



Penerapan *Support Vector Machine* Dalam Peramalan Nilai Tukar Petani Provinsi Bengkulu

Yosep Oktavianus Sitohang¹

¹ BPS Provinsi Bengkulu

* Corresponding Author: yosepoktavianus@bps.go.id

Article Information

Article History:

Submitted: November, 28 2022

Accepted: December, 14 2022

Published: December, 31 2022

Key Words:

NTP

Peramalan

SVM

MAPE

Abstract

Pertanian masih menjadi sektor dominan di Provinsi Bengkulu, oleh sebab itu sasaran pembangunan disektor ini, khususnya kesejahteraan petani perlu mendapat perhatian lebih. Salah satu indikator kesejahteraan petani adalah Nilai Tukar Petani (NTP). Agar pembangunan tepat sasaran diperlukan perencanaan menggunakan kajian nilai NTP dimasa depan, sehingga metode peramalan dibutuhkan. Metode konvensional yang sering digunakan adalah ARIMA, namun metode ini memiliki banyak keterbatasan. Teknik *Machine Learning* yang berkembang di era *Society 5.0* menjadi alternatif solusi dari permasalahan tersebut. Salah satu metode terbarunya untuk peramalan data deret waktu adalah *Support Vector Machine (SVM)*. Dari hasil penghitungan didapat MAPE sebesar 8,55, termasuk dalam *highly accurate forecasting* sehingga SVM layak digunakan dalam peramalan.

1. PENDAHULUAN

Sektor pertanian masih merupakan sektor yang dominan di Provinsi Bengkulu. Tercermin dari jumlah penduduk usia 15 tahun keatas yang bekerja di sektor pertanian pada Agustus 2020 mencapai 48,88 (BPS, 2021). Sektor pertanian juga menyumbang diatas 28 persen dari Produk Domestik Regional Bruto (PDRB) Provinsi Bengkulu selama 5 tahun berturut-turut. Oleh karenanya, penting bagi Pemerintah Provinsi Bengkulu untuk tetap menjadikan sektor pertanian sebagai sasaran pembangunan, khususnya dalam hal peningkatan kesejahteraan petani.

Salah satu indikator dari kesejahteraan petani adalah Nilai Tukar Petani (NTP). NTP dapat diartikan sebagai pengukur kemampuan tukar produk pertanian yang dihasilkan petani dengan barang/jasa yang diperlukan untuk konsumsi rumah tangga maupun untuk keperluan dalam memproduksi produk pertanian (BPS, 2020).

Agar tujuan pembangunan dapat tepat sasaran, Pemerintah Provinsi Bengkulu harus mengkaji data NTP dimasa depan untuk melihat potensi keberhasilan sekaligus merencanakan pembangunan pertanian, sehingga dapat menjadi bahan pertimbangan pengambilan keputusan dalam pembangunan dimasa yang akan datang. Untuk itu diperlukan metode peramalan deret waktu.

Salah satu metode konvensional yang sering digunakan dalam peramalan deret waktu adalah metode *Autoregressive Integrated Moving Average (ARIMA)* yang diperkenalkan Box dan Jenkins (1976). Namun metode ini memiliki keterbatasan, yaitu harus memenuhi asumsi stasioner, residual bersifat *white noise* serta asumsi bahwa model yang terbentuk bersifat linier (Wei, 2006).

Teknik *Machine Learning* yang berkembang di era *Society 5.0* menjadi salah satu solusi dari permasalahan tersebut. Salah satu metode terbarunya untuk peramalan data deret waktu adalah *Support Vector Machine (SVM)*. Metode ini diperkenalkan oleh Vapnik (1995) dan memiliki banyak kelebihan, yaitu mampu mengatasi permasalahan *overfitting* dan selalu menghasilkan solusi yang terbaik (*global solution*). Metode ini juga tidak membutuhkan asumsi-asumsi seperti metode peramalan yang konvensional. Oleh karenanya metode ini

sering digunakan dalam melakukan peramalan data deret waktu (Cao, 2003; Sapankevych dan Sankar, 2009; Hidayat, 2013; Hermanianto, 2014; Mustakim, 2015).

Berdasarkan hal-hal diatas, maka metode SVM akan digunakan untuk meramalkan nilai NTP di Provinsi Bengkulu semester 2 tahun 2021.

2. METODE

2.1 Support Vector Machine (SVM)

Metode SVM merupakan salah satu metode machine learning terbaru yang diperkenalkan oleh Vapnik (1995). Dalam aplikasinya metode ini digunakan untuk keperluan klasifikasi dan regresi. Misalkan x adalah data deret waktu dan nilai yang akan diprediksi adalah y , sehingga terdapat $(x_1, y_1), \dots, (x_l, y_l), x \in R^m, y \in R$, maka persamaan regresi yang dapat terbentuk adalah:

$$y = f(x) + \varepsilon \quad (1)$$

Pada SVM, fungsi $f(x)$ dapat dinyatakan sebagai berikut:

$$f(x) = \omega \cdot \Phi(x) + b \quad (2)$$

ω merupakan penimbang (*weight*), b adalah bias dan fungsi $\Phi(x)$ merupakan fungsi tidak linier yang memetakan x kedalam ruang yang berdimensi tinggi (*high dimensional space*). Fungsi $f(x)$ terbaik didapatkan dengan meminimalkan risiko kesalahan. Menggunakan prinsip *Structural Risk Minimization* (SRM) dan ε -*Insensitive Loss Functions* didapat persamaan:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (3)$$

dimana α_i dan α_i^* merupakan pengali *Lagrange* yang didapat menggunakan *quadratic programming*. Fungsi $K(x', x)$ merupakan fungsi kernel yang berfungsi memetakan data kedalam suatu ruang dimensi yang lebih tinggi sehingga dapat dibentuk *hyperplane* yang optimal. Fungsi kernel yang digunakan didalam penelitian ini adalah *Gaussian Radial Basis Function* (RBF). Hal ini dikarenakan dari beberapa literatur yang ada, RBF memberikan performa yang lebih baik dibandingkan fungsi-fungsi kernel lainnya (Hidayat (2013), Hermanianto (2014), Mustakim (2015)). Adapun fungsi kernel dari RBF adalah sebagai berikut:

$$K(x', x) = \exp\left(-\frac{\|x' - x\|^2}{2\sigma^2}\right) \quad (4)$$

Metode SVM pada penelitian ini merujuk pada referensi Vapnik (1995). Oleh karenanya untuk lebih detail dapat melihat pada literatur tersebut.

2.2 Metode Grid Search

Metode ini merupakan salah satu metode umum yang sering digunakan untuk memperoleh estimasi dari parameter dengan hasil yang optimal (Rao, 2009). Cara kerja dari metode ini ialah menentukan beberapa titik grid parameter pada rentang tertentu, mengevaluasi fungsi $f(x)$ menggunakan seluruh titik grid dan menentukan titik grid yang memberikan akurasi terbaik. Adapun penerapannya pada penelitian ini dikombinasikan dengan metode *Cross Validation* (CV) yang khusus diaplikasikan pada data deret waktu (Hyndman, 2014) dengan tahapan:

1. Tentukan nilai k minimum dari observasi pada data *training* yang akan digunakan untuk membangun model.
2. Memilih observasi $k + i$ sebagai *test data* dan menghitung eror yang dihasilkan dari data peramalan terhadap *test data*.
3. Ulangi hingga $i = 0, 1, \dots, T - k$ dimana T adalah banyaknya data observasi.

4. Menghitung rata-rata dari keseluruhan eror yang didapat.

Adapun nilai k yang digunakan pada penelitian ini adalah 18.

2.3 Ukuran Akurasi

Ukuran akurasi yang digunakan adalah MSE dan MAPE. Adapun rumusnya sebagai berikut:

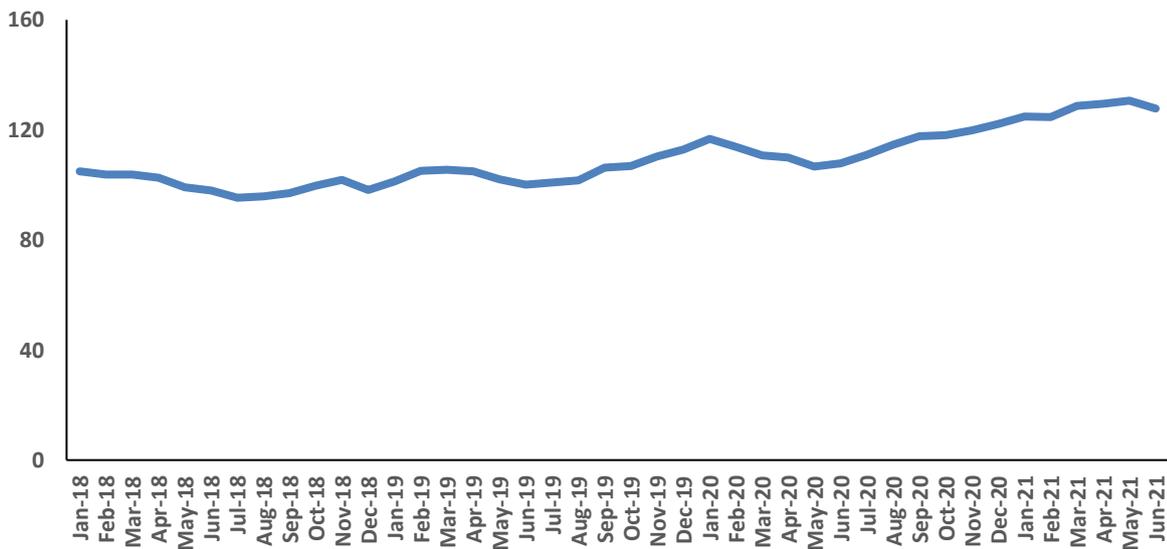
$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{5}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{6}$$

dimana y_i adalah nilai aktual, \hat{y}_i adalah nilai estimasi dan n adalah banyaknya data. MSE digunakan pada proses *grid search* sedangkan MAPE digunakan sebagai ukuran akurasi peramalan.

3. HASIL DAN PEMBAHASAN

Data yang digunakan adalah data NTP Provinsi Bengkulu periode Januari 2018 hingga Juni 2021 yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Bengkulu. Selanjutnya data dibagi menjadi 2, yaitu data *training*, menggunakan periode Januari 2018 hingga Desember 2020 dan data *testing* menggunakan data periode Januari 2021 hingga Juni 2021.



Gambar 1. Data NTP Provinsi Bengkulu Periode Januari 2018-Juni 2021

Adapun rentang nilai yang digunakan untuk parameter yang digunakan pada penelitian ini adalah {100, 200, 300} untuk C , {0.1,0.2,...,0.5} untuk ϵ dan {1,2,...,5} untuk σ . Alasan penggunaan rentang tersebut agar waktu pengolahan tidak terlalu lama.

Dari hasil pengolahan menggunakan metode *grid search* didapat nilai optimal untuk parameter $C = 300$, $\epsilon = 0.5$ dan $\sigma = 1$. Dengan menggunakan nilai tersebut, adapun nilai MAPE untuk peramalan data testing adalah sebesar 8,55. Adapun menurut Lewis (1982), nilai MAPE tersebut dapat diinterpretasikan sebagai *highly accurate forecasting*.

Adapun hasil peramalan nilai NTP Provinsi Bengkulu semester 2 tahun 2021 adalah sebagai berikut :

Tabel 1. Peramalan Nilai NTP Provinsi Bengkulu Semester 2 Tahun 2021

Bulan	NTP
Juli	125,47
Agustus	124,19
September	123,10
Oktober	122,09
November	121,08
Desember	120,06

Sumber: Data yang diolah

4. SIMPULAN

Kesimpulan yang diperoleh adalah metode SVM sangat baik dalam melakukan peramalan NTP Provinsi Bengkulu, dibuktikan dari nilai MAPE 8,55. Adapun dari hasil peramalan terlihat bahwa nilai NTP disemester 2 mengalami penurunan. Pemerintah daerah perlu mengambil langkah-langkah untuk mengantisipasinya.

Untuk penelitian selanjutnya perlu untuk mereduksi noise yang dikandung oleh data deret waktu, sehingga diharapkan hasil peramalan dapat semakin akurat.

TERIMA KASIH

Ucapan terima kasih kepada Bapak Win Rizal selaku Kepala BPS Provinsi Bengkulu serta seluruh pihak yang membantu dalam penyusunan riset ini.

REFERENSI

- Box, G.E.P. dan Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- BPS. (2021). *Provinsi Bengkulu Dalam Angka 2021*. Bengkulu.
- BPS. (2020). *Perkembangan Nilai Tukar Petani Provinsi Bengkulu 2019*. Bengkulu.
- Cao, L. (2003). *Support Vector Machines Experts for Time Series Forecasting*. *Neurocomputing* 51: 321-339.
- Hidayat, R. (2013). *Sistem Prediksi Status Gizi Balita dengan Menggunakan Support Vector Regression*. Skripsi: Institut Pertanian Bogor.
- Hermanianto, A.,B. (2014). *Prediksi Panjang Musim Hujan Menggunakan Support Vector Regression*. Skripsi: Institut Pertanian Bogor.
- Hyndman, Rob. (2014). *Forecasting: Principles & Practice*. Robjhyndman.com/uwa
- Lewis, C.D. (1982). *Industrial and Business Forecasting Methods*. London: Butterworths.
- Mustakim, Buono, A, dan Hermadi, I. (2015). *Support Vector Regression Untuk Prediksi Produktivitas Kelapa Sawit di Provinsi Riau*. *Jurnal Sains, Teknologi dan Industri* 12(2): 179-188.
- Rao, S.,S. (2009). *Engineering Optimization: Theory and Practice*, fourth edition. J Wiley, New Jersey (US).
- Sapankevych, N dan Sankar, R. (2009). *Time Series Prediction Using Support Vector Machines: A Survey*. *IEEE Computational Intelligence Magazine* 4(2): 24-30.
- Vapnik, V.N. (1995). *The Nature of Statistical Learning Theory*. New York: Springer.
- Wei, W.W.S. (2006). *Time Series Analysis Univariate and Multivariate Methods; Second Edition*. USA: Pearson Addison Wesley.