ANDI.Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia.
- [5] Simanjuntak, Payaman J,1998. *Pengantar Ekonomi Sumber Daya Manusia*. Jakarta: Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia.