
PENCARIAN KERNEL TERBAIK SUPPORT VECTOR REGRESSION PADA KASUS DATA KEMISKINAN DI INDONESIA DENGAN USER INTERFACE (GUI) MATLAB

Muhammad Ghazali¹, Ramdani Purnamasari²

¹ Pendidikan Matematika, STKIP Taman Siswa Bima

² Pendidikan Teknologi Informasi, STKIP Taman Siswa Bima

Alamat e-mail : m.ghazali11@gmail.com, ramdanipurnamasari5@gmail.com

ABSTRACT

Poverty is a topic that is often discussed in various scientific study forums. Facts on the ground show that increasing development has not been able to reduce the increasing number of poor people. Statistical studies on poverty data need to be carried out to assist the government in mapping policy-making patterns. One of the variables in mapping poverty data patterns is the Poverty Depth Index. The poverty depth index is a measure of the average gap in the distribution of each population's expenditure on the poverty line. Many factors affect the poverty depth index, especially from health, human resources and economic indicators. Therefore, a statistical modeling is needed to analyze the factors that affect the poverty depth index in Indonesia. The poverty data used in this study were sourced from the 2019 SUSENAS data in the form of data with individual observations of all provinces in Indonesia. Several previous studies using the Support Vector Regression (SVR) method to estimate the Poverty Depth Index as a response variable with several variables from health and economic indicators showed a very good level of model accuracy. However, SVR is constrained by choosing the right kernel to find the optimum prediction accuracy. So it is necessary to create a user interface that automatically selects the best type of kernel to facilitate the modeling process. The user interface will also help users to use the SVR even if they do not know the programming language. This study aims to: (1) produce a statistical analysis that makes it easier to map the pattern of factors that influence poverty in Indonesia, (2) produce a user interface that makes it easier for users to analyze poverty data in Indonesia. The conclusion obtained from this study is that the most accurate estimation is to use a degree 1 Gaussian kernel (RBF) SVR model while using the Polynomial kernel is not enough to provide a good estimate.

Keywords: Data Mining, Kernel, Regression, SVR.

PENDAHULUAN

Kemiskinan di Indonesia adalah permasalahan berat yang masih dihadapi oleh pemerintah. Kemiskinan selain dipengaruhi oleh dimensi ekonomi, juga berkaitan dengan berbagai dimensi antara lain dimensi sosial, budaya, sosial politik, lingkungan, kesehatan, pendidikan, agama, dan budi pekerti. Menelaah kemiskinan secara multidimensional sangat diperlukan untuk perumusan kebijakan pengentasan kemiskinan [1]. Selama ini kemiskinan lebih cenderung dikaitkan dengan dimensi

ekonomi karena dimensi ini paling mudah diamati, diukur dan diperbandingkan. Padahal kemiskinan berkaitan juga dengan berbagai dimensi antara lain dimensi sosial, budaya, sosial politik, lingkungan, kesehatan, pendidikan, agama, dan budi pekerti [2].

Berdasarkan data Badan Pusat Statistik pada bulan Maret 2019, jumlah penduduk miskin di Indonesia adalah 25,14 juta orang atau sekitar 9,41 persen. Sebenarnya angka ini menurun 0,25 persen dari persentase bulan September 2018 atau turun sebanyak

0,80 jiwa. Tetapi angka ini masih tergolong sangat besar karena Garis Kemiskinan pada Maret 2019 tercatat sebesar 425.250,-/kapita/bulan dengan komposisi Garis Kemiskinan Makanan sebesar 313.231,-. Pada bulan Maret 2019 tercatat bahwa rata-rata rumah tangga miskin di Indonesia memiliki 4,68 anggota keluarga [3].

Penelitian sebelumnya yang melakukan pemodelan terhadap kemiskinan di Indonesia diantaranya adalah pemodelan kemiskinan di Jawa Timur menggunakan metode *Geographically Weighted Regression* (GWR) [4], pemodelan kemiskinan di Jawa Timur dengan menggunakan *Structural Equation Modeling-Partial Least Square* [5]. *Meta-Analytic Structural Equation Modeling* (MASEM) untuk memodelkan kemiskinan di pulau Jawa [6], kemudian pemodelan dengan pendekatan *Multivariate Adaptive Regression SPLINES* (MARS) pada data penduduk miskin di Indonesia tahun 2012 [7], kemudian dilanjutkan dengan penelitian menggunakan pendekatan *Generalized Methods of Moments* (GMM) dengan Regresi Data Panel model *Fixed Effect* untuk data tahun 2008-2012 [8]. Tetapi penelitian diatas hanya memberikan ketepatan prediksi dengan menggunakan koefisien determinasi (R^2) hanya berkisar antara 40-60 persen.

Kemudian dicoba menggunakan *Support Vector Regression* (SVR) ternyata menghasilkan nilai koefisien determinasi lebih baik dimana penaksiran optimumnya mencapai 80-90 persen pada data indeks kedalaman kemiskinan dan angka harapan hidup di Indonesia [9][10]. Kendala utama pada SVR adalah pemilihan berbagai macam kernel untuk mencari penaksiran optimum. Sehingga perlu dilakukan penelitian ini untuk memilih kernel optimum dengan *user interface* untuk memodelkan faktor-faktor yang mempengaruhi indeks kedalaman kemiskinan di Indonesia.

Support Vector Regression

Algoritma *Support Vector Machine* (SVM) pertama kali diusulkan oleh Vapnik (1995) untuk klasifikasi dua kategori atau binomial. Pada bentuk yang paling sederhana, SVM memisahkan titik-titik dari kelas yang berbeda, misalkan kelas $\{+1\}$ dan $\{-1\}$, dengan *hyperplane* tunggal pada ruang berdimensi banyak yang pada akhirnya partisi-partisi tersebut diselesaikan secara *nonlinear*. *Hyperplane* yang optimum diperoleh melalui program *nonlinear*, tepatnya *quadratic programming* [11].

Support Vector Regression (SVR) adalah penerapan *Support Vector Machine* (SVM) pada kasus regresi. SVM yang merupakan metode klasifikasi dengan variabel respons adalah variabel ordinal dikembangkan menjadi metode regresi dimana variabel bebasnya adalah variabel numerik berupa bilangan riil dan kontinu.

Misalkan terdapat sekumpulan data $D = \{(x^1, y^1), (x^2, y^2), \dots, (x^l, y^l)\}, x \in R^n, y \in R$ (1)

dengan fungsi linear $f(x) = \langle w, x \rangle + b$ (2)

Fungsi optimal regresi dari persamaan diatas adalah sebagai berikut:

$$\Phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_i (\xi_i^- + \xi_i^+) \quad (3)$$

dimana C adalah nilai yang ditentukan sebelumnya, dan (ξ_i^-, ξ_i^+) adalah variabel *slack* yang menunjukkan batas atas dan batas bawah pada output dalam sistem [12].

Faktor $\|w\|^2$ dinamakan faktor regularisasi. Meminimalkan $\|w\|^2$ akan membuat fungsi setipis (*flat*) mungkin, sehingga dapat mengontrol kapasitas fungsi (*function capacity*). Menggunakan ide *ϵ -insensitive loss function* yang diperkenalkan Vapnik kita meminimalkan *norm* dari w agar mendapatkan *generalisasi* yang baik untuk fungsi regresi f [13].

Persamaan *ϵ -insensitive loss function*, untuk menyelesaikan optimisasi persamaan (3) adalah

$$L_\varepsilon(y) = \begin{cases} 0 & \text{untuk } |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon & \text{selainnya} \end{cases} \quad (4)$$

sehingga solusinya adalah sebagai berikut

$$\max_{\alpha, \alpha^*} W(\alpha, \alpha^*) = \max_{\alpha, \alpha^*} -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{i=1}^l \alpha_i (y_i - \varepsilon) \alpha_i^* - \alpha_i^* (y_i + \varepsilon) \quad (5)$$

dengan batasan

$$\begin{aligned} 0 \leq \alpha_i, \alpha_i^* \leq C, \quad i = 1, \dots, l \\ \sum_{j=1}^l (\alpha_i - \alpha_i^*) = 0 \end{aligned} \quad (6)$$

Selesaikan persamaan (5) dengan batasan (6) dengan menggunakan *lagrange multiplier* maka kondisi optimal dari fungsi regresi dituliskan sebagai berikut:

$$\bar{w} = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_j) \quad (7)$$

$$\bar{b} = -\frac{1}{2} \langle \bar{w}, \langle K(\mathbf{x}_i, \mathbf{x}_r) + K(\mathbf{x}_i, \mathbf{x}_s) \rangle \rangle \quad (8)$$

Jika $\varepsilon = 0$ maka kita peroleh optimisasi loss function dalam bentuk persamaan yang lebih sederhana sebagai berikut

$$\min_{\beta} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\beta_i)(\beta_j) K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{i=1}^l \beta_i y_i \quad (9)$$

dengan batasan

$$\begin{aligned} -C \leq \beta_i \leq C, \quad i = 1, \dots, l \\ \sum_{i=1}^l \beta_i = 0 \end{aligned}$$

dimana $K(\mathbf{x}_i, \mathbf{x}_j)$ adalah fungsi kernel dari $(\mathbf{x}_i, \mathbf{x}_j)$

Fungsi regresi optimal dari persamaan (2) dituliskan sebagai berikut [14]:

$$\langle \bar{w}, x \rangle = \sum_{i=1}^l \beta_i K(\mathbf{x}_i, \mathbf{x}_j) \quad (10)$$

$$\bar{b} = -\frac{1}{2} \langle \bar{w}, \langle K(\mathbf{x}_i, \mathbf{x}_r) + K(\mathbf{x}_i, \mathbf{x}_s) \rangle \rangle \quad (11)$$

Kernel

Kernel dalam ilmu data mining adalah fungsi perkalian dalam (*inner product*) pada ruang fitur. Fungsi kernel memungkinkan kita untuk mengimplementasikan suatu model pada ruang dimensi lebih tinggi (ruang fitur) tanpa harus mendefinisikan fungsi pemetaan dari ruang input ke ruang fitur. Pada penelitian ini, optimasi dibantu dengan fungsi kernel diantaranya kernel *polynomial* dan *Gaussian Radial Basis Function* (RBF).

Persamaan fungsi kernel *polynomial* adalah

$$K(x, x') = ((x, x') + 1)^d \quad (12)$$

Sedangkan fungsi kernel RBF adalah

$$K(x, x') = \exp\left(-\frac{\|x-x'\|^d}{2\sigma^2}\right) \quad (13)$$

dimana d adalah derajat kernel [15].

Ukuran Kebaikan Model

Untuk memilih model yang terbaik digunakan prosedur validasi model menggunakan *Mean Square Error* (MSE) dan R^2 . Koefisien determinasi R^2 merupakan persentase pengaruh dari variabel prediktor terhadap variabel respon. Persamaannya ditulis sebagai berikut [16]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (14)$$

dan

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (15)$$

y_i menyatakan objek pengamatan ke- i dan \hat{y}_i adalah prediksi data ke- i . Maka model terbaik adalah model yang memiliki nilai MSE paling kecil dan nilai R^2 yang paling maksimal.

METODOLOGI PENELITIAN

Sumber Data dan Variabel Penelitian

Data yang digunakan pada penelitian ini merupakan publikasi data sekunder dari Survei Sosial Ekonomi Nasional (SUSENAS) pada seluruh propinsi di Indonesia tahun 2019. Berdasarkan konsultasi peneliti dengan BPS Kabupaten Bima. Variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

1. indeks kedalaman kemiskinan (Y)
2. Persentase pengeluaran per kapita untuk non makanan (X_1)
3. Rata-rata lama sekolah (X_2)
4. Angka harapan hidup (X_3)

Pengolahan data pada penelitian ini menggunakan *software* GUI Matlab

Metode Analisis

Penelitian ini bertujuan untuk membangun sebuah *interface* atau *software* simulasi yang akan membantu menganalisa ketiga variabel tersebut terhadap faktor kemiskinan dengan memodelkan secara

matematis persamaannya dan memberikan grafik yang lebih akurat. Metode pemodelan matematika yang digunakan menggunakan Support Vector Regression dengan kernel Polinomial dan RBF dengan bantuan toolbox Matlab buatan Steve R. Gunn.

Untuk memilih model yang terbaik digunakan prosedur validasi model menggunakan Koefisien determinasi (R^2) merupakan persentase pengaruh dari variabel prediktor terhadap variabel respon. Persamaannya ditulis sebagai berikut [13] :

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

y_i menyatakan objek pengamatan ke- i dan \hat{y}_i adalah prediksi data ke- i . Maka model terbaik adalah model yang memiliki nilai R^2 yang paling maksimal.

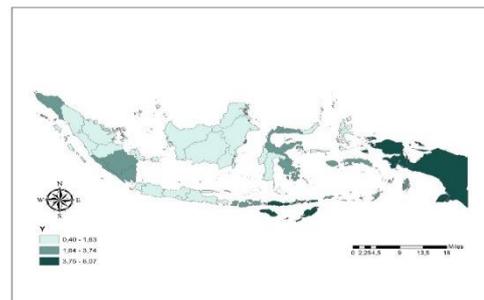
HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini merupakan publikasi data sekunder dari Survei Sosial Ekonomi Nasional (SUSENAS) pada seluruh propinsi di Indonesia tahun 2019. Berdasarkan konsultasi peneliti dengan BPS Kabupaten Bima, variabel yang disarankan untuk penelitian ini tetap terdiri dari variabel terikat (Y) yaitu indeks kedalaman kemiskinan [17], sedangkan variabel bebasnya terdiri dari Persentase pengeluaran per kapita untuk non makanan (X_1) yang mewakili faktor ekonomi [18], Rata-rata lama sekolah (X_2) yang mewakili faktor SDM [19], Angka harapan hidup (X_3) yang mewakili faktor kesehatan [20]. Deskripsi dari data penelitian ditunjukkan dalam tabel berikut

Tabel 3.1 Statistik Deskripsi

	Min	Max	Mean	Std	Var
Y	.40	6.07	1.80	1.362	1.856
X1	3.21	12.78	6.05	1.856	3.446
X2	6.65	11.06	8.54	.9292	.864
X3	64.82	74.92	69.89	2.548	6.493

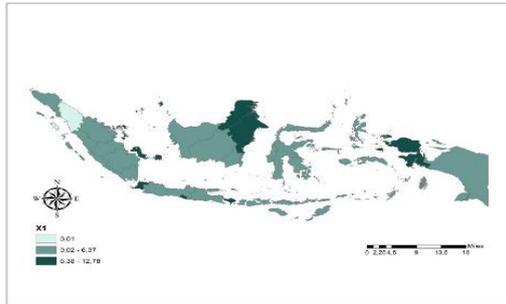
Rata pengeluaran perkapita untuk produk makanan di Indonesia adalah Rp 605.350 dengan standard deviasi 185.633, sedangkan Rata-rata lama sekolah di Indonesia adalah 11 tahun atau kelas XI SMA, sedangkan angka harapan hidup adalah 69,9 tahun dengan variansi 1,3 tahun.



Gambar 3.1 Peta Indeks Kedalaman Kemiskinan di Indonesia (Y)

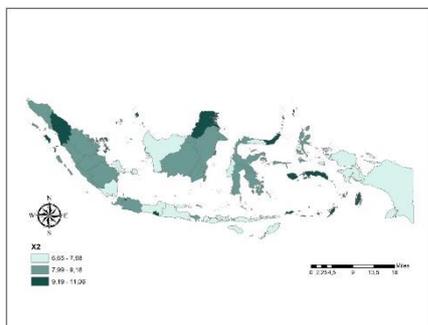
Sebaran Indeks Kedalaman Kemiskinan yang paling tinggi berada di daerah Papua dan Nusa Tenggara, sedangkan yang paling rendah adalah daerah Jawa dan Kalimantan. Sedangkan beberapa propinsi di Sulawesi memiliki Indeks Kedalaman Kemiskinan yang beragam.

Gambar 3.2 Peta (X1) Rata-rata Pengeluaran per Kapita Sebulan di Perkotaan dan Perdesaan - Bukan Makanan (per Rp



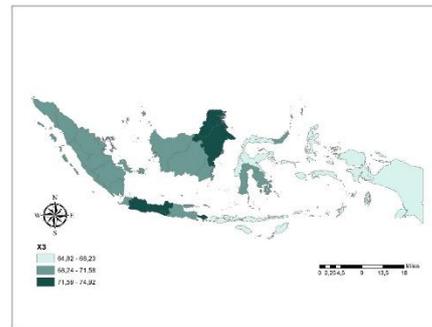
100.000) di Indonesia

Secara umum sebaran daya konsumsi masyarakat yang ditunjukkan oleh oleh variabel Rata-rata Pengeluaran Non Makanan per Kapita Sebulan cukup tinggi untuk sebagian besar provinsi di Indonesia, hanya Propinsi Sumatera Barat yang memiliki Rata-rata Pengeluaran Non Makanan yang rendah.



Gambar 3.3 Peta Rata-rata lama sekolah di Indonesia tahun 2019

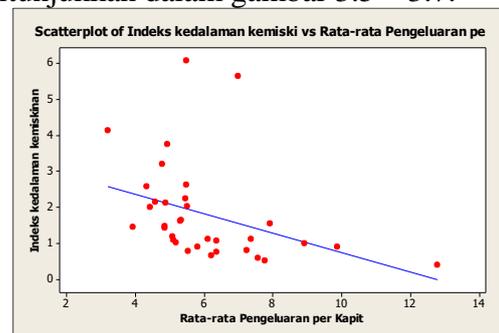
Sebaran pendidikan di Indonesia yang diukur dengan variabel rata-rata lama sekolah menunjukkan bahwa sebagian besar propinsi di wilayah timur Indonesia memiliki indeks pendidikan yang masih rendah, terutama di Wilayah Papua, Nusa Tenggara (kecuali Bali), Kalimantan Barat, Bengkulu, Jawa Tengah dan Jawa Timur.



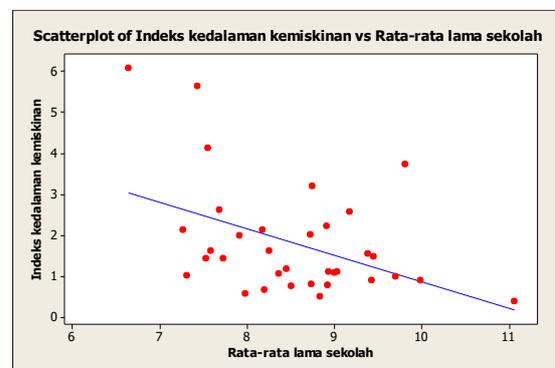
Gambar 3.4 Angka Harapan Hidup di Indonesia tahun 2019

Faktor kesehatan masyarakat diukur dengan variabel Angka Harapan hidup di Indonesia menunjukkan bahwa beberapa propinsi di Indonesia wilayah timur memiliki angka harapan hidup yang lebih rendah dibandingkan propinsi di bagian barat Indonesia.

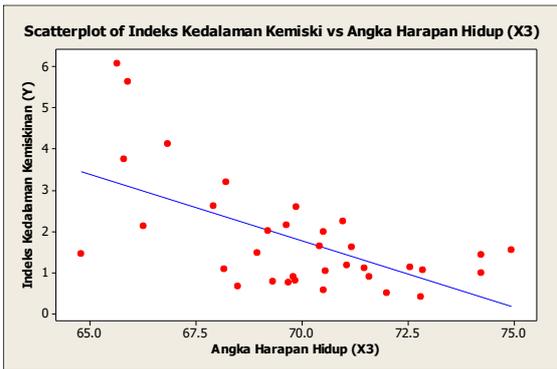
Adapun pola regresi antara variabel Y dengan variabel X1, X2 dan X3 ditunjukkan dalam gambar 3.5 – 3.7.



Gambar 3.5 Plot Y vs X1



Gambar 3.6 Plot Y vs X2



Gambar 3.7 Plot Y vs X3

Berdasarkan Gambar 3.4 sampai Gambar 3.7 terlihat bahwa hubungan variabel Y dan variabel X1, X2 dan X3 adalah berbanding terbalik yang ditunjukkan dengan garis regresi linear yang menurun. Hal ini berarti bahwa semakin baik penghasilan masyarakat yang ditunjukkan oleh besarnya biaya belanja non konsumsi, faktor pendidikan yang ditunjukkan oleh lamanya menempuh pendidikan, faktor kesehatan yang ditunjukkan oleh lamanya harapan hidup, maka tingkat kedalaman kemiskinan akan semakin menurun.

Penelitian ini bertujuan untuk membangun sebuah *interface* atau software simulasi yang akan membantu menganalisa ketiga variabel tersebut terhadap faktor kemiskinan dengan memodelkan secara matematis persamaannya dan memberikan grafik yang lebih akurat. Metode pemodelan matematika yang digunakan menggunakan Support Vector Regression dengan kernel Polinomial dan RBF dengan bantuan toolbox Matlab buatan Steve R. Gunn [15].

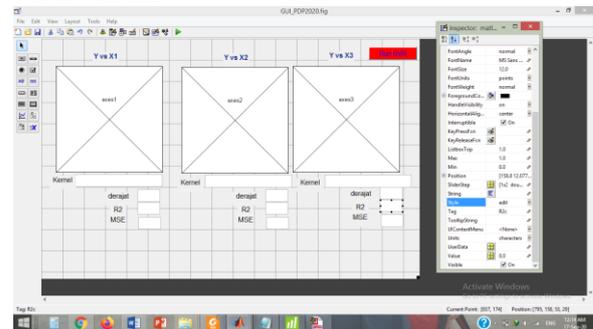
Tabel 3.2 Hasil pencarian kernel

Metode	Kernel	derajat	R ²		
			X1	X2	X3
OLS			8.33%	15.14%	13.92%
SVR	Polynomial	1	8.33%	15.14%	13.92%
SVR	Polynomial	2	7.53%	25.38%	14.10%
SVR	Polynomial	3	7.32%	25.21%	14.72%
SVR	RBF	1	57.35%	62.43%	87.10%
SVR	RBF	2	57.22%	62.43%	86.59%
SVR	RBF	3	57.22%	62.43%	86.01%

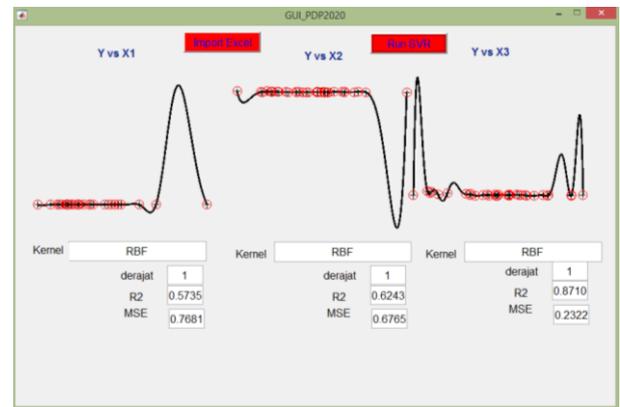
Dengan menggunakan SVR kernel Polynomial tidak memberikan penaksiran

model yang lebih baik dari Regresi Linear (OLS), tetapi jika menggunakan kernel RBF akan menghasilkan penaksiran yang lebih baik

Penelitian ini membuat GUI yang memudahkan pengguna untuk mencari kernel dan penaksiran terbaik dengan menu yang lebih mudah tanpa harus memahami bahasa pemrograman Matlab. Tampilan dari rancangan interface menggunakan GUI Matlab ditunjukkan oleh gambar dibawah ini.



Gambar 3.8 Rancangan Interface dengan Matlab



Gambar 3.9 Output pencarian kernel terbaik

Penggunaan Support Vector regression (SVR) untuk menaksir hubungan antara variabel Y dan X1 memperlihatkan bahwa kernel yang terbaik untuk menaksir hubungan diantara keduanya adalah kernel Polynomial derajat 1 yang menghasilkan R² = 13, 58%. Sedangkan untuk menaksir hubungan antara variabel Y dan X2 memperlihatkan bahwa kernel yang terbaik untuk menaksir hubungan diantara

keduanya adalah kernel Polynomial derajat 1 yang menghasilkan $R^2 = 14.53\%$. Sedangkan antara variabel Y dan X3 memperlihatkan bahwa kernel yang terbaik untuk menaksir hubungan diantara keduanya adalah kernel Polynomial derajat 1 yang menghasilkan $R^2 = 87.10\%$.

KESIMPULAN

Model terbaik untuk mengestimasi pengaruh faktor pendidikan masyarakat yang diukur melalui rata-rata lama sekolah terhadap angka harapan hidup di Indonesia dengan metode *Support Vector Regression* (SVR) diperoleh dengan menggunakan kernel *Polynomial* derajat 1 untuk variabel Y terhadap variabel X1, X2 dan X3.

SARAN

Saran untuk penelitian selanjutnya adalah membandingkan data dengan beberapa kernel SVM yang berbeda diantaranya kernel *Splines*, *B-Splines*, *Fouries Series*, dll. Perlu ada penelitian lanjutan tentang pola hubungan kedua variabel berdasarkan kuartil data untuk melihat pengaruh antar variable

UCAPAN TERIMAKASIH

Ucapan terimakasih kami khaturkan kepada Ristekdikti yang membiayai Penelitian kami dalam hibah skema Penelitian Dosen Pemula (PDP) Tahun pembiayaan 2020. Ucapan terimakasih pula kami sampaikan kepada BPS Kabupaten Bima yang membantu peneliti dalam memperoleh data dan membantu pemilihan variabel yang akan diteliti, beserta seluruh pihak yang membantu terlaksananya penelitian ini.

DAFTAR PUSTAKA

- [1] C. Suryawati, "Memahami Kemiskinan Secara Multidimensional," *J. Manaj. Pelayanan Kesehat.*, vol. 8, no. 3, pp. 121–129, 2005.
- [2] G. Anuraga, *Pemodelan Kemiskinan di Jawa Timur Dengan Structural Equation Modeling-Partial Least Square (Tesis)*. Surabaya: Institut Teknologi Sepuluh Nopember, 2005.
- [3] BPS, "Persentase Penduduk Miskin Maret 2019 Sebesar 9,41 Persen," *Badan Pusat Statistika*, 2019. [Online]. Available: <https://www.bps.go.id/pressrelease/2019/07/15/1629/persentase-penduduk-miskin-maret-2019-sebesar-9-41-persen.html>.
- [4] Y. Damayanti and V. Ratnasari, "Pemodelan Penduduk Miskin di Jawa Timur Menggunakan Metode Geographically Weighted Regression (GWR)," *J. Sains dan Seni ITS*, vol. 2, no. 2, 2013.
- [5] G. Anuraga and B. W. Otok, "PEMODELAN KEMISKINAN DI JAWA TIMUR DENGAN STRUCTURAL EQUATION MODELING-PARTIAL LEAST SQUARE," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 1, no. 2, 2013.
- [6] A. N. Ngafiyah, *Meta-Analytic Structural Equation Modeling (MASEM) Pada Faktor-Faktor Yang Mempengaruhi Kemiskinan Di Pulau Jawa (Tesis)*. Surabaya: Institut Teknologi Sepuluh Nopember, 2014.
- [7] E. D. A. A. Sita and B. W. Otok, "Pendekatan Multivariate Adaptive Regression SPLINES (MARS) pada Pemodelan Penduduk Miskin di Indonesia Tahun 2008-2012," in *Prosiding Seminar Nasional Matematika*, 2014.

- [8] M. Ghazali and B. W. Otok, "Pemodelan fixed effect pada regresi data longitudinal dengan estimasi generalized method of moments (studi kasus data penduduk miskin di indonesia)," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 4, no. 1, 2016.
- [9] M. Ghazali and I. Fitriati, "Aplikasi Support Vector Regression pada Pemodelan Kemiskinan di Indonesia," in *Proceeding Seminar Nasional Riset Ilmu Komputer 2016*, 2016, pp. 129–133.
- [10] I. Fitriati and M. Ghazali, "PEMODELAN PENGARUH RATA-RATA LAMA SEKOLAH TERHADAP INDEKS KEDALAMAN KEMISKINAN DI INDONESIA MENGGUNAKAN SUPPORT VECTOR REGRESSION," *Pros. Semin. Nas. Ris. Kuantitatif Terap. 2017*, vol. Vol. 1, No, no. April, pp. 100–105, 2017.
- [11] B. Santoso, *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis, Teori dan Aplikasi*. Yogyakarta: Graha Ilmu, 2007.
- [12] S. R. Gunn, *Support Vector Machines For Classification and Regression*. Southampton: Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science, University of Southampton, 1998.
- [13] H. Drucker *et al.*, "Support Vector Regression Machines," *Bell Labs Monmouth Univ.*, vol. 1, pp. 155–161, 1997.
- [14] V. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. Springer-Verlag New York, 2000.
- [15] S. R. Gunn, "Support Vector Machines for classification and regression," *ISIS Tech. Rep.*, p. 14.1: 5-16., 1998.
- [16] R. . Walpole, *Pengantar Statistika*, 3rd ed. Jakarta: Gramedia Utama, 1995.
- [17] BPS (2020) ` Kemiskinan (P1) Menurut Provinsi, 2007 – 2020. <https://www.bps.go.id/linkTableDinamis/view/id/1121> (diakses 16 September 2020)
- [18] BPS, Rata-rata Pengeluaran per Kapita Sebulan Makanan dan Bukan Makanan di Daerah Perkotaan dan Perdesaan Menurut Provinsi (rupiah), 2019, https://www.bps.go.id/indikator/indikator/view_data_pub/0000/api_pub/155/da_13/1 (diakses 16 September 2020)
- [19] BPS, Rata-Rata Lama Sekolah Menurut Provinsi [Metode Baru], 2010-2019 <https://www.bps.go.id/dynamictable/2020/02/18/1773/rata-rata-lama-sekolah-menurut-provinsi-metode-baru-2010-2019.html> (diakses 16 September 2020)
- [20] BPS, Umur Harapan Hidup Saat Lahir (UHH) Menurut Provinsi [Metode Baru], 2010-2019 <https://www.bps.go.id/dynamictable/2018/04/16/1298/angka-harapan-hidup-saat-lahir-menurut-provinsi-2010-2017.html> (diakses 16 September 2020)