

BAB I

PENDAHULUAN

1.1 Latar Belakang

Saat ini dunia sedang dilanda pandemi yang sangat mempengaruhi sektor kesehatan. *Severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2) atau yang lebih dikenal masyarakat luas dengan sebutan corona, merupakan virus yang dapat menular dari satu individu ke individu lainnya. Virus corona ini sendiri dapat menyerang berbagai golongan usia dan dapat memberikan dampak yang lebih besar pada sebagian individu yang mengidap penyakit tertentu (Pane, 2020).

Kasus pertama COVID-19 di Indonesia tercatat pada tanggal 2 Maret 2020 (Putri et al., 2020). Terhitung sejak awal pertama masuknya COVID-19 di Indonesia, terdapat peningkatan kasus positif yang tergolong sangat signifikan. Pada tanggal 31 Agustus 2020 kasus positif COVID-19 di Indonesia mencapai angka sebesar 174.554 kasus terkonfirmasi (Satgas COVID-19, 2021).

Guna memutus rantai penyebaran COVID-19 di Indonesia, beberapa pemerintah daerah mengeluarkan kebijakan-kebijakan seperti pembatasan aktivitas di luar rumah, ibadah yang dilakukan dirumah saja, pembelajaran yang dilakukan daring dari rumah, dan lain sebagainya (Sari et al., 2020). Dengan tingkat penyebaran dan kasus positif yang terus meningkat, masih belum dipastikan kapan tepatnya titik tertinggi kasus positif COVID-19 di Indonesia. Menurut Wiku Adisasmito (2020), pada wawancaranya bersama BBC News Indonesia, puncak pandemi masih belum dapat diprediksi dikarenakan kasus yang sangat dinamis

dengan perilaku masyarakat Indonesia, sehingga dibutuhkan penelitian yang dapat menghitung maupun menghasilkan data prediksi puncak kasus COVID-19 di Indonesia. Hasil dari penelitian tersebut dapat membantu bagi peneliti maupun ahli lainnya dalam penelitian atau rancangan yang dapat membantu mengurangi maupun mencegah penyebaran virus COVID-19 di Indonesia.

Menurut Kamus Besar Bahasa Indonesia (KBBI) Online, “Prediksi adalah ramalan atau prakiraan”. Dalam melakukan prediksi terdapat metode *machine learning* yang biasa digunakan untuk melakukan prediksi, seperti misalnya *Support Vector Regression (SVR)*. Dalam penelitian ini, akan menggunakan metode *Support Vector Regression (SVR)*. SVR digunakan untuk memecahkan kasus regresi dan dataset *time series*, dengan mencari persamaan untuk menghasilkan nilai kesalahan paling kecil (Suranart et al, 2014). SVR dapat melakukan prediksi dengan akurasi yang lebih baik dibandingkan dengan metode regresi lain (Suranart et al., 2014). SVR dapat diterapkan pada data yang kontinu (Agustina et al., 2018). Kelebihan dari metode SVR adalah dapat memanfaatkan model *nonlinear* secara implisit melalui penerapan fungsi kernel yang memetakan vektor ke dimensi yang lebih tinggi, selain itu SVR dapat mengatasi masalah *overfitting* lebih baik dibandingkan dengan model regresi biasa (Maharesi, 2013).

Penelitian yang dilakukan oleh Khanoksin Suranart, Supaporn Kiattisin dan Adisorn Leelasntitham (Suranart et al, 2014), mengenai perbandingan prediksi harga emas menggunakan *Neural Network (NN)*, *Radial Basis Function Netowk (RBFN)* dan *Support Vector Regression (SVR)* menunjukkan bahwa SVR memiliki hasil yang lebih dibandingkan dengan 2 metode lainnya dengan nilai *Mean Absolute*

Percentage Error (MAPE) sebesar 0.908 dan *Mean Square Error* (MSE) sebesar 210.401, sedangkan metode *Radial Basis Function* (RBF) memiliki nilai MAPE sebesar 3.589 dan nilai MSE sebesar 3341.428 dan untuk metode *Neural Network* (NN) memiliki nilai MAPE sebesar 5.756 dan nilai MSE sebesar 6991.224. Angka tersebut menunjukkan bahwa metode SVR memiliki akurasi lebih baik dibandingkan dengan 2 metode lainnya.

Penelitian mengenai prediksi sudah dilakukan oleh Olivia Bonita, Lailil Muflikhah dan Ratih Kartika Dewi (Bonita et al., 2018) dalam penelitiannya melakukan prediksi harga batu bara dengan menggunakan metode *Support Vector Regression* (SVR). Penggunaan metode SVR pada penelitian ini adalah untuk memprediksi harga batu bara dimana SVR diterapkan melalui tahapan normalisasi data, menghitung matriks hessian, pencarian nilai (α) dengan *sequential learning*, dan perhitungan fungsi regresi. Dengan menggunakan kernel Gaussian RBF dan kernel ANOVA untuk dapat dianalisis pengaruhnya supaya dapat menentukan akurasi hasil prediksi. Hasil yang diperoleh oleh kedua kernel dilakukan pengujian dan evaluasi dengan *Mean Absolute Percentage Error* (MAPE) untuk memperoleh hasil prediksi yang baik. Pengujian memberikan rata-rata nilai MAPE sebesar 9,64% dengan kernel Gaussian RBF dan 8,38% dengan kernel ANOVA, hasil ini dapat dikategorikan sangat baik menurut peneliti.

Penelitian mengenai prediksi lainnya juga pernah dilakukan oleh Muhammad Nasrul Tsalatsa Putra, Agus Eko Minarno dan Setio Basuki (Tsalatsa Putra et al., 2020). Penelitian ini dilakukan karena beberapa kendala pada berkas rekam medis yang dikarenakan beberapa faktor seperti kurangnya pengawasan,

informasi, dan tidak meratanya pemberian layanan kesehatan serta beberapa kendala lainnya. Penelitian yang dilakukan adalah melakukan prediksi diagnosa berdasarkan data rekam medis pasien. Dalam penelitian tersebut peneliti menggunakan metode *Support Vector Regression* (SVR). Kernel yang digunakan yaitu kernel *Linear*, kernel *Polynomial*, serta kernel *Radial Basis Function* (RBF). Pengujian dilakukan menggunakan *Normalized Root Mean Squared Error* (NRMSE) dan menunjukkan hasil untuk kernel *Linear* sebesar 0.4167, kernel *Polynomial* sebesar 0.4281 dan kernel RBF sebesar 0.0797. Hasil prediksi yang dilakukan mendapatkan hasil yang dapat menunjukkan bahwa ketepatan akurasi dari ketiga kernel tersebut sangat baik.

Berdasarkan data dari penelitian-penelitian sebelumnya, telah dibuktikan bahwa penggunaan metode *Support Vector Regression* (SVR) untuk melakukan prediksi memiliki reputasi akurasi yang sangat baik. Dalam melakukan prediksi dibutuhkan fungsi kernel untuk memetakan data ke ruang fitur agar data dapat diproses dengan lebih terstruktur. Berdasarkan penelitian sebelumnya yang dilakukan oleh (Olivia, et al., 2018) menggunakan SVR dan kernel ANOVA serta RBF menunjukkan hasil MAPE sebesar 9.64% untuk kernel ANOVA serta 8.37% untuk kernel RBF. Pada penelitian ini akan membahas bagaimana SVR dengan kernel *Polynomial*, *Linear*, serta RBF dengan data kasus COVID-19 yang tercatat dari bulan Agustus 2020 hingga bulan Maret 2021. Penelitian akan dilakukan dengan data COVID-19 di kota Tangerang, pemilihan kota Tangerang dikarenakan peningkatan kasus di Kota Tangerang menjadi salah satu yang tercepat dan juga diharapkan dapat memperkirakan angka kasus COVID-19 yang akan datang.

Diharapkan penelitian ini dapat melakukan perhitungan performa masing-masing kernel dan menunjukkan kernel yang memiliki performa terbaik dalam melakukan prediksi titik puncak COVID-19 di kota Tangerang.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan diatas, maka rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana memprediksi titik puncak COVID-19 di kota Tangerang menggunakan metode *Support Vector Regression* (SVR) dengan kernel *Linear*, *Polynomial*, dan *Radial Basis Function* (RBF)?
2. Bagaimana cara mengukur hasil yang dihasilkan menggunakan perhitungan *Mean Absolute Percentage Error* (MAPE), *Mean Absolute Error* (MAE), dan *Mean Squared Error* (MSE)?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini dapat dijabarkan menjadi sebagai berikut:

1. Batasan area yang dijadikan objek penelitian adalah wilayah yang berada di kota Tangerang, Indonesia.
2. Sampel data pada penelitian ini menggunakan data yang tercatat sejak bulan Agustus 2020 sampai dengan bulan Maret 2021. [sumber: <https://infocorona.bantenprov.go.id/kasus-konfirmasi>]

1.4 Tujuan Penelitian

Tujuan dari penelitian ini berdasarkan rumusan masalah yang telah dipaparkan adalah sebagai berikut:

1. Menerapkan metode *Support Vector Regression* (SVR) untuk melakukan prediksi titik puncak COVID-19 di Kota Tangerang dengan menggunakan kernel *Linear*, *Polynomial*, dan *Radial Basis Function* (RBF).
2. Melakukan evaluasi terhadap hasil prediksi menggunakan tiga metode yaitu, *Mean Absolute Percentage Error* (MAPE), *Mean Absolute Error* (MAE), dan *Mean Squared Error* (MSE).

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini dijabarkan sebagai berikut:

1. Memberikan ilmu pengetahuan tentang algoritma *Support Vector Regression* (SVR) dengan kernel *Linear*, *Polynomial* dan *Radial Basis Function* (RBF) pada data kasus COVID-19 di kota Tangerang.
2. Menunjukkan hasil evaluasi *Mean Absolute Percentage Error* (MAPE) paling baik dan memberikan gambaran titik puncak COVID-19 di kota Tangerang.

1.6 Sistematika Penelitian

Sistematika penulisan pada penulisan skripsi ini terdiri dari lima bab, yaitu:

1. BAB 1 PENDAHULUAN

Bab ini berisikan latar belakang, rumusan masalah, Batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

2. BAB 2 LANDASAN TEORI

Bab ini berisikan landasan yang menjadi teori dalam penelitian yang terdiri dari penjelasan mengenai COVID-19, regresi, *Support Vector Regression*, kernel, MAPE, MAE dan MSE.

3. BAB 3 METODOLOGI PENELITIAN

Bab ini membahas metode penelitian yang digunakan, perancangan sistem prediksi, gambaran program yang dibuat.

4. BAB 4 HASIL DAN DISKUSI

Bab ini membahas spesifikasi system yang digunakan dalam penelitian, implementasi sistem yang dirancang dan uji coba sistem.

5. BAB 5 SIMPULAN DAN SARAN

Bab ini berisikan simpulan yang dihasilkan dari implementasi dan saran untuk pengembangan sistem yang akan dilakukan di masa mendatang.