

## Prediksi Pasien Covid-19 yang Meninggal di Jawa Timur Menggunakan Metode Regresi Spline Truncated

Qurrotul Aini<sup>1</sup>, Tony Yulianto<sup>2</sup>, Faisol<sup>3</sup>

<sup>1</sup>Universitas Islam Madura, [gurrotul188@gmail.com](mailto:gurrotul188@gmail.com)

<sup>2</sup>Universitas Islam Madura, [toniyulianto65@gmail.com](mailto:toniyulianto65@gmail.com)

<sup>3</sup>Universitas Islam Madura, [faisol.munif@gmail.com](mailto:faisol.munif@gmail.com)

DOI 10.31102/zeta.2022.7.2.79-83

### ABSTRACT

*Corona Virus Disease 2019 (COVID-19) was first discovered at the end of 2019 in Wuhan, China. The virus attacks the respiratory system with pneumonia-like symptoms. This virus is classified as a new virus so it does not have an antidote and has spread throughout the world until it is out of control. There are more than 200 countries reporting covid-19 cases including Indonesia. Because the biggest effect of COVID-19 is death, so researchers need to predict whether the mortality rate of COVID-19 patients. One method for predicting covid-19 patients who died in East Java is using the spline truncated regression method. Spline regression is a regression analysis that is able to estimate data that does not have a specific pattern and has a tendency to seek data estimates from the formed pattern on its own, while truncated is a function that can be interpreted as a function of slice. And produces a GCV value of 51.0167, a MAPE value 47.0096, MSE 32.9825, and R-Sq 0.9998. Based on the MAPE value, the prediction model is said to be still reasonable.*

**Keywords:** *COVID-19, spline truncated regression, positive patients, cured patients, deceased patients.*

### ABSTRAK

*Corona Virus Disease 2019 (COVID-19) pertama kali diketahui pada akhir tahun 2019 di Wuhan, China. Virus ini menyerang sistem pernapasan dengan gejala seperti pneumonia. Virus ini yang tergolong virus baru sehingga belum memiliki penawar dan sudah menyebar ke seluruh penjuru dunia hingga tak terkendali. Ada lebih dari 200 negara yang melaporkan adanya kasus COVID-19 ini termasuk Indonesia. Karena efek terbesar COVID-19 adalah kematian, jadi peneliti perlu memprediksi apakah tingkat kematian pasien COVID-19. Salah satu metode untuk memprediksi pasien covid-19 yang meninggal di Jawa Timur adalah menggunakan metode regresi Spline Truncated. Regresi Spline merupakan analisis regresi yang mampu mengestimasi data yang tidak memiliki pola tertentu dan memiliki kecenderungan dalam mencari sendiri estimasi data dari pola yang terbentuk, sedangkan truncated merupakan sebuah fungsi yang dapat diartikan sebagai fungsi potongan. Dan menghasilkan nilai GCV 51.0167, nilai MAPE sebesar 47.0096, MSE 32.9825, dan R-Sq 0.9998. berdasarkan nilai MAPE tersebut model prediksi dikatakan masih dalam tingkat kewajaran.*

**Kata Kunci:** *COVID-19, regresi spline truncated, pasien positif, pasien sembuh, pasien meninggal.*

## 1. PENDAHULUAN

*Corona Virus Disease 2019* (COVID-19) pertama kali diketahui pada akhir tahun 2019 di Wuhan, China. Virus ini menyerang sistem pernapasan dengan gejala seperti *pneumonia*. Virus ini yang tergolong virus baru sehingga belum memiliki penawar dan sudah menyebar ke seluruh penjuru dunia hingga tak terkendali. Ada lebih dari 200 negara yang melaporkan adanya kasus COVID-19 ini termasuk Indonesia. Dalam kondisi saat ini, COVID-19 yang telah ditetapkan oleh WHO (*World Health Organization*) sebagai pandemi. Per 12 April 2020, terdapat setidaknya 1.775.210 kasus dengan jumlah kematian mencapai 108.544 dan sembuh sebanyak 401.517 jiwa. Dan juga total kasus orang yang terinfeksi Corona di Indonesia mencapai 4.241 dengan jumlah pasien yang sembuh sebanyak 359 dan kematian sebanyak 373 hingga tanggal 12 April 2020. Dengan skala kasus yang massif tersebut, COVID-19 ditetapkan sebagai epidemik bencana nasional (Albana & Azhari, 2020).

Kasus yang terjadi di Indonesia memang tidak sebanyak negara yang lain, namun terdapat kemungkinan bahwa akan terjadi kenaikan terus menerus dalam beberapa waktu dekat. Tingginya tingkat kematian akibat virus corona ini disebabkan oleh beberapa faktor, diantaranya seperti penyakit bawaan yang telah dialami dan kurangnya kasadaran dari masing-masing individu terhadap virus ini, kurang mentaati peraturan pemerintah untuk mematuhi protokol kesehatan dan faktor eksternal seperti fasilitas rumah sakit yang kurang memadai (Ilpaj & Nurwati, 2020).

Efek terbesar COVID-19 adalah kematian. Oleh karena itu, peneliti perlu memprediksi apakah tingkat kematian pasien covid-19 mengalami peningkatan atau penurunan. Jika ada peningkatan maka harus dilakukan upaya-upaya pencegahan untuk menurunkan peningkatan kematian tersebut, jika mengalami penurunan berarti kebijakan pemerintah sangat efektif. Salah satu metode untuk memprediksi pasien covid-19 yang meninggal di Jawa Timur adalah menggunakan metode regresi *Spline Truncated*.

Regresi *Spline* merupakan analisis regresi yang mampu mengestimasi data yang tidak memiliki pola tertentu dan memiliki kecenderungan dalam mencari sendiri estimasi data dari pola yang terbentuk. Salah satu kelebihan pendekatan *Spline* adalah model ini cenderung mencari sendiri estimasi data kemanapun pola data tersebut bergerak. Kelebihan ini terjadi karena dalam *Spline* terdapat titik-titik knot, yaitu titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data. *Truncated* merupakan sebuah fungsi yang dapat diartikan sebagai fungsi potongan (Fadhilah, dkk, 2016).

Penelitian menggunakan regresi *Spline Truncated* sebelumnya telah dilakukan oleh

Khoirunnisa Nur Fadhilah, Suparti, Tarno (2016) berjudul “Pemodelan Regresi *Spline Truncated* Untuk Data Longitudinal” yang menghasilkan Model spline truncated data longitudinal terbaik untuk harga saham bulanan adalah model regresi orde 2 dengan 3 titik knot. Nilai MAPE untuk subyek 1 (Bank Mandiri) terletak di antara 20% hingga 50%, maka dapat dikatakan bahwa model pada subyek 1 cukup untuk digunakan untuk melakukan prediksi ke depan. Nilai MAPE untuk subyek 2 (Bank Bukopin) dan subyek 3 (Bank Bumi Arta) tersebut terletak di antara 10% hingga 20%, maka dapat dikatakan bahwa model pada subyek 2 dan subyek 3 baik digunakan untuk melakukan prediksi ke depan. Dan pada penelitian lainnya dilakukan oleh Edy Widodo, Adisti Nurul Irmayanti (2019) berjudul “Perbandingan Metode Regresi *Spline Truncated* Dengan Regresi Linear Sederhana Untuk Kasus Harga Saham Perusahaan Pertambangan di Indonesia” yang menghasilkan model terbaik untuk pemodelan kasus harga saham perusahaan tambang adalah menggunakan regresi metode Regresi *Spline Truncated*, dengan nilai MAPE yang dihasilkan 7,81% untuk ADRO.JK, 8,51% untuk ANTM.JK, dan 12,74% untuk ITMG.JK.

## 2. TINJAUAN PUSTAKA

Corona virus adalah keluarga besar virus yang menyebabkan penyakit mulai dari gejala ringan sampai berat. Ada setidaknya dua jenis coronavirus yang diketahui menyebabkan penyakit yang dapat menimbulkan gejala berat seperti *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS). *Corona Virus Disease 2019* (COVID-19) adalah penyakit jenis baru yang belum pernah diidentifikasi sebelumnya pada manusia. Virus penyebab Covid-19 ini dinamakan Sars-CoV-2. COVID-19 adalah zoonosis (ditularkan antara hewan dan manusia). Penelitian menyebutkan bahwa SARS yang menjadi sumber penularan Covid-19 ini masih belum diketahui. Tanda dan gejala umum infeksi Covid-19 antara lain gejala gangguan pernapasan akut seperti demam, batuk dan sesak napas. Masa inkubasi rata-rata 5-6 hari dengan masa inkubasi terpanjang 14 hari. Pada kasus Covid-19 yang berat dapat menyebabkan pneumonia, sindrom pernapasan akut, gagal ginjal, dan bahkan kematian. Tanda-tanda dan gejala klinis yang dilaporkan pada sebagian besar kasus adalah demam, dengan beberapa kasus mengalami kesulitan bernapas, dan hasil rontgen menunjukkan infiltrat pneumonia luas di kedua paru (Yuliawati & Djannah, 2020).

Kasus yang terjadi di Indonesia memang tidak sebanyak negara yang telah penulis sampaikan diatas, namun terdapat kemungkinan bahwa akan terjadi kenaikan terus menerus dalam beberapa waktu dekat. Banyaknya jumlah kematian yang bertambah hari demi hari akibat virus corona ini tidak hanya menimbulkan gejala dan penyakit fisik saja akan

tetapi, berpengaruh besar terhadap kesejahteraan masyarakat Indonesia yang didalamnya mencakup kesehatan mental. Seperti berdampak pada kondisi sosial ekonomi keluarga yang ditinggalkan, hal ini dapat berpengaruh secara signifikan apabila seseorang yang terjangkit virus corona lalu meninggal dunia adalah tulang punggung dalam keluarganya. Lalu pengaruhnya terhadap masyarakat yaitu membuat menjadi lebih mudah panik, cemas dan stress. Rasa cemas atau khawatir secara berlebihan karena terlalu banyak menerima informasi tersebut yang akhirnya menyebabkan tubuh menciptakan gejala mirip coronavirus. Padahal gejala tersebut hanya perwujudan dari rasa cemas berlebihan, bukan terinfeksi coronavirus. Kondisi seperti itu dikenal dengan istilah psikosomatik akibat virus corona. Adanya peraturan pemerintah yang semakin ketat seperti *physical distancing* untuk mencegah tingginya penyebaran virus corona ini tidak bisa dipungkiri dapat menyebabkan kesehatan mental yang kurang baik. Hal ini termasuk kedalam pengaruh tingginya tingkat kematian di Indonesia (Ilpaj & Nurwati, 2020).

Metode regresi nonparametrik merupakan metode regresi yang digunakan ketika kurva regresi antara variabel respon dan variabel prediktor tidak diketahui bentuknya atau polanya. Regresi nonparametrik *Spline* memiliki fleksibilitas yang tinggi dimana data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh subjektifitas peneliti. (Bintariningrum & Budiantara, 2014).

Regresi *Spline Truncated* merupakan suatu pendekatan ke arah pengepasan data dengan tetap memperhitungkan kemulusan kurva. *Spline* merupakan model polinomial yang tersegmen. Sifat tersegmen akan memberikan fleksibilitas yang baik. Sifat ini memungkinkan model Regresi *spline* menyesuaikan diri secara efektif terhadap karakteristik data. Pada data longitudinal terdapat  $i = 1, 2, \dots, m$  subjek dan  $j = 1, 2, \dots, n_i$  pengamatan dalam setiap subjek, maka fungsi *Spline Truncated* berorde  $p$  dengan titik knot  $K = \{K_1, K_2, \dots, K_r\}$  (Widodo & Irmayanti, 2019)

### 3. METODE PENELITIAN

Pada tahapan ini diuraikan beberapa tahap penelitian yang akan digunakan untuk mencapai tujuan penelitian. Beberapa tahap penelitian sebagai berikut:

#### 1. Studi Literatur

Tahap ini akan melakukan identifikasi permasalahan, yaitu mencari referensi berupa jurnal dan buku yang berkaitan dengan metode regresi *spline truncated* untuk dijadikan landasan.

#### 2. Pengambilan Data

Data yang diambil dalam penelitian skripsi ini adalah data sekunder dari pasien COVID-19 yang positif, sembuh dan meninggal di Jawa Timur.

### 3. Pengolahan Data Statistik Deskriptif

Data yang sudah diambil diolah dan disesuaikan dengan variabel yang dibutuhkan. Dan variabel yang dibutuhkan adalah pasien COVID-19 yang positif, sembuh dan meninggal. Adapun variabel-variabel dari studi kasus peneliti adalah:

$$\begin{aligned}y &= \text{Pasien Meninggal} \\x_1 &= \text{Pasien Positif} \\x_2 &= \text{Pasien Sembuh}\end{aligned}$$

#### 4. Penerapan Regresi *Spline Truncated*

Dari data yang sudah ada di pengolahan data akan diterapkan menggunakan metode Regresi *Spline Truncated*. Untuk mendapatkan *output* (hasil), maka terdapat 9 langkah/tahapan sebagai berikut:

- a. Memasukkan data variabel respon dan variabel prediktor.
- b. Pemilihan data *in sample* dan *out sample*.
- c. Membuat *scatterplot* antara variabel respon dan variabel prediktor data *in sample*.
- d. Menentukan orde dan banyaknya titik knot.
- e. Menentukan model Regresi *Spline Truncated* untuk setiap orde dan titik knot.
- f. Menghitung GCV (*Generalized Cross Validation*) untuk setiap orde dan titik knot.
- g. Menentukan titik knot optimal dengan GCV (*Generalized Cross Validation*) minimum.
- h. Membentuk model *Spline* terbaik berdasarkan titik knot optimal.
- i. Menghitung ketepatan peramalan model terbaik menggunakan MAPE pada data *out sample*.

#### 5. Simulasi

Simulasi pada penelitian ini mengaplikasikan permasalahan dalam metode regresi *spline truncated*.

#### 6. Validasi Hasil Simulasi

Untuk memvalidasi hasil simulasi menggunakan metode regresi *spline truncated* dengan hasil real yang ada di lapangan.

#### 7. Penarikan Kesimpulan

Pada tahap ini, merupakan tahapan terakhir dalam penyelesaian penelitian. Setelah penelitian mendapatkan hasil dari metode regresi *spline truncated* dalam memprediksi pasien covid-19 yang meninggal di Jawa Timur. Maka ditarik kesimpulan dan saran penelitian ini.

## 4. HASIL PENELITIAN

### 4.1 Pemilihan Titik Knot Optimal

Pada regresi *spline truncated*, sangat penting untuk menentukan titik knot optimal. Jika diperoleh titik knot optimal maka akan diperoleh model *spline truncated* terbaik. Pemilihan titik knot optimal dilakukan dengan melihat nilai Generalized Cross Validation (GCV). Lalu percobaan dilakukan sampai memperoleh nilai GCV paling minimum. Terdapat beberapa langkah yang dilakukan sebelum mendapatkan model *spline* terbaik berdasarkan GCV minimum, yaitu dengan membandingkan nilai GCV

antara model *spline* linier dengan 1 titik knot sampai 100 titik knot. Ketika penambahan titik knot menghasilkan GCV lebih besar, maka penambahan titik knot tidak optimal.

#### 4.2 Interpretasi Model Terbaik

Dari hasil pemilihan titik knot optimum yang dilakukan, maka model terbaik yang terbentuk adalah regresi *spline* linear dengan 13 titik knot yaitu sebagai berikut:

$$\hat{y} = -3.7699 + 0.16259 \cdot x_1 - 0.014566(x_1 - 6835.722) - 0.038757(x_1 - 17385.3037) - 0.082044(x_1 - 2610.7108) - 0.068452(x_1 - 734.8661) + 0.1057(x_1 - 22802.8324) + 0.087295(x_1 - 3317.9406) + 0.045275(x_1 - 17930.1844) - 0.064429(x_1 - 18771.2949) - 0.040321(x_1 - 4873.8593) - 0.29907 \cdot x_2 - 0.089745(x_2 - 15415.9449) + 0.28814(x_2 - 119.8397) + 0.036919(x_2 - 9214.7685) + 0.048055(x_2 - 975.865) + 0.025162(x_2 - 2485.5234) - 0.049357(x_2 - 5693.2885) - 0.014149(x_2 - 11611.8513)_+$$

Berdasarkan hasil uji parameter didapat bahwa  $x_1$  dan  $x_2$  berpengaruh signifikan terhadap  $y$ . Koefisien dari  $x_1$  dan  $x_2$  bernilai positif. Hal ini berarti terdapat hubungan yang positif antara  $x_1$  maupun  $x_2$  terhadap  $y$ .

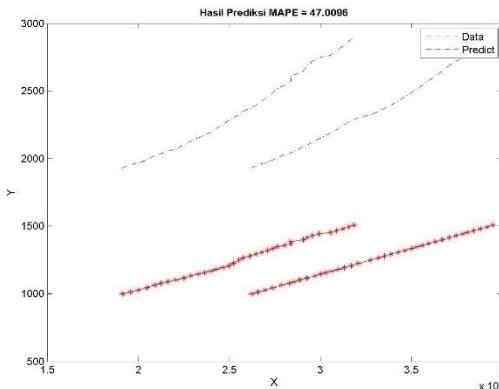
Ketika  $x_2$  konstan maka penambahan  $x_1$  sebesar satu satuan akan menyebabkan penambahan  $y$  sebesar 0.16259.

Ketika  $x_1$  konstan maka penambahan  $x_2$  sebesar satu satuan akan menyebabkan penambahan  $y$  sebesar -0.29907.

Artinya jika pasien yang positif bertambah maka kemungkinan akan menyebabkan penambahan pasien yang meninggal sebesar 0.16259. dan jika pasien yang sembuh semakin bertambah maka kemungkinan akan menyebabkan menurunan pasien yang meninggal sebesar -0.29907.

#### 4.3 Hasil Prediksi

Berdasarkan model terbaik di atas maka didapat hasil diprediksi pada Gambar 1, berdasarkan gambar tersebut terlihat bahwa hasil prediksi pasien Covid-19 merupakan model yang baik dengan nilai GCV 51.0167, nilai MAPE sebesar 47.0096, MSE 32.9825, dan R-Sq 0.9998. berdasarkan nilai MAPE tersebut model prediksi dikatakan masih dalam tingkat kewajaran.



Gambar 1. Hasil Prediksi Pasien Covid-19 di Jawa Timur

### 5. KESIMPULAN

#### 5.1 Kesimpulan

Kesimpulan yang diperoleh berdasarkan pembahasan sebelumnya adalah di dapat model yang terbaik dalam regresi *spline truncated* untuk memprediksi pasien Covid-19 yang meninggal di Jawa Timur adalah sebagai berikut:

$$\hat{y} = -3.7699 + 0.16259 \cdot x_1 - 0.014566(x_1 - 6835.722) - 0.038757(x_1 - 17385.3037) - 0.082044(x_1 - 2610.7108) - 0.068452(x_1 - 734.8661) + 0.1057(x_1 - 22802.8324) + 0.087295(x_1 - 3317.9406) + 0.045275(x_1 - 17930.1844) - 0.064429(x_1 - 18771.2949) - 0.040321(x_1 - 4873.8593) - 0.29907 \cdot x_2 - 0.089745(x_2 - 15415.9449) + 0.28814(x_2 - 119.8397) + 0.036919(x_2 - 9214.7685) + 0.048055(x_2 - 975.865) + 0.025162(x_2 - 2485.5234) - 0.049357(x_2 - 5693.2885) - 0.014149(x_2 - 11611.8513)_+$$

Model tersebut merupakan model terbaik dengan 13 titik knot yang menghasilkan nilai GCV 51.0167, nilai MAPE sebesar 47.0096, MSE 32.9825, dan R-Sq 0.9998. berdasarkan nilai MAPE tersebut model prediksi dikatakan masih dalam tingkat kewajaran.

#### 5.2 Saran

Saran yang dapat diusulkan dalam penelitian ini adalah:

- Peneliti dapat mengembangkan lebih lanjut dengan permasalahan yang berbeda dan dapat menggunakan metode lainnya yang lebih sesuai dengan kriteria regresi *spline truncated*.
- Peneliti dapat mengembangkan lebih lanjut dengan variabel yang lebih banyak misalkan menambahkan variabel respon dan prediktor.

## DAFTAR PUSTAKA

- Albana, A. S., & Azhari, S. (2020). Prediksi Penyebaran COVID-19 Kota Surabaya Dengan Simulasi Monte Carlo. *Journal of Advances in Information and Industrial Technology*, 2(1), 36-42.
- Fadhillah, K. N., Suparti, , & Tarno, . (2016). Pemodelan Regresi Spline Truncated Untuk Data Longitudinal. *Jurnal Gaussian*, 5(3), 447-454.
- Ginting, F., Buulolo, E., & Siagian, E. R. (2019). Implementasi Algoritma Regresi Linear Sederhana Dalam Memprediksi Besaran Pendapatan Daerah (Studi Kasus: Dinas Pendapatan Kab. Deli Serdang). *Konferensi Nasional Teknologi Informasi dan Komputer*, 3(1), 274-279.
- Handayani, D., Hadi, D. R., Isbaniah, F., Burhan, E., & Agustin, H. (2020). Penyakit virus Corona 2019. *Jurnal Respirologi Indonesia*, 40(2), 119-129.
- Ilpa, S. M., & Nurwati, N. (2020). Analisis Pengaruh Tingkat Kematian Akibat Covid-19 Terhadap Kesehatan Mental Masyarakat Di Indonesia. *Jurnal Pekerjaan Sosial*, 3(1), 16-28.
- Katemba, P., & Djoh, R. K. (2017). Prediksi Tingkat Produksi Kopi Menggunakan Regresi Linear. *Jurnal Ilmiah Flash*, 3(1), 42-51.
- Riskiyah, S. (2017). *Penerapan Regresi Nonparametrik Spline Truncated Untuk Mengetahui Faktor-faktor yang Mempengaruhi Hasil Tangkapan Ikan Nelayan di Kabupaten Pamekasan*. Pamekasan: Universitas Islam Madura.
- Rory, . (2016). *Regresi Campuran Nonparametrik Spline Linear Truncated dan Fungsi Kernel Untuk Pemodelan Data Kemiskinan di Provinsi Papua*. Surabaya.
- Widodo, E., & Irmayanti, A. N. (2019). Perbandingan Metode Regresi Spline Truncated dengan Regresi Linear Sederhana untuk Kasus Harga Saham Perusahaan Pertambangan di Indonesia. *Eksakta*, 19(02), 143-53.
- Yulianto, T., Faisol, F., Zahroh, F., Suryanti, S., & Tafrikan, M. (2021). Forecasting the recovery of COVID-19 patients in East Java using the Fuzzy time series Cheng method. *Journal of Natural Sciences and Mathematics Research*, 7(1), 44-50. doi:<https://doi.org/10.21580/jnsmr.2021.7.1.11289>
- Yuliawati, K., & Djannah, S. N. (2020). Bagaimana Pengetahuan, Sikap Dan Perilaku Masyarakat Tentang Konsumsi Multivitamin/Suplemen selama Pandemi Covid-19. *Jurnal Kesehatan Masyarakat Khatulistiwa*, 7(3), 123-134.
- Zahrah, M. . (2020). *Prediksi Tingkat Hunian Hotel di Kabupaten Pamekasan Berdasarkan Estimator Deret Fourier Dengan dan Tanpa tren*. Universitas Islam Madura.