

PREDIKSI JUMLAH KASUS COVID-19 DI JAWA BARAT MENGGUNAKAN METODE ARIMA

Nadiza Lediwara¹, Aulia Khamas Heikmakhtiar², Sembada Denrineksa Bimorogo³,
Allodya Nadra Xaviera⁴, Alya Hijrianisa⁵, and Amadea Kenyoning Hapsari⁶

^{1,2,3}Informatika, Fakultas Sains dan Teknologi Pertahanan,
Universitas Pertahanan Republik Indonesia

^{4,5,6}Kimia, Fakultas MIPA Militer, Universitas Pertahanan Republik Indonesia

¹nadizalediwara@gmail.com; ²auliakphd@gmail.com; ³denri5693@gmail.com;

⁴allodya.nadra@gmail.com; ⁵ahijrianisa@gmail.com; ⁶amadeasubiakto11@gmail.com

Abstrak— Penelitian ini bertujuan untuk memprediksi seberapa banyak penambahan jumlah kasus Covid-19 di Jawa Barat kurun waktu tertentu menggunakan metode ARIMA. Data yang digunakan berupa data total kasus pasien positif Covid-19 di Jawa Barat dari bulan Mei 2021 sampai Desember 2021. Hasil dari penelitian ini adalah untuk data total kasus pasien positif Covid-19 model terbaiknya ARIMA (2,1,0) dengan metode manual. Diperoleh hasil yaitu kasus Covid-19 cenderung menurun tidak terlalu drastis.

Kata Kunci: Covid-19, ARIMA, Jawa Barat

1. PENDAHULUAN

Dalam penyelenggaraan sistem pertahanan negara, maka masyarakat Indonesia harus mempunyai kemampuan dasar dalam bela negara. Pasal 9 ayat 1 UU No.3 Tahun 2002 berisi tentang hak dan kewajiban setiap warga negara untuk ikut serta dalam upaya bela negara sebagai yang diwujudkan dalam penyelenggaraan pertahanan negara. Bela negara diwujudkan tidak harus dengan ikut berperang di medan pertempuran, tetapi mengabdikan sesuai dengan profesi masing-masing demi memperkuat sistem pertahanan negara adalah suatu implementasi dari bela negara. Maka dari itu, bela negara tidak dilaksanakan oleh militer saja, tetapi seluruh rakyat ikut serta dalam pelaksanaan bela negara. Perwujudan bela negara cukup luas, diwujudkan dengan bekerja sesuai profesi untuk menangkalkan ancaman-ancaman yang nyata. Salah satu wujud bela negara yaitu bekerja dalam merespon fenomena bencana yang ada di Indonesia. Salah satunya bencana terhadap Covid 19. Virus Covid-19 telah menyebar dengan cepat

sejak pertama kali ditemukan, dan dalam situasi saat ini, kasus virus telah terdeteksi di hampir setiap negara. Organisasi Kesehatan Dunia telah mengidentifikasi bencana virus yang akhirnya dikenal sebagai Covid-19 atau virus corona, dan menyebutnya sebagai pandemi. Pada tanggal 28 Desember 2020, dunia telah mencatat sekitar 1,58 juta nyawa hilang akibat virus corona. Sedangkan infeksi virus corona pertama kali terdeteksi di Indonesia pada Maret 2020. Sejak infeksi ini pertama kali terdeteksi, 21.452 orang telah meninggal di Indonesia per 28 Desember 2020 [1]. Infeksi COVID-19 umumnya menyebabkan penyakit pernapasan ringan hingga berat dan kematian, namun beberapa orang yang terinfeksi virus ini mungkin tidak menunjukkan gejala. Proporsi kasus tanpa gejala bervariasi antara 6% dan 41%. Proporsi pasien tanpa gejala berubah pada kelompok usia yang lebih tua karena meningkatnya penyakit penyerta. Gejala umum COVID-19 antara lain batuk, nafas cepat/sesak, demam/menggigil, nyeri otot/badan, mual, diare, dan hilangnya rasa/bau.

Selain gejala yang biasa terjadi, seperti kesulitan bernapas, nyeri dada terus-menerus, kelelahan, tidak mampu berdiri, dan bibir atau wajah membiru, ada juga tanda darurat yang memerlukan perhatian medis segera. Gejala berkembang secara bertahap (dari ringan hingga parah) dan dapat terjadi 2 hingga 14 hari setelah terpapar virus [2]. Jumlah orang yang terinfeksi virus corona baru juga semakin meningkat di Indonesia. Per 1 Mei 2021, jumlah kasus baru virus corona di Indonesia sebanyak 743.198 kasus. Penyebaran penyakit ini dapat dikendalikan dengan menggunakan teknik pemodelan statistik untuk membuat prediksi berdasarkan perkiraan jumlah orang yang terinfeksi COVID-19. Model-model ini memungkinkan perkiraan prevalensi penyakit dalam jangka pendek dan jangka panjang serta memberikan gambaran umum mengenai jumlah layanan kesehatan yang diperlukan untuk mengurangi dampak dari berbagai sektor secara signifikan [3]. Prediksi penyebaran virus ini dapat dikendalikan dengan menggunakan teknik pemodelan statistik. Model-model ini dapat digunakan untuk memperkirakan prevalensi suatu penyakit dalam jangka pendek dan jangka Panjang. Data rangkaian waktu yang menjadi dasar pengambilan keputusan. Dari pemodelan *data science* yang ada dapat diprediksi kasus Covid 19 ini dan dapat dilihat apakah terjadi penurunan ataukah peningkatan. ARIMA (*Autoregressive Integrated Moving Average*) adalah salah satu model deret waktu yang paling populer. Model ARIMA terdiri dari tiga fase utama: identifikasi, evaluasi dan pengujian, dan pengendalian diagnostik. Jika model yang dihasilkan memuaskan maka model ARIMA dapat digunakan untuk melakukan prediksi.

2. LANDASAN TEORI

Penelitian pertama yaitu membahas prediksi jumlah pendaftaran mahasiswa baru dengan membandingkan model ARIMA dan Prophet [4]. Hasil dari penelitian ini adalah model Prophet bekerja lebih baik dibandingkan dengan model ARIMA. Kinerja yang lebih

baik terlihat dari hasil RMSE yang semakin mendekati angka 0, meskipun dari algoritmanya model Prophet bekerja lebih sederhana dibandingkan dengan ARIMA. Penelitian kedua yaitu membahas mengenai prediksi harga produksi karet [5]. Penelitian ini menggunakan model ARIMA. Penelitian ini dilakukan karena dalam tiga tahun terakhir harga produksi karet mengalami fluktuasi. Hasil dari prediksi tersebut didapatkan model ARIMA (1,1,2) merupakan model terbaik untuk prediksi harga produksi karet. Penelitian ketiga yaitu membahas prediksi harga emas dengan model ARIMA [6]. Hasil dari sepuluh percobaan maka didapatkan prediksi harga emas terbaik. Hasil yang didapatkan itu juga dapat membantu pedagang dalam pengambilan keputusan.

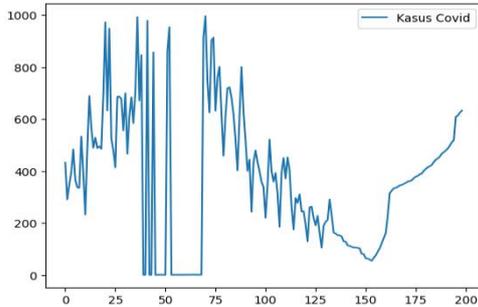
3. METODE PENELITIAN

Tahapan pengerjaan penelitian ini menggunakan tahapan dalam CRISP-DM [7]. Tahapan awal yaitu *Business Understanding*. Tahapan ini merupakan tahapan untuk mengetahui ruang lingkup yang akan dikerjakan, apa saja yang dibutuhkan yang akan dikerjakan [8]. Tahapan berikutnya yaitu *Data Understanding*. Tahapan ini peneliti perlu mengetahui dapat apa saja yang akan diproses pada tahapan penelitian [9]. Adapun data yang akan digunakan menggunakan data time series yang berisi mengenai jumlah kasus Covid 19. Tahapan berikutnya yaitu *data preparation*. Penyediaan data ini adalah langkah yang paling penting dalam alur penelitian [10]. Penyediaan data seperti pengecekan *missing value* dan hal lainnya sangat diperlukan. Tahapan berikutnya yaitu *modelling*. Model yang digunakan pada penelitian ini adalah model ARIMA. Setelah pemodelan maka tahapan berikutnya adalah Evaluasi. Pada tahapan evaluasi ini mencari nilai RMSE terbaik. Nilai RMSE terbaik adalah yang semakin mendekati nilai 0.

4. HASIL PENELITIAN DAN PEMBAHASAN

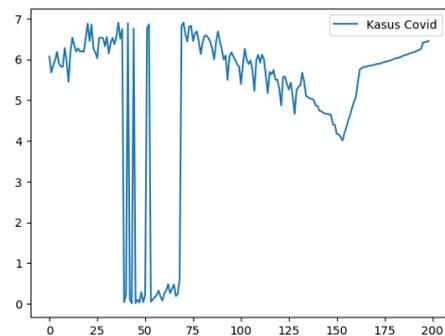
Tahap pertama sebelum membentuk model

ARIMA yang dilakukan adalah membuat plot data. Plot data total kasus pasien positif COVID-19 di Jawa Barat ditunjukkan pada Gambar 1. Data diambil dari 1 Mei 2021 hingga 31 Desember 2021. Berdasarkan Gambar 1, diketahui bahwa data total kasus pasien Covid-19 merupakan bukan data musiman sehingga dapat diprediksi dengan metode ARIMA.



Gambar 1. Plot Data Kasus Covid-19

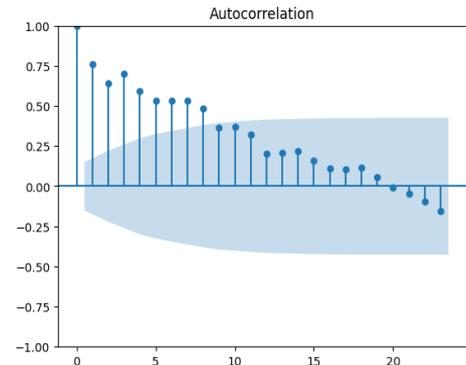
Langkah selanjutnya adalah mengecek kestasioneran data, hal ini bertujuan untuk memperkecil kesalahan model. Data dapat dikategorikan stasioner apabila memenuhi kestasioneran terhadap ragam dan rerata. Untuk membuat kestasioneran data dilakukan dengan pemilihan metode logaritma natural dengan basis e sehingga berdasarkan Gambar 2, diperoleh bahwa data dengan fluktuasi yang relatif berkurang dan terdapat perubahan persentase.



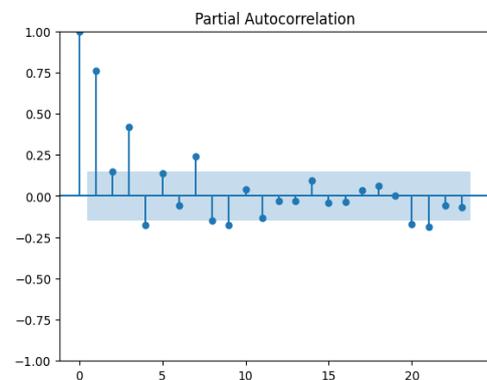
Gambar 2. Stasioner Data dengan Logaritma

Disimpulkan bahwa, data yang ditransformasikan lebih stasioner terhadap ragam. Selanjutnya adalah mengecek kestasioneran data terhadap nilai rerata. Pada Gambar 3 dan Gambar 4, terlihat bahwa setelah dilakukan autokorelasi dan parsial autokorelasi terdapat beberapa lag yang

masih relatif tinggi dan cenderung berada pada garis konfidensi interval. Hal ini berarti data tidak stasioner terhadap rerata.



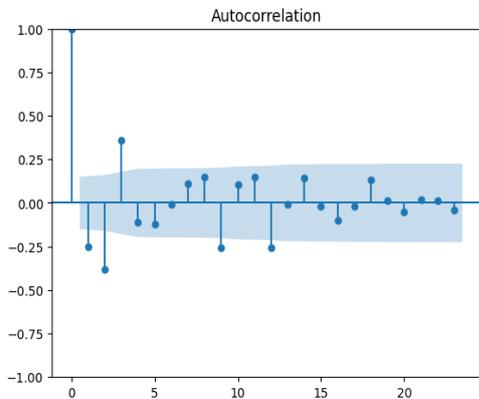
Gambar 3. Autokorelasi Data Transformasi Covid-19



Gambar 4. Parsial Autokorelasi pada Data Hasil Transformasi Kasus Covid-19

Selanjutnya, untuk lebih memvalidasi stasioneritas suatu data dilakukan uji ADF. Uji *Augmented Dickey Fuller* (ADF) adalah uji yang dilakukan untuk memperkuat adanya deskripsi pengujian kestasioneran data [10]. Untuk memutuskan data stasioner atau tidak, dilakukan perbandingan hasil output dari interpretasi p-value. Apabila p-value kurang dari tingkat signifikansi (0,05), maka dapat diartikan bahwa data cenderung stasioner. Pada pengujian ADF, diperoleh hasil p-value sebesar 0.13987. Dari hasil pengujian diperoleh hasil bahwa data tersebut cenderung tidak stasioner terhadap rerata. Agar data stasioner, dilakukan differencing terhadap data hasil transformasi agar lag tidak terlalu tinggi dan berada dalam batas garis konfidensi interval, setelah itu dilakukan autokorelasi kembali. Gambar 5 menandakan bahwa p-value diperoleh hasil lebih rendah dan sudah berada pada jangkauan garis konfidensi interval. Agar lebih tervalidasi stasioneritasnya, dilakukan pengujian ADF

kembali. Uji ADF menghasilkan nilai p-value sebesar 0.0001495. Dengan nilai p-value yang sangat kecil mendekati nol, deret waktu yang diperoleh kemungkinan besar sudah stasioner. Oleh karena itu sudah dapat ditentukan model-model terbaik dari ARIMA yang dapat digunakan.



Gambar 5. Autokorelasi Hasil Differencing Kasus Positif Covid-19

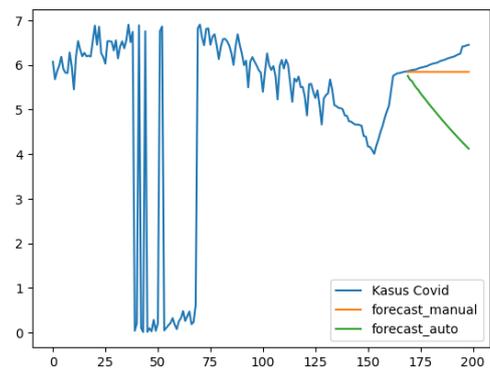
Selanjutnya, dilakukan pembuatan model arima dengan metode manual. Kelebihan dari metode manual adalah memberikan kontrol penuh kepada pengguna untuk menentukan orde dan parameter model. Dari data akan diperoleh beberapa informasi mengenai nilai parameter, uji statistik, dan lain-lain. Dari codingan python pun di bandingkan hasil dari ARIMA dengan manual dan auto untuk mendapatkan hasil dengan nilai RMSE terkecil.

| Manual | | | Auto | | |
|--------|------|------|------|------|------|
| MAE | MAPE | RMSE | MAE | MAPE | RMSE |

Tabel 1. Perbandingan Hasil Manual dan Auto ARIMA

Selain dengan pemodelan ARIMA metode manual, dilakukan kembali pemodelan arima dengan metode auto atau sering disebut auto ARIMA. Kelebihan dari metode auto ARIMA adalah mudah digunakan, tidak memerlukan pengetahuan mendalam tentang ARIMA, serta Cepat dan efisien. Perbandingan penggunaan dari kedua metode tersebut adalah orde ARIMA yang digunakan. Pada percobaan ini, metode manual menggunakan orde ARIMA (2,1,0) sedangkan metode auto ARIMA memperoleh hasil orde arima (3,0,0). Pada Tabel 1 diperoleh hasil perbandingan forecast dari metode arima manual dan auto

arima. Dilihat dari data bahwa MAE; MAPE; RMSE pada metode ARIMA manual lebih kecil daripada auto ARIMA. Semakin rendah nilai dari MAE; MAPE; RMSE, maka akan semakin baik model dalam melakukan prediksi. Untuk memprediksi kasus Covid-19 maka digunakan visualiasi seperti yang terlihat pada Gambar 6. Dari Gambar 6 diperoleh hasil bahwa dengan metode auto ARIMA kasus positif Covid-19 cenderung menurun secara drastis, sedangkan pada metode manual kasus positif Covid-19 cenderung sedikit menurun.



Gambar 6. Prediksi ARIMA Manual dan Auto

5. KESIMPULAN DAN REKOMENDASI

- Dari hasil prediksi banyaknya penam bahan jumlah kasus Covid-19 di Jawa Barat pada kurun waktu tertentu menggunakan ARIMA.
- Hasil dari penelitian ini adalah pada data dalam kurun waktu 1 Mei 2021 hingga 31 Desember 2021 diperoleh hasil kasus COVID-19 yang cenderung turun namun tidak terlalu drastis.
- Prediksi tersebut benar adanya, karna pada kenyataan periode COVID-19 kurun waktu yang di analisis, mengalami penurunan jumlah kasus infeksi virus ini.

6. REFERENSI

- [1] A. Mufriyani, "Strategi bisnis multi level marketing pada masa pandemi covid-19

- perspektif marketing syariah: studi pada PT. Melia Sehat Sejahtera Kota Mataram, "Jun.2022,doi.10.1101/2020.05.20.20107763V1.
- [2] B. G. Masyah Jurusan Keperawatan Poltekkes Kemenkes Palangka Raya Kampus Jln Obos, "Pandemi Covid 19 Terhadap Kesehatan Mental Dan Psikososial," *Mnj (Mahakam Nursing Journal)*, Vol.2, No.8, Pp.353–362, Dec. 2020, Doi: 10.35963/Mnj.V2i7.180.
- [3] S. Mervin, T. Sekolah, T. Bisnis, M. Dua, and S. Bitung, "Dampak Pandemi Covid 19 Terhadap Bisnis Resort di Kelurahan Mawali," *Jurnal Syntax Fusion*, vol. 1, no. 04, pp. 52–64, Apr. 2021, doi: 10.54543/FUSION.V1i04.113.
- [4] C. Chandra and S. Budi, "Analisis Komparatif ARIMA dan Prophet dengan Studi Kasus Dataset Pendaftaran Mahasiswa Baru," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 2, pp. 2443–2229, Aug. 2020, doi: 10.28932/JUTISI.V6i2.2676.
- [5] "Penerapan Metode ARIMA Dalam Peramalan Harga Produksi Karet di Provinsi Riau | Jurnal Teknologi dan Manajemen Industri Terapan." Acces sed: Jan. 16, 2024. [Online].
- [6] Y. M. Pranoto, R. A. Harianto, And I. Iswanto, "Pemanfatan Arima Untuk Prediksi Harga Emas Dalam Sistem Rekomendasi Trading Gold Option," *Jurnal Media Informatika Budidarma*, Vol. 4, No. 4, Pp. 863–871, Oct. 2020, Doi: 10.30865/Mib.V4i4.2246.
- [7] F. Nuraeni, N. N. Febriani Sm, L. Listiani, and E. Rahmawati, "Implementation of K-Means Algorithm with Distance of Euclidean Proximity in Clustering Cases of Violence Against Women and Children," *2019 1st International Conference on Cybernetics and Intelligent System, ICORIS 2019*, pp. 162–167, Aug. 2019, doi: 10.1109/ICORIS.2019.8874883.
- [8] C. Schröer, F. Kruse, and J. M. Gómez, "A Systematic Literature Review on Applying CRISP-DM Process Model," *Procedia Comput Sci*, vol. 181, pp. 526–534, Jan. 2021, doi: 10.1016/J.PROCS.2021.01.199.
- [9] A. Khumaidi, "Data Mining For Predicting The Amount Of Coffee Production Using Crisp-Dm Method," *Jurnal Techno Nusa Mandiri*, vol. 17, no. 1, pp. 1–8, Feb. 2020, doi: 10.33480/TECHNO.V17i1.1240.
- [10] J. A. Solano, D. J. Lancheros Cuesta, S. F. Umaña Ibáñez, and J. R. Coronado-Hernández, "Predictivemodels assessment based on CRISP-DM methodology for students performance in Colombia Saber 11 Test," *Procedia Comput Sci*, vol. 198, pp. 512–517, Jan. 2022, doi: 10.1016/J.PROCS.2021.12.278.