ANALISIS KEMISKINAN DAN KERENTANAN KEMISKINAN DENGAN REGRESI RIDGE, LASSO, DAN *ELASTIC-NET* DI PROVINSI JAWA TENGAH TAHUN 2017

(*POVERTY AND VULNERABILITY TO POVERTY ANALYSIS WITH RIDGE REGRESSION, LASSO, AND ELASTIC-NET IN CENTRAL JAVA PROVINCE 2017*)

Guntur Wahyu Kusuma¹, Ika Yuni Wulansari²

¹Politeknik Statistika STIS

²Politeknik Statistika STIS

Jalan Otto Iskandardinata No. 64C, Jakarta Timur 13330, Indonesia

E-mail: 15.8640@stis.ac.id

ABSTRAK

Kemiskinan bukanlah kondisi yang statis melainkan kondisi yang dinamis. Seseorang yang dianggap miskin saat ini bisa saja tidak miskin di masa mendatang, begitu juga sebaliknya. Berdasarkan metode BPS, pada tahun 2017 Provinsi Jawa Tengah memiliki persentase penduduk rentan miskin tertinggi di Pulau Jawa. Perhitungan kerentanan kemiskinan oleh BPS belum bisa menjelaskan kemiskinan yang dinamis, metode yang dapat menjelaskan adalah Vulnerability as Expected Poverty (VEP). Analisis kemiskinan dan kerentanan kemiskinan bersifat multidimensi sehingga digunakan Regresi Ridge, LASSO, dan Elastic-net yang dapat mengatasi multikolinieritas yang sering muncul pada high dimensional data dan menghasilkan model yang parsimoni. Penelitian ini bertujuan untuk menduga persentase rumah tangga miskin maupun rentan miskin dengan metode VEP dan mengetahui variabel-variabel yang memengaruhi kemiskinan dan kerentanan kemiskinan di Provinsi Jawa Tengah tahun 2017 menggunakan model terbaik yang terpilih dari Regresi Ridge, LASSO, dan Elastic-net. Hasil penelitian menunjukkan bahwa 10,4 persen rumah tangga di Provinsi Jawa Tengah miskin dan 4,8 persen rumah tangga rentan miskin. Model terbaik untuk kemiskinan dan kerentanan kemiskinan adalah Elastic-net. Variabel-variabel yang signifikan memengaruhi kemiskinan adalah AHH, pertanian, pengeluaran non makanan, jenis lantai, jenis kloset, dan upah minimum. Sedangkan variabel yang signifikan memengaruhi kerentanan kemiskinan adalah RLS, angka morbiditas, pengeluaran makanan, dan jenis kloset.

Kata kunci: kerentanan kemiskinan, VEP, regresi ridge, LASSO, elastic-net

ABSTRACT

Poverty is not a static condition but a dynamic condition. Someone who is considered poor today may not be poor in the future, and vice versa. Based on the BPS method, in 2017 Central Java Province has the highest percentage of poor vulnerable population on Java. Calculation of vulnerability to poverty by BPS has not been able to explain dynamic poverty, the method that can explain is Vulnerability as Expected Poverty (VEP). Poverty and vulnerability to poverty analysis is multidimensional so Ridge Regression, LASSO, and elastic-net are used which can overcome multicollinearity that often appears in high dimensional data and produce parsimony models. This study aims to estimate the percentage of poor and vulnerable households with the VEP method and find out the variables that affect poverty and vulnerability to poverty in Central Java Province in 2017 using the best models selected from Ridge Regression, LASSO, and elastic-net. The results showed that 10,4 percent of households in Central Java Province were poor and 4,8 percent of households were vulnerable to poor. The best model for poverty and vulnerability to poverty is Elastic-net. The variables that significantly influence poverty are AHH, agriculture, non-food expenditure, type of floor, toilet type, and minimum wage. While the variables that significantly influence vulnerability to poverty are RLS, food expenditure, morbidity rate and type of toilet.

Keywords: vulnerability to poverty, VEP, ridge regression, LASSO, elastic-net

PENDAHULUAN

Pengentasan kemiskinan telah menjadi salah satu target dalam *Sustainable Development Goals* (SDGs), yaitu pada tujuan pertama "*end poverty in all its forms everywhere*" yang berarti mengakhiri segala bentuk kemiskinan di semua tempat pada tahun 2030. Di Indonesia sendiri pengentasan

kemiskinan secara rinci dijabarkan dalam sasaran pokok Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2015-2019 yaitu pada sasaran dimensi pemerataan, dengan target mengurangi tingkat kemiskinan dari 11,25 persen pada tahun 2014 menjadi 5 – 6 persen pada tahun 2019. Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan di bawah garis kemiskinan (BPS, 2017). Hasil pengukuran kemiskinan dari BPS menunjukkan bahwa kemiskinan di Indonesia dari tahun ke tahun mengalami penurunan atau tren yang membaik. Menurut Haughton dan Khandker (2009), Perubahan persentase penduduk miskin bukanlah perubahan satu arah. Hal ini yang menyebabkan strategi penanggulangan kemiskinan masih belum maksimal. Perubahan persentase penduduk miskin penuh dengan pergerakan keluar dan masuk kemiskinan.

Pergerakan keluar dan masuk kemiskinan tersebut menunjukkan bahwa kemiskinan bukanlah kondisi yang statis. Namun, kemiskinan adalah kondisi yang dinamis yang selalu bergerak, bisa bergerak ke atas ke arah yang lebih baik atau justru ke bawah ke arah yang lebih buruk. Hal ini sejalan dengan Dariwardani (2014), yang menyatakan bahwa kemiskinan bersifat dinamis. Artinya, seseorang yang dianggap miskin pada saat ini bisa saja tidak miskin di masa mendatang. Begitu juga sebaliknya, seseorang yang tidak miskin bisa saja menjadi miskin di masa mendatang. Hal ini disebabkan oleh banyaknya kelompok masyarakat yang memiliki pendapatan di sekitar garis kemiskinan. Akibatnya, apabila terjadi sedikit saja guncangan dalam perekonomian dapat membuat kelompok masyarakat tersebut menjadi tergolong miskin. Sehingga dibutuhkan adanya kebijakan yang tepat sasaran demi pengentasan kemiskinan yang maksimal.

Jumlah penduduk miskin antarprovinsi di Indonesia masih sangat beragam. Jika dilihat berdasarkan pulau, pada tahun 2017 Pulau Jawa memiliki jumlah penduduk miskin terbesar daripada pulau lainnya yaitu sebesar 14,79 juta orang. Jika dilihat lebih lanjut lagi, pada tahun 2017 terdapat tiga provinsi di Pulau Jawa yang memiliki persentase penduduk miskin lebih dari angka persentase penduduk miskin nasional yang menurut Badan Pusat Statistik sebesar 10,64 persen. Ketiga provinsi tersebut adalah Daerah Istimewa Yogyakarta dengan persentase penduduk miskin sebesar 13,02 persen, Jawa Tengah sebesar 13,01 persen, dan Jawa Timur sebesar 11,77 persen. Kemudian, berdasarkan metode pengukuran rentan miskin BPS, dimana dikategorikan rentan miskin bila garis kemiskinan $\leq$ pengeluaran per kapita $\leq$ 1,6 garis kemiskinan menunjukkan bahwa dari tahun 2015-2017 persentase penduduk rentan miskin Provinsi Jawa Tengah selalu menempati posisi pertama di Pulau Jawa (30,498 persen, 29,527 persen, dan 26,469 persen) dan selalu berada di atas persentase penduduk rentan miskin nasional (26,832 persen, 24,927 persen, dan 23,219 persen). Penduduk rentan miskin ini sangat sensitif terhadap gejolak dan guncangan ekonomi. Menurut Kumala dan Agustini (2011), adanya guncangan dalam perekonomian, pekerjaan kepala rumah tangga, ataupun kondisi yang merugikan lainnya, dapat mengakibatkan penduduk menjadi miskin pada periode tertentu.

Perhitungan rentan miskin BPS belum bisa menjelaskan kemiskinan yang dinamis. Salah satu metode yang dapat menghitung peluang setiap rumah tangga untuk menjadi miskin pada masa mendatang adalah metode *Vulnerability as Expected Poverty* (VEP), metode yang dikembangkan oleh Chaudhuri, Jalan, dan Suryahadi (2002). Metode tersebut dimulai dengan membangun model pengeluaran per kapita rumah tangga. Dimana, perhitungan peluang setiap rumah tangga menjadi miskin pada masa mendatang memerlukan beberapa variabel yang memiliki hubungan dengan pengeluaran per kapita rumah tangga. Kemiskinan bersifat multidimensi, artinya masalah kemiskinan berkaitan dan dipengaruhi oleh banyak variabel. Banyaknya variabel yang diduga memengaruhi kemiskinan ini membawa pada konsep *high dimensional data*. Dimana saat melakukan analisis pada *high dimensional data* yang melibatkan banyak variabel independen sering terjadi pelanggaran asumsi non multikolinieritas. Selain itu banyaknya variabel independen dalam suatu model juga tidak baik karena akan melanggar prinsip parsimoni yang menyatakan bahwa semakin sederhana suatu model, semakin baik model tersebut.

Kemudian, terkait dengan kebijakan dalam upaya pengentasan kemiskinan dan upaya preventif untuk mencegah kemiskinan yang dipengaruhi oleh banyak variabel. Perlu dibuat prioritas berdasarkan variabel mana yang mempunyai pengaruh signifikan terhadap kemiskinan dan kerentanan kemiskinan tersebut. Sehingga perlu dilakukan seleksi untuk memperoleh variabel-variabel yang mempunyai pengaruh signifikan dari sekian banyak variabel yang diduga

memengaruhi kemiskinan dan kerentanan kemiskinan. Beberapa metode yang dapat digunakan untuk mengatasi masalah multikolinieritas dan dapat melakukan seleksi variabel sehingga didapat model yang parsimoni adalah Regresi Ridge, LASSO, dan *Elastic-net*. Oleh karena itu, untuk dapat mengetahui variabel-variabel yang memengaruhi kemiskinan dan kerentanan kemiskinan pada *high dimensional data* digunakan Regresi Ridge, LASSO, dan *Elastic-net*.

Berdasarkan uraian di atas, maka dalam penelitian ini akan dilakukan pendugaan persentase rumah tangga miskin maupun rentan miskin pada level provinsi dan kabupaten/kota di Provinsi Jawa Tengah tahun 2017 dengan metode VEP. Kemudian mengetahui model terbaik dari Regresi Ridge, LASSO, dan *Elastic-net* dalam analisis variabel-variabel yang memengaruhi kemiskinan dan kerentanan kemiskinan, selanjutnya menggunakan model terbaik yang terpilih mengetahui variabel-variabel yang memengaruhi kemiskinan dan kerentanan kemiskinan di Provinsi Jawa Tengah tahun 2017.

METODE

Data yang digunakan pada penelitian ini adalah data sekunder yang merupakan *raw data* hasil Survei Sosial Ekonomi Nasional (Susenas) KOR dan Modul bulan Maret tahun 2017 dan beberapa publikasi yang bersumber dari *website* Badan Pusat Statistik (BPS). Variabel yang digunakan dari *raw data* Susenas KOR dan Modul bulan Maret tahun 2017 di provinsi Jawa Tengah adalah data pada unit rumah tangga, yaitu pengeluaran perkapita rumah tangga, jumlah ART, status perkawinan KRT, jenis kelamin KRT, umur KRT, ijazah tertinggi KRT, lapangan usaha utama KRT, luas lantai, jenis lantai, jenis kloset, cara memperoleh air minum, dan bahan bakar memasak. Sedangkan variabel yang digunakan dari *website* BPS Jawa Tengah adalah rata-rata lama sekolah, angka partisipasi sekolah 16-18 tahun, angka morbiditas, angka harapan hidup, TPT, TPAK, pengeluaran makanan, pengeluaran bukan makanan, jumlah penduduk, pertumbuhan ekonomi, inflasi, dan upah minimum.

Metode analisis yang digunakan pada penelitian ini adalah *Vulnerability as Expected Poverty* (VEP) yang digunakan untuk pengukuran kerentanan kemiskinan rumah tangga sehingga diperoleh kelompok rumah tangga menurut status kemiskinannya. Kemudian, Regresi Linier Berganda, Regresi Ridge, LASSO, dan *Elastic-net* yang digunakan untuk mengetahui variabel-variabel yang memengaruhi kemiskinan dan kerentanan kemiskinan.

VULNERABILITY AS EXPECTED POVERTY (VEP)

Identifikasi status kemiskinan rumah tangga dengan menggunakan metode VEP diawali dengan membangun model pengeluaran per kapita rumah tangga sebagai berikut:

$$\ln c_h = X_h \beta + \varepsilon_h \quad (1)$$

dimana c_h adalah pengeluaran per kapita rumah tangga h , X_h adalah vektor dari karakteristik rumah tangga yang diobservasi, β adalah vektor parameter, dan ε_h *disturbance term* atau gangguan yang juga mencakup faktor idiosinkratik (*shock* atau guncangan) yang berkontribusi dalam perbedaan pengeluaran per kapita dari rumah tangga yang memiliki karakteristik yang sama. Pada metode VEP, varians dari residual ε_h pada persamaan 1 dibiarkan terdapat heteroskedastisitas dan diasumsikan bergantung atau dipengaruhi oleh karakteristik rumah tangga yang diobservasi. Oleh karena itu, varians dari residual dimodelkan sebagai berikut:

$$\sigma_{\varepsilon,h}^2 = X_h \theta \quad (2)$$

perkiraan parameter β dan θ pada persamaan 1 dan 2 diestimasi menggunakan prosedur *Feasible Generalized Least Squares* (FGLS) tiga tahap.

Tahap pertama FGLS adalah mengestimasi residual dari persamaan 3 menggunakan prosedur *Ordinary Least Squares* (OLS). Selanjutnya, membuat model residual kuadrat yang digunakan untuk menghitung varians dari pengeluaran per kapita setiap rumah tangga sebagai berikut:

$$\hat{\varepsilon}_{OLS,h}^2 = X_h \theta + \eta_h \quad (3)$$

Tahap kedua FGLS yaitu menggunakan hasil estimasi persamaan 3 untuk mentransformasi persamaan 3 itu sendiri sehingga diperoleh penduga θ yang efisien secara asimtot sebagai berikut:

$$\frac{\hat{\varepsilon}_{OLS,h}^2}{X_h \hat{\theta}_{OLS}} = \left(\frac{X_h}{X_h \hat{\theta}_{OLS}} \right) \theta + \frac{\eta_h}{X_h \hat{\theta}_{OLS}} \quad (4)$$

Tahap ketiga FGLS yaitu mentransformasi persamaan 1 dengan standar deviasi dari komponen idiosinkratik untuk pengeluaran per kapita rumah tangga sebagai berikut:

$$\frac{\ln c_h}{\sqrt{X_h \hat{\theta}_{FGLS}}} = \left(\frac{X_h}{\sqrt{X_h \hat{\theta}_{FGLS}}} \right) \beta + \frac{\varepsilon_h}{\sqrt{X_h \hat{\theta}_{FGLS}}} \dots \dots \dots (5)$$

didapatkan penduga β dan θ yang terbaik, maka nilai harapan dari log pengeluaran per kapita rumah tangga h adalah:

$$\hat{E}[\ln c_h | X_h] = X_h \hat{\beta}_{FGLS} \dots \dots \dots (6)$$

kemudian nilai varians dari log pengeluaran per kapita rumah tangga h diperoleh adalah:

$$\hat{V}[\ln c_h | X_h] = \hat{\sigma}_{\varepsilon, h}^2 = X_h \hat{\theta}_{FGLS} \dots \dots \dots (7)$$

Selanjutnya nilai estimasi pada persamaan 8 dan 9 digunakan untuk menghitung nilai peluang suatu rumah tangga untuk menjadi miskin dengan formula yaitu:

$$\hat{V} = \Phi \left[\frac{\ln z - \ln X_h \hat{\beta}_{FGLS}}{\sqrt{X_h \hat{\theta}_{FGLS}}} \right] \dots \dots \dots (8)$$

dimana Φ kerapatan kumulatif standar normal dan z adalah garis kemiskinan pada waktu t .

Setelah diperoleh nilai peluang setiap rumah tangga untuk jatuh miskin pada masa mendatang, selanjutnya setiap rumah tangga dapat ditentukan status kemiskinannya. Untuk menentukan status kemiskinan suatu rumah tangga dibutuhkan suatu titik potong (*cut off point*) atau bisa juga disebut *threshold*. Dalam penelitian ini ditetapkan *threshold* yaitu angka kemiskinan Provinsi Jawa Tengah tahun 2017 yaitu sebesar 0,1301.

Tabel 1. Pengkategorian status kemiskinan rumah tangga.

		Pengeluaran per kapita rumah tangga saa ini (c_t)			
		$c_t < z$	$c_t \geq z$		
		A	D		
Peluang rentan miskin rumah tangga (VEP_t)	$VEP_t \geq 0,1301$	B	E	$E[c_t] < z$	Prediksi pengeluaran per kapita rumah tangga $E[c_t]$
				$E[c_t] \geq z$	
	$VEP_t < 0,1301$	C	F		

Sumber: Suryahadi dan Sumanto, 2001

Berdasarkan **Tabel 1.** pengelompokkan status kerentanan kemiskinan rumah tangga yaitu:

- *Poor* = A + B + C
- *Chronic poor* = A
- *Transient poor* = B + C
- *Non poor* = D + E + F
- *High vulnerability non poor* = D +
- *High vulnerability group* = A + B + D + E
- *Low level of consumption* = A + D
- *High variability of consumption* = B + E
- *Low vulnerability group* = C + F
- *Total vulnerable group* = A + B + C + D + E

Dimana z adalah garis kemiskinan dan *high vulnerability non poor* didefinisikan sebagai rentan miskin.

REGRESI RIDGE

Regresi Ridge adalah suatu teknik yang dikembangkan untuk menstabilkan nilai koefisien regresi karena adanya multikolinieritas. Metode Regresi Ridge pertama kali diperkenalkan oleh A. E. Hoerl pada tahun 1962. Regresi Ridge merupakan modifikasi dari metode OLS yang menghasilkan penduga bias dari koefisien regresi (Kutner, et al, 2005). Regresi Ridge mengurangi dampak multikolinieritas dengan menentukan penduga yang bias tetapi mempunyai varians yang lebih kecil dari varians penduga OLS. Menurut Ohlyver (2011), meskipun metode ini menghasilkan penduga koefisien regresi bias, penduga ini bisa mendekati nilai parameter yang sebenarnya. Hal ini dapat

diketahui dari perbandingan *mean square error* (MSE) antara penduga Ridge dengan penduga kuadrat terkecil (*least square*), dimana MSE penduga Ridge lebih kecil daripada MSE penduga OLS.

Regresi Ridge serupa dengan OLS yang meminimumkan *sum of squares error* (SSE) atau jumlah kuadrat sisaan pada pendugaan koefisien regresi. Namun, dalam metode ini menambahkan penalti penyusutan dalam meminimumkan SSE. Sehingga persamaan yang terbentuk adalah sebagai berikut:

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p \beta_j^2 \dots \dots \dots (9)$$

Dimana $\lambda \geq 0$ adalah parameter penyusutan dan $\sum_{j=1}^p \beta_j^2$ adalah penalti penyusutan. Jika $\lambda=0$, maka penalti penyusutan tidak memberikan pengaruh apapun sehingga regresi ridge akan menghasilkan penduga koefisien regresi yang sama dengan OLS. Namun, jika $\lambda \rightarrow \infty$ akan berdampak pada penalti penyusutan yang semakin besar dan koefisien pendugaan yang semakin mendekati nol. Pada OLS hanya dihasilkan satu set dugaan koefisien, sedangkan pada regresi ridge dihasilkan set dugaan koefisien berbeda-beda untuk setiap nilai λ . Parameter penyusutan (λ) yang optimal ditentukan dengan metode validasi silang.

LASSO (LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR)

LASSO (*Least Absolute Shrinkage and Selection Operator*) adalah salah satu metode penyusutan seperti regresi ridge yang dapat mengatasi permasalahan multikolinearitas yang diperkenalkan oleh Tibshirani tahun 1996. Dalam regresi ridge, penduga koefisien regresi disusutkan ke arah nol seiring dengan peningkatan nilai λ . Satu hal yang tidak dapat dilakukan oleh regresi ridge adalah melakukan seleksi peubah secara otomatis.

Menurut Tibshirani (1996), dua teknik standar untuk meningkatkan pendugaan OLS adalah pemilihan subset (*subset selection*) dan regresi ridge, tapi keduanya memiliki kelemahan. Pemilihan subset menyediakan model yang dapat diinterpretasikan tetapi bisa sangat bervariasi karena merupakan proses diskrit yang dipertahankan atau dikeluarkan dari model. Perubahan kecil dalam data dapat menghasilkan model terpilih yang sangat berbeda dan ini dapat mengurangi akurasi prediksi. Regresi Ridge adalah proses kontinyu yang mengecilkan koefisien dan karenanya lebih stabil. Namun, tidak menetapkan koefisien apapun menjadi 0 dan karenanya tidak memberikan model yang mudah diinterpretasikan. Sedangkan LASSO dapat menyusutkan beberapa koefisien dan menetapkan yang lain ke 0. Oleh karena itu LASSO dapat mempertahankan fitur-fitur yang baik dari pemilihan subset dan regresi ridge.

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \dots \dots \dots (10)$$

LASSO dan Regresi Ridge memiliki persamaan yang hampir sama, hanya penaltinya saja yang berbeda. Penalti penyusutan pada regresi ridge $\sum_{j=1}^p \beta_j^2$ diganti dengan $\sum_{j=1}^p |\beta_j|$ sebagai penalti penyusutan LASSO. Oleh karena itu, LASSO dapat menyusutkan koefisien menjadi nol. Hal tersebut menyebabkan LASSO dapat menghasilkan model dengan peubah penjelas yang lebih sedikit dan parsimoni. Menurut Prabowo, Wilandari, dan Rusgiyono (2015), penalti penyusutan pada metode LASSO menyebabkan nilai penduga koefisien parameter menyusut sehingga variabel prediktor yang penting atau berpengaruh terhadap model tetap dimasukkan ke dalam model, sedangkan variabel prediktor yang kurang penting akan disusutkan sampai nol dan terseleksi dari model sehingga model menjadi lebih efisien.

ELASTIC-NET

Elastic-net adalah salah satu metode penyusutan seperti Regresi Ridge dan LASSO yang dapat mengatasi permasalahan multikolinearitas yang diperkenalkan oleh Zou dan Hastie pada tahun 2005. *Elastic-net* berfungsi seperti LASSO, yaitu secara simultan melakukan seleksi variabel dan penyusutan koefisien penduga, serta dapat menangani kekurangan dari LASSO. *Elastic-net* mengkombinasikan penalti penyusutan dari regresi ridge dan LASSO, sehingga persamaan yang terbentuk adalah sebagai berikut:

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda_2 \sum_{j=1}^p \beta_j^2 + \lambda_1 \sum_{j=1}^p |\beta_j| \dots \dots \dots (11)$$

$$\text{dengan } \alpha = \frac{\lambda_1}{\lambda_1 + \lambda_2}, 0 \leq \alpha \leq 1$$

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda [(1 - \alpha) \sum_{j=1}^p \beta_j^2 + \alpha \sum_{j=1}^p |\beta_j|] \dots \dots \dots (12)$$

Dimana α adalah parameter gabungan antara Regresi Ridge ($\alpha = 0$) dan LASSO ($\alpha = 1$). Jika $\alpha = 0$, maka penalti penyusutan pada *Elastic-net* menjadi penalti penyusutan regresi ridge. Sedangkan ketika $\alpha = 1$, maka penalti penyusutan pada *Elastic-net* menjadi penalti penyusutan LASSO.

VALIDASI SILANG (CROSS VALIDATION)

Validasi silang merupakan metode yang paling sederhana dan banyak dipakai secara luas untuk menduga kesalahan prediksi. Idealnya, ketika data yang dimiliki memadai, akan dapat ditentukan suatu gugus data validasi yang akan digunakan untuk mengukur ketepatan model yang dimiliki. Namun sering kali data yang dimiliki terlalu sedikit sehingga tidak memungkinkan untuk dilakukan validasi secara langsung. Solusinya adalah melakukan validasi silang yang menggunakan sebagian data yang tersedia untuk membuat model dan sebagian data yang lain digunakan sebagai data pengujian model (Hastie, Tibshirani, dan Friedman, 2008). Pada validasi silang, data dibagi menjadi dua kelompok, yaitu data pemodelan dan data validasi. Pengelompokan data pemodelan dan data validasi dilakukan secara acak dan setiap pengamatan memiliki peluang yang sama untuk menjadi data validasi.

Salah satu jenis validasi silang adalah validasi silang lipat-k. Metode ini baik digunakan ketika jumlah data amat sedikit. Dalam validasi silang lipat-k, amatan dibagi ke dalam k gugus data secara acak dengan ukuran yang hampir sama. Dimana salah satu gugus sebagai gugus validasi dan k-1 gugus digunakan sebagai pembentuk model. Proses validasi silang dilakukan berulang sampai k kali, dengan masing-masing k gugus data digunakan satu kali sebagai validasi model. Proses ini menghasilkan k MSE (*Mean of Squares Error*). $MSE_1, MSE_2, \dots, MSE_k$. Sehingga CVE (*Cross Validation Error*) lipat-k diperoleh dengan cara:

$$CVE(k) = \frac{1}{k} \sum_{i=1}^k MSE_i \dots \dots \dots (13)$$

Validasi silang lipat-k sering kali menggunakan k = 5 atau k = 10 dibandingkan nilai k = n, *leave-one-out cross-validation* (LOOCV). Hal ini dikarenakan lebih menguntungkan dalam proses komputasi. LOOCV membutuhkan waktu yang lebih lama yang berpotensi menyulitkan proses komputasi. Dengan validasi silang lipat-k, maka dapat diperoleh parameter penyusutan yang optimum dilihat dari nilai CVE yang paling kecil.

HASIL DAN PEMBAHASAN

IDENTIFIKASI STATUS KEMISKINAN RUMAH TANGGA MENGGUNAKAN METODE VEP

Setelah diperoleh model pengeluaran per kapita rumah tangga dengan FGLS tiga tahap, kemudian dilakukan pendugaan nilai harapan dan varians pengeluaran per kapita rumah tangga. Selanjutnya, dengan menggunakan nilai harapan dan varians tersebut dihitung nilai peluang setiap rumah tangga untuk jatuh miskin di masa mendatang. Nilai peluang setiap rumah tangga selanjutnya dibandingkan dengan *cut off* sebesar 0,1301 untuk menentukan status kemiskinan rumah tangga tersebut. Berdasarkan proses tersebut, diperoleh persentase rumah tangga menurut status kemiskinannya di Provinsi Jawa Tengah tahun 2017 adalah sebesar 10,4 persen miskin, 4,8 persen rentan miskin dan sisanya 84,8 persen merupakan rumah tangga tidak miskin. Sebagaimana tujuan dari penelitian ini yaitu, untuk mengetahui variabel-variabel yang memengaruhi tingkat kemiskinan dan kerentanan kemiskinan maka dilakukan pendugaan persentase rumah tangga miskin dan rentan miskin untuk setiap kabupaten/kota yang ada di Jawa Tengah.

Sumber: Susenas 2017, diolah

Gambar 1. Persentase rumah tangga miskin di Provinsi Jawa Tengah tahun 2017.

Sumber: Susenas 2017, diolah

Gambar 2. Persentase rumah tangga rentan miskin di Provinsi Jawa Tengah.

Gambar 1 menunjukkan persentase rumah tangga miskin setiap kabupaten/kota di Provinsi Jawa Tengah tahun 2017. Kota Semarang memiliki persentase rumah tangga miskin sebesar 3,8 persen dan merupakan yang terendah di antara kabupaten/kota lainnya. Terendah kedua adalah Kota Salatiga sebesar 3,9 persen dan terendah ketiga adalah Kota Magelang dengan persentase rumah tangga miskin sebesar 4,3 persen. Sementara itu, kabupaten/kota yang memiliki persentase rumah tangga miskin tertinggi adalah Kabupaten Wonosobo sebesar 17,7 persen, tertinggi kedua adalah Kabupaten Brebes sebesar 14,8 persen, dan disusul kabupaten Rembang sebesar 13,6 persen.

Berdasarkan Gambar 2, dapat dilihat persentase rumah tangga rentan miskin di Provinsi Jawa Tengah tahun 2017 untuk setiap kabupaten/kota. Dapat dilihat bahwa kabupaten/kota yang memiliki persentase rumah tangga rentan miskin terendah adalah Kabupaten Kudus sebesar 0,5 persen. Terendah kedua adalah Kabupaten Temanggung sebesar 0,9 persen dan terendah ketiga adalah Kabupaten Sukoharjo sebesar 1 persen. Sementara itu, Kabupaten Pemalang memiliki persentase rumah tangga rentan miskin tertinggi di antara kabupaten/kota lainnya sebesar 14,1 persen. Tertinggi kedua adalah Kabupaten Rembang sebesar 13,2 persen dan disusul Kabupaten Wonosobo sebesar 11,4 persen.

ANALISIS KEMISKINAN DAN KERENTANAN KEMISKINAN

Analisis regresi linier berganda dengan menggunakan metode estimasi OLS untuk mengetahui variabel-variabel yang memengaruhi kemiskinan dan kerentanan kemiskinan dengan cara meregresikan 16 variabel independen dengan persentase rumah tangga miskin kabupaten/kota dan persentase rumah tangga rentan miskin kabupaten/kota di Jawa Tengah sebagai variabel dependen.

Pada *high dimensional data* dalam hal ini kemiskinan dan kerentanan kemiskinan yang bersifat multidimensi yang dipengaruhi oleh banyak variabel sering terjadi masalah multikolinieritas. Hal ini dibuktikan dengan uji VIF pada model kemiskinan dan kerentanan kemiskinan yang menunjukkan terjadinya multikolinieritas. Selain itu, prinsip parsimoni pada *high dimensional data* juga tidak dapat terpenuhi. Lebih lanjut lagi terkait dengan kebijakan masalah kemiskinan, seleksi variabel untuk mengetahui variabel mana yang signifikan memengaruhi kemiskinan dan kerentanan kemiskinan perlu dilakukan. Agar dapat disusun prioritas kebijakan dalam upaya pengentasan kemiskinan dan upaya preventif untuk mencegah kemiskinan, sehingga kebijakan tersebut dapat dilaksanakan secara efektif dan efisien. Beberapa metode yang dapat digunakan untuk mengatasi masalah multikolinieritas dan dapat melakukan seleksi variabel sehingga didapat model yang parsimoni adalah Regresi Ridge, LASSO, dan *Elastic-net*.

REGRESI RIDGE, LASSO, DAN ELASTIC-NET

a. Nilai λ optimum untuk Regresi Ridge

Parameter penyusutan (λ) untuk Regresi Ridge diperoleh dengan menggunakan metode validasi silang (*Cross Validation/CV*), yang menghasilkan CVE minimum. CV dilakukan dengan *ten-fold cross validation* atau validasi silang lipat-10. Dalam validasi silang lipat-10, amatan dibagi ke dalam 10 gugus data secara acak dengan ukuran yang hampir sama. Dimana salah satu gugus sebagai gugus validasi dan 9 gugus digunakan sebagai pembentuk model. Proses validasi silang dilakukan berulang sampai 10 kali, dengan masing-masing gugus data digunakan satu kali sebagai validasi model. Sehingga akan diperoleh sebanyak 10 MSE yang kemudian dari 10 MSE tersebut akan dihitung nilai

Gambar 3. Validasi silang parameter penyusutan (λ) Regresi Ridge pada model kemiskinan.

Gambar 4. Validasi silang parameter penyusutan (λ) Regresi Ridge pada model kerentanan kemiskinan.

CVE. Dimana dari perhitungan nilai CVE ini dapat diperoleh parameter penyusutan yang optimum dilihat dari nilai CVE yang paling kecil.

Gambar 3 menunjukkan validasi silang parameter penyusutan (λ) untuk Regresi Ridge pada model kemiskinan sedangkan Gambar 4 menunjukkan validasi silang parameter penyusutan (λ) untuk Regresi Ridge pada model kerentanan kemiskinan. Dua garis vertikal pada CV menggambarkan nilai λ optimum dengan CVE minimum. Garis vertikal pertama paling kiri menunjukkan nilai $\log \lambda$ minimum (λ_{min}) dan garis vertikal kedua menunjukkan nilai λ_{1se} . Namun, nilai λ_{1se} lebih direkomendasikan dijadikan λ optimum karena nilai terlindungi terhadap *overfitting* dengan memilih nilai λ lebih besar dan juga dapat menghasilkan model dengan parameter yang lebih sedikit (dapat melakukan seleksi variabel). Pada akhirnya diperoleh nilai λ optimum untuk model kemiskinan, yaitu λ_{1se} sebesar 15,051 dan $\log \lambda_{1se}$ sebesar 2,711. Nilai λ optimum untuk model kerentanan kemiskinan, yaitu λ_{1se} sebesar 24,389 dan $\log \lambda_{1se}$ sebesar 3,194.

b. Nilai λ optimum untuk LASSO

Parameter penyusutan (λ) untuk LASSO diperoleh dengan menggunakan metode validasi silang (*Cross Validation*/CV), yang menghasilkan CVE minimum. Proses CV pada metode ini sama dengan CV pada Regresi Ridge. Hasil CV dapat dilihat pada Gambar 5 nilai λ optimal yang diperoleh pada model kemiskinan adalah λ_{1se} sebesar 0,749 dan $\log \lambda_{1se}$ sebesar -0,289. Kemudian, pada Gambar 6 nilai λ optimal yang diperoleh pada model kerentanan kemiskinan adalah λ_{1se} sebesar 0,918 dan $\log \lambda_{1se}$ sebesar -0,085.

Gambar 5. Validasi silang parameter penyusutan (λ) LASSO pada model kemiskinan.

Gambar 6. Validasi silang parameter penyusutan (λ) LASSO pada model kerentanan kemiskinan.

c. Nilai λ optimum untuk *Elastic-net*

Elastic-net menggunakan dua penalti penyusutan yaitu, penalti Regresi Ridge $\sum_{j=1}^p \beta_j^2$ dan penalti LASSO $\sum_{j=1}^p |\beta_j|$. Parameter untuk penalti Regresi Ridge adalah λ dikali dengan α , sedangkan parameter untuk penalti LASSO adalah λ dikali dengan $(1-\alpha)$ dengan α bernilai antara nol sampai satu. Pemilihan nilai α dan λ optimum dengan CV yang menghasilkan CVE minimum. Pada model kemiskinan nilai α optimum adalah 0,99 dengan λ_{1se} sebesar 0,689 dan $\log \lambda_{1se}$ sebesar -0,372 (Gambar 7). Kemudian, pada model kerentanan kemiskinan nilai α optimum adalah 0,93 dengan λ_{1se} sebesar 0,747 dan $\log \lambda_{1se}$ sebesar -0,292 (Gambar 8).

Gambar 7. Validasi silang parameter penyusutan (λ) Elastic-net pada model kemiskinan.

Gambar 8. Validasi silang parameter penyusutan (λ) Elastic-net pada model kerentanan kemiskinan.

Koefisien Regresi

Perbandingan koefisien regresi hasil analisis dengan RLB, Regresi Ridge, LASSO, dan *Elastic-net* dapat dilihat pada Tabel 4.

Tabel 4. Koefisien regresi pada Regresi Ridge, LASSO, dan *Elastic-net* pada model kemiskinan dan kerentanan kemiskinan.

Variabel	Model Kemiskinan				Model kerentanan Kemiskinan			
	RLB	Regresi Ridge	LASSO	<i>Elastic-net</i>	RLB	Regresi Ridge	LASSO	<i>Elastic-net</i>
(1)	(2)	(3)	(3)	(5)	(6)	(7)	(8)	(9)
Dimensi Pendidikan								
RLS	0,716	-0,128	0	0	-2,044	-0,101	-0,049	-0,273
APS	-0,009	-0,010	0	0	0,148	-0,008	0	0
Dimensi kesehatan								
Angka morbiditas	-0,035	0,012	0	0	0,261	0,039	0,041	0,059
AHH	-0,421	-0,089	0	-0,011	-0,341	-0,065	0	0
Dimensi ketenagakerjaan								
TPT	0,027	-0,018	0	0	-0,591	0,018	0	0
TPAK	0,030	0,007	0	0	0,031	-0,014	0	0
Persentase rumah tangga dengan lapangan usaha utama KRT pertanian								
	0,002	0,013	0,017	0,015	-0,127	0,004	0	0
Dimensi kualitas hidup								
Rata-rata pengeluaran makanan	-0,874	-1,219	0	0	30,769	1,245	2,767	6,485
Rata-rata pengeluaran non makanan	-1,259	-0,696	-0,814	-0,886	-2,623	-0,142	0	0
Jumlah penduduk	-0,001	0,001	0	0	-0,001	0,001	0	0
Dimensi fasilitas rumah tinggal								
Persentase rumah tangga dengan jenis lantai kualitas rendah	0,079	0,014	0,023	0,026	0,052	0,006	0	0
Persentase rumah tangga dengan jenis kloset tidak layak	0,147	0,039	0,158	0,159	0,207	0,025	0,103	0,141
Persentase rumah tangga dengan bahan bakar memasak kualitas rendah	-0,021	0,012	0	0	0,118	0,002	0	0
Dimensi ekonomi								
Pertumbuhan ekonomi	-0,711	-0,045	0	0	-0,065	0,091	0	0
Inflasi	-1,115	-0,142	0	0	-0,530	-0,003	0	0
Upah minimum	-12,031	-2,043	-2,594	-3,071	0,953	-1,227	0	0
Jumlah parameter	16	16	5	6	16	16	4	4

d. Menentukan model terbaik

Pemilihan model terbaik dari model Regresi Ridge, LASSO, dan *Elastic-net* berdasarkan nilai MSE. Dimana model terbaik adalah yang mempunyai nilai MSE terkecil. Berdasarkan Tabel 5, nilai MSE paling kecil pada model kemiskinan dan kerentanan kemiskinan adalah pada metode *Elastic-net* sebesar 4,214 dan 9,539. Berdasarkan model terbaik yang telah terpilih, dari 16 variabel yang pada awalnya dimasukkan ke dalam model kemiskinan dan kerentanan kemiskinan dengan menggunakan metode *Elastic-net* yang mana menghasilkan model terbaik dan dapat mengatasi masalah multikolinieritas terdapat beberapa variabel yang mengalami penyusutan ke arah nol atau dengan kata lain terseleksi dari model. Pada model kemiskinan terdapat 10 variabel yang mengalami penyusutan ke arah nol dan pada model kerentanan kemiskinan terdapat 12 variabel yang mengalami penyusutan ke arah nol. Variabel-variabel independen yang terseleksi dari model tersebut merupakan variabel independen yang tidak berpengaruh signifikan terhadap variabel dependen, sedangkan variabel-variabel independen yang masih tetap ada dalam model merupakan variabel yang mempunyai pengaruh signifikan terhadap variabel dependennya. Pada model kemiskinan terdapat 6 variabel yang signifikan memengaruhi persentase rumah tangga miskin, yaitu AHH, persentase rumah tangga dengan lapangan usaha utama KRT pertanian, rata-rata pengeluaran non makanan, persentase rumah tangga dengan jenis lantai kualitas rendah, persentase rumah tangga dengan jenis kloset tidak layak, dan upah minimum regional. Kemudian, pada model kerentanan kemiskinan terdapat 4 variabel yang signifikan memengaruhi persentase rumah tangga rentan miskin, yaitu rata-rata lama sekolah, angka morbiditas, rata-rata pengeluaran makanan, dan persentase rumah tangga dengan jenis kloset tidak layak.

Tabel 5. Perbandingan metode Regresi Ridge, LASSO, dan *Elastic-net* pada model kemiskinan dan kerentanan.

Metode (1)	Model Kemiskinan	Model Kerentanan Kemiskinan
	MSE (2)	MSE (3)
Regresi Ridge	4,980	11,816
LASSO	4,391	11,227
<i>Elastic-net</i>	4,214	9,539

KESIMPULAN

Berdasarkan hasil dan pembahasan sebelumnya, maka beberapa kesimpulan yang dapat diambil adalah pertama, persebaran rumah tangga menurut status kemiskinannya di Provinsi Jawa Tengah adalah sebesar 10,4 persen rumah tangga miskin, 4,8 persen rumah tangga rentan miskin, dan 84,8 persen rumah tangga tidak miskin. Kemudian, jika dilihat menurut kabupaten/kota persentase rumah tangga miskin paling tinggi ada di Kabupaten Wonosobo, Kabupaten Brebes, dan Kabupaten Rembang. Kemudian, yang paling rendah ada di Kota Semarang, Kota Salatiga, dan Kota Magelang. Persentase rumah tangga rentan miskin paling tinggi ada di Kabupaten Pemalang, Kabupaten Rembang, dan Kabupaten Wonosobo. Kemudian, yang paling rendah ada di Kabupaten Kudus, Kabupaten Temanggung, dan Kabupaten Sukoharjo. Kedua, model terbaik dari model kemiskinan dan kerentanan dengan metode *Elastic-net*. Variabel-variabel yang signifikan memengaruhi persentase rumah tangga miskin di Provinsi Jawa Tengah tahun 2017 adalah AHH, persentase rumah tangga dengan lapangan usaha utama KRT pertanian, rata-rata pengeluaran non makanan, persentase rumah tangga dengan jenis lantai kualitas rendah, persentase rumah tangga dengan jenis kloset tidak layak, dan upah minimum regional. Selanjutnya, variabel-variabel yang signifikan memengaruhi persentase rumah tangga rentan miskin di Provinsi Jawa Tengah tahun 2017 adalah rata-rata lama sekolah, angka morbiditas, rata-rata pengeluaran makanan, dan persentase rumah tangga dengan jenis kloset tidak layak.

DAFTAR PUSTAKA

- Badan Pusat Statistik. (2017). *Perhitungan dan Analisis Kemiskinan Makro Indonesia Tahun 2017*. Jakarta: Badan Pusat Statistik.
- Chaudhuri, S., Jalan, J., & Suryahadi, A. (2002). Assessing Household Vulnerability to Poverty from Cross-sectional Data: A Methodology and Estimates from Indonesia. *World*, 0102–52(April), 36.
- Chaudhuri, S. (2003). Assessing vulnerability to poverty: concepts, empirical methods and illustrative examples. *Department of Economics*, 56.
- Dariwardani, N. M. I. (2014). Analisis Dinamika Kemiskinan (Poverty Dynamics) di Bali Berdasarkan Data Susenas Panel 2008 2010. *Jurnal Ekonomi Kuantitatif Terapan*, 7(1).
- Draper, N. R., & Smith, H. (1981). *Applied Regression Analysis*, John Wiley & Sons. New York, 407.
- Fanny, R., Djuraidah, A., & Alamudi, A. (2018). Pendugaan Produktivitas Bagan Perahu dengan Regresi Gulud, LASSO dan Elastic-net. *Xplore: Journal of Statistics*, 2(2), 7–14.
- Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. New York: Springer.
- Hoerl, A. E., & Kennard, R. W. (2000). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 42(1), 80–86.
- Haughton, Jonathan dan Shahidur R Khandker. (2009). *Handbook on Poverty and Inequality*. Washington D.C: The World Bank.
- Kumala, A. Z., Agustini, H. N., & Rais. (2010). Dinamika Kemiskinan dan Pengukuran Kerentanan Kemiskinan dalam Upaya Melindungi Anak-Anak dari Dampak Kemiskinan, (1776), 1–35.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (Vol. 5). Boston: McGraw-Hill Irwin.
- Ohyver, M. (2011). Metode Regresi Ridge Untuk Mengatasi Kasus Multikolinear. *ComTech: Computer, Mathematics and Engineering Applications*, 2(1), 451-457.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 67(2), 301–320.