

Algoritma SARIMA sebagai Pendukung Strategi Peramalan HPS dalam Persaingan Tender di LPSE Indonesia

Daud Fernando^{1,*}, Indira Syawanodya¹, Raditya Muhammad¹

¹ Program Studi Rekayasa Perangkat Lunak, Universitas Pendidikan Indonesia, Indonesia

* Correspondence: daudfernando@upi.edu

Copyright: © 2024 by the authors

Received: 31 Oktober 2024 | Revised: 8 November 2024 | Accepted: 4 Desember 2024 | Published: 19 Desember 2024

Abstrak

Tender di Layanan Pengadaan Secara Elektronik (LPSE) Indonesia merupakan pengadaan barang/jasa yang berupa fasilitas publik dan dikelola oleh penyedia dengan nilai Harga Perkiraan Sendiri (HPS) terendah ketika proses *reverse auction*. Nilai HPS yang fluktuatif serta ketatnya persaingan kompetitor membuat kemenangan bagi penyedia semakin sulit dan kompetitif. Tujuan dari penelitian ini adalah membuat model peramalan dari nilai HPS tender di LPSE Indonesia menggunakan algoritma *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Jenis penelitian ini adalah penelitian eksperimental untuk menentukan ordo dari model SARIMA terbaik. Variabel penelitian yang digunakan adalah tanggal publikasi tender dan nilai HPS sebanyak 747.098 data tender dari data historis hasil *web scraping* situs web LPSE dengan rentang tanggal penarikan 7 Januari 2013 hingga 30 November 2022. Teknik analisis data menggunakan analisis eksplorasi data untuk mengetahui karakteristik sebaran data kemudian implementasi algoritma peramalan SARIMA. Hasil penelitian ini menunjukkan model SARIMA((5,1,1),(4,1,1,7)) adalah model optimal dengan nilai evaluasi *mean absolute percentage error* (MAPE) persentase nilai galat 33,56% yang kaitannya dengan LPSE dapat memberikan nilai peramalan yang masuk akal. Hasil peramalan 30 hari ke depan menunjukkan sebaran nilai HPS berada di rentang 680 juta – 700 juta rupiah di periode Desember 2022.

Kata kunci: algoritma sarima; peramalan; prediksi nilai hps; tender; web scraping

Abstract

Tender in Indonesia's Electronic Procurement Service (LPSE) is the procurement of goods/services in the form of public facilities and managed by the provider with the lowest estimated price (HPS) value during the reverse auction process. The fluctuating value of HPS and the tight competition of competitors make winning for providers increasingly difficult and competitive. The purpose of this research is to create a forecasting model of the HPS value of tenders in LPSE Indonesia using the Seasonal Autoregressive Integrated Moving Average (SARIMA) algorithm. This type of research is experimental research to determine the order of the best SARIMA model. The research variables used are tender publication date and HPS value as much as 747,098 tender data from historical data from web scraping of the LPSE website with a withdrawal date range of January 7, 2013 to November 30, 2022. The data analysis technique uses data exploration analysis to determine the characteristics of the data distribution and then the implementation of the SARIMA forecasting algorithm. The results of this study show that the SARIMA((5,1,1),(4,1,1,7)) model is the optimal model with an evaluation value of mean absolute percentage error (MAPE) percentage error value of 33.56% which in relation to LPSE can provide a reasonable forecasting value. The results of forecasting for the next 30 days show that the distribution of HPS values is in the range of 680 million - 700 million rupiah in the period December 2022.

Keywords: sarima algorithm; forecasting; hps value prediction; tender; web scraping

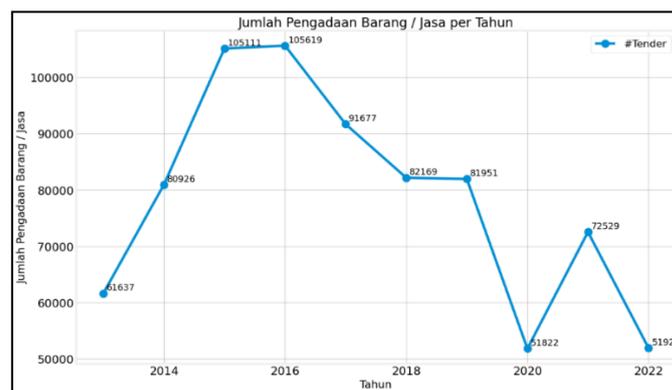


PENDAHULUAN

Layanan Pengadaan Secara Elektronik (LPSE) adalah sistem pengembangan barang jasa atau tender yang sudah terdigitalisasi di berbagai kementerian / lembaga / pemerintah daerah (KLPD). Tujuan utama dari sistem ini memberikan aksesibilitas tinggi sebagai *electronic procurement* serta upaya reformasi birokrasi pengadaan barang / jasa (tender) kepada masyarakat (Satria et al., 2021). Kepala Lembaga Kebijakan Pengadaan Barang/Jasa Pemerintah (LKPP), menafsirkan 52,1% dari total APBN (atau sekitar Rp 1.214,1 triliun) untuk tahun 2021 (Susanto, 2021).

Tender adalah peran utama dalam negara Indonesia karena berkontribusi langsung pada pengadaan barang serta penyediaan fasilitas umum, sebagai contohnya dalam pengerjaan konstruksi (Almahdali et al., 2022). Peserta lelang tender yang berpartisipasi perlu mengajukan nilai harga perkiraan sendiri (HPS) yang dijadikan penentuan apakah layak memenangkan tender milik KLPD (Aksan & Nurfadilah, 2020). Pada tahun 2022, lelang tender tersedia di 617 situs web KLPD yang perlu diakuisisi sebagai data dalam proses analisis peramalan.

Akuisisi data pada situs web memerlukan teknik khusus seperti *web scraping*. *Web scraping* merupakan proses pengambilan data dari situs web. Teknik ini akan mengambil data yang ada di internet (tepatnya pada situs web) untuk disimpan ke dalam basis data agar bisa dijadikan sumber atau pendukung analisis data (Djufri, 2020). Metode *web scraping* juga akan menjamin kualitas data yang dikumpulkan terpercaya dan sesuai sumbernya (Ristijana et al., 2021). Hasil dari teknik pengumpulan data *web scraping* adalah terkumpulnya 785.361 tender nasional di Indonesia dari rentang tahun 2013 hingga tahun 2022 yang pergerakan kuantitas per tahunnya dapat dilihat pada gambar 1. Hasil pada gambar 1 menunjukkan pola fluktuatif jumlah tender dari tahun ke tahun dapat memengaruhi kemampuan penyedia dalam mengajukan penawaran kompetitif, sehingga sangat diperlukan analisis tren data ini lebih lanjut dalam merumuskan strategi pengadaan.



Gambar 1. Pergerakan jumlah tender per tahun

Permasalahan penelitian ini adalah beberapa peserta lelang tender terkadang kalah dalam mengajukan nilai HPS di masa pelelangan karena meleset dalam memproyeksikan pola kenaikan atau penurunan harganya (Abimantara & Purwito, 2019). Selain itu, peserta lelang tender membutuhkan alat yang dapat membantu peramalan nilai HPS kedepannya karena merasa peningkatan kuantitas tender yang ada dari tahun ke tahun dalam meminimalisir pengajuan harga yang terlalu tinggi atau pun rendah (Subagia et al., 2020). Permasalahan dalam proses implementasi peramalan terhadap data berderet waktu yang sudah ada hanya berbasis algoritma *Autoregressive Moving-Average* (ARIMA) (Sirisha et al., 2022). Model ini menggabungkan tiga komponen yaitu Auto Regressive (p), Integrated (d), dan Moving Average (q). Sehingga masing – masing dari komponen tersebut memiliki variabelnya sendiri dan kerap kali dinotasikan menjadi ARIMA (p, d, q) (Arta et al., 2022). Penelitian berbasis

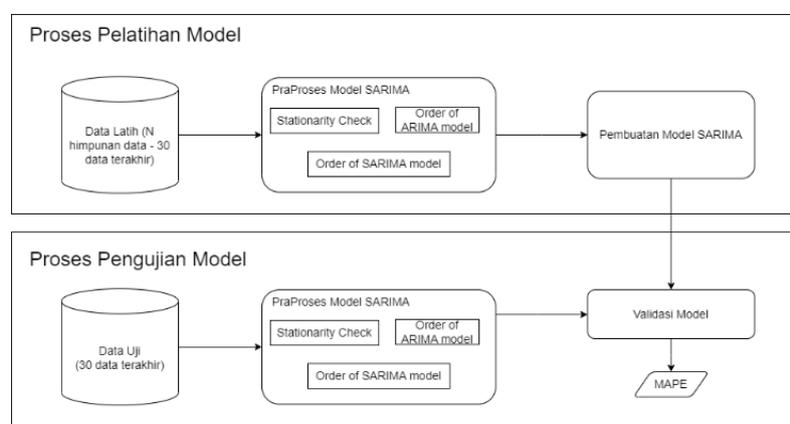
model ARIMA menghasilkan secara objektif model yang dapat dipercaya dalam melakukan peramalan data berderet waktu dengan nilai evaluasi Mean Absolute Error (MAE) sebesar 13,23 (Noureen et al., 2019). Penelitian sebelumnya menyarankan untuk menggunakan model yang mampu menangkap aspek musiman dari data *time series* karena hasilnya lebih sesuai dengan kasus dunia nyata yang dapat meningkatkan nilai akurasi prediksi (Qonita et al., 2017).

Solusi atas semua permasalahan tersebut adalah model peramalan menggunakan algoritma SARIMA dengan himpunan data tender dari LPSE Indonesia. Penelitian sebelumnya menggunakan model SARIMA karena menjadi model yang lebih sensitif dalam mempelajari data deret waktu, khususnya yang memiliki pola berulang karena musiman namun memiliki keterbatasan untuk peramalan pada periode waktu bulanan saja (Liu et al., 2022). Penelitian ini akan menutupi dari temuan sebelumnya melalui peramalan harian pada nilai HPS tender di LPSE Indonesia. Tujuan dari penelitian ini adalah mengimplementasi algoritma SARIMA dari data hasil akuisisi situs web LPSE pada periode tahun 2013 hingga tahun 2022 untuk mendukung penyedia dalam proses lelang tender.

Dampak dari hasil penelitian ini adalah membantu penawaran nilai HPS yang lebih tepat bagi para penyedia dan tidak merugikan. Selain itu, terbatasnya penelitian mengenai peramalan nilai HPS pada media LPSE akan menjadi kebaruan dari ranah penelitian dalam segi pembelajaran mesin. Kebermanfaatan penelitian ini menambah literatur yang kurang memperhatikan peramalan HPS untuk platform tender elektronik di Indonesia atau pengadaan publik pada umumnya.

METODE

Jenis penelitian ini adalah penelitian eksperimental untuk memilih dan menguji model SARIMA terbaik dalam memprediksi data yang fluktuatif. Pemilihan model SARIMA untuk meramalkan nilai HPS di platform LPSE adalah pilihan yang tepat dibandingkan dengan model peramalan yang lain. Hasil penelitian pada data berderet waktu yang fluktuatif yang membandingkan dua model peramalan SARIMA dengan *Holt-Winters* menunjukkan bahwa SARIMA sebagai model dengan evaluasi terbaik dalam melakukan *forecasting*. Sehingga dalam penelitian ini akan mengimplementasikan teknik peramalan berbasis model SARIMA. Gambar 2 adalah alur penelitian yang dilakukan yang membagi dua bagian himpunan data untuk melatih serta menguji model peramalan SARIMA.



Gambar 2. Rancangan model komputasi

Proses analisis dan prediksi data menggunakan metode SARIMA akan memerlukan tujuh proses utama. Pertama, menganalisis data *moving average* menggunakan grafik garis yaitu dengan mengambil nilai rata-rata dalam nilai per periode waktu harian di mana yang akan berada pada sumbu y adalah nilai HPS dan dari sumbu x adalah periode waktunya. Kedua, Pemeriksaan stasioneritas himpunan data menggunakan *Augmented Dickey-Fuller Test* (ADF)

(Fauziah et al., 2021). Tabel 1 adalah variabel yang digunakan terdiri dari kolom data nilai HPS dan tanggal publikasi tendernya.

Tabel 1. Variabel model SARIMA

Tanggal Publikasi Tender	Nilai HPS (dalam Rupiah)
2022-09-08	1939998469
2021-03-04	1798500000
2019-07-12	198440000
2016-06-28	84000000
2013-05-15	597800000

Ketiga, pengidentifikasian model (penentuan ordo SARIMA) dengan parameter atau ordo p , d , q , P , D , Q , dan m berbasis nilai dari *Partial Auto Correlation Function* (PACF) dan juga *AutoCorrelation Function* (ACF). Persamaan (1) adalah cara menentukan nilai dari ordo ACF model SARIMA. Persamaan (1) juga menjadi penentu apakah proses *differencing* perlu dilakukan atau tidak bila mana nilai rentangnya tidak berdistribusi normal.

$$\rho_k = \frac{\text{covariance}(Z_t, Z_{t-k})}{\sqrt{\text{Variance}(Z_t)}\sqrt{\text{Variance}(Z_{t-k})}} \quad (1)$$

Keterangan:

- ρ_k = nilai ACF
- Z_t = nilai data deret waktu di lag ke-t
- Z_{t-k} = nilai deret waktu lag ke-t dikurangi lag ke-k

Keempat, proses validasi model. Validasi model berbasis *time series cross validation* dengan membagi himpunan data menjadi dua bagian. Himpunan data nilai HPS Kelima, Penentuan Ambang Batas MAPE Model SARIMA. MAPE dipercaya oleh peneliti untuk menguji hasil evaluasi dari model peramalan pada data berderet waktu melalui persamaan (3) yang dapat menentukan apakah model termasuk kategori yang masuk akal dalam meramal.

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|y_t - \hat{y}_t|}{y_t} \times 100\% \quad (3)$$

Di mana nilai y_t adalah nilai aktual himpunan data dari poin ke-t variabel independen, \hat{y}_t adalah hasil prediksi dari poin ke-t variabel independen, dan N adalah total himpunan data observasi. Setidaknya dalam kategorisasi Tabel 2, suatu model dapat digunakan untuk peramalan apabila berada di level prediksi yang masuk akal di mana nilai rentang MAPE di antara 20 hingga 50.

Tabel 2. Interpretasi metrik mape

MAPE	Interpretasi
< 10	Prediksi dengan akurasi sangat tinggi
10 – 20	Prediksi yang baik
20 – 50	Prediksi yang masuk akal
> 50	Prediksi yang tidak akurat

Selanjutnya tahapan terakhir adalah melakukan kelayakan model dengan cara pemeriksaan kelayakan model. Apabila residual memenuhi perilaku *white noise*, maka model tersebut dapat ditetapkan sebagai model yang layak. Implementasi pengujian menggunakan *Q-Ljung Box* yang berdasar pada nilai ujinya pada penerimaan atau penolakan hipotesis nul yang residualnya bersifat *white noise*.

HASIL DAN PEMBAHASAN

Hasil

Hasil dari pengakuisisian data melalui teknik *web scraping* berhasil diimplementasikan pada 617 KLPD di Indonesia per tahun 2022. Setiap tender nasional dilakukan peledangan di masing – masing situs web KLPD yang mayoritas hampir memiliki pola elemen data yang sama, sehingga sangat memungkinkan untuk dilakukan pengambilan data sekaligus. 3 dari 617 KLPD yang ditarik datanya dapat dilihat pada Tabel 2 menunjukkan data yang terdiri dari tiga kolom, di antaranya nama LPSE atau KLPD asalnya, situs web dalam bentuk URL, serta jumlah tender yang ada di setiap situs webnya. Hasil datanya dapat digunakan untuk melakukan perulangan dalam akuisisi data tender di berbagai KLPD LPSE Indonesia.

Tabel 2. Hasil agregasi tiga situs web lpse

Nama LPSE	URL	Jumlah Tender
LPSE Kabupaten Flores Timur	http://103.131.61.17/eproc4	1120
LPSE Kota Padang Sidempuan	http://103.131.61.187/eproc4	1214
LPSE Kabupaten Rokan Hilir	http://103.131.61.54/eproc4	2312

Permasalahan perubahan struktur situs web, kesalahan dalam *Extensible Path Markup Language* (XPath), dan juga masalah koneksi berhasil dihadapi dengan menganalisis lebih detail lokasi atribut atau variabel yang digunakan menggunakan XPath-nya. Selain itu, penggunaan *library* Pandas dalam melakukan manipulasi data seperti melakukan *slicing* pada teks, memisahkan nominal rupiah dengan logo mata uangnya, serta mengubah karakter o dengan angka 0 pada kolom nilai HPS telah dilakukan. Proses ekstraksi data pun berlangsung dengan bantuan *library selenium* khususnya *from selenium.webdriver.common.by import By* yang akan mengakses metode XPath untuk mendapatkan atribut dari file HTML yang sudah didapatkan oleh *library BeautifulSoup*. Dalam melakukan proses peramalan data berderet waktu diperlukan setidaknya dua atribut yang digunakan, di antaranya adalah data tanggal dan juga data numerikal. Tujuannya untuk memenuhi syarat dasar pada himpunan data berderet waktu.

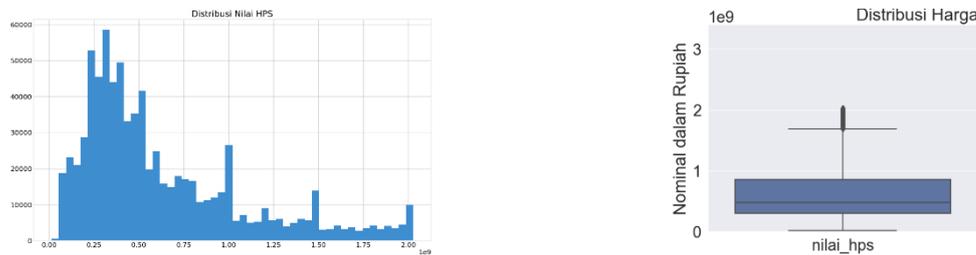
Pengakuisisian data pada situs web LPSE hanya akan mengambil dua atribut berupa tanggal pembuatan tendernya dan kemudian berapa nilai HPS yang telah diatur dalam satuan mata uang rupiah. Tabel 3 menunjukkan bagaimana setiap atribut bisa didapatkan menggunakan URL dan juga XPathnya. Hasil pengaplikasian XPath karena berbasis pada situs web adalah elemen yang dianalisis dapat dilakukan perulangan pada nilai data yang sama di situs web KLPD lainnya karena polanya berulang. Hasilnya dapat diterapkan dengan mengambil id “main”, kolom ke-5, dan baris ke-1 untuk variabel tanggal pembuatan tender.

Tabel 3. Analisis Pola XPath Variabel di situs web LPSE

Nama Atribut	URL	XPath
Tanggal Pembuatan	http://{domain_lpse}/eproc4/lelang/{kode_tender}/pengumumanlelang	//*[@id="main"]/div/able/tbody/tr[5]/td
Nilai HPS (dalam Rupiah)	http://{domain_lpse}/eproc4/lelang/{kode_tender}/pengumumanlelang	//*[@id="main"]/div/able/tbody/tr[13]/td[2]

Hasil akuisisi data yang berhasil dilakukan adalah sebanyak 785.361 tender nasional di Indonesia dari rentang tahun 2013 hingga tahun 2022, contoh lima data tendernya dapat dilihat pada Tabel 4. Tabel 4 menunjukkan atribut atau kolom tanggal pembuatan dari tender yang berada di situs web LPSE sedangkan nilai HPS adalah penjelasan dari nilai harga perkiraan sendiri yang dipublikasi oleh KLPD bersangkutan dalam satuan mata uang rupiah.

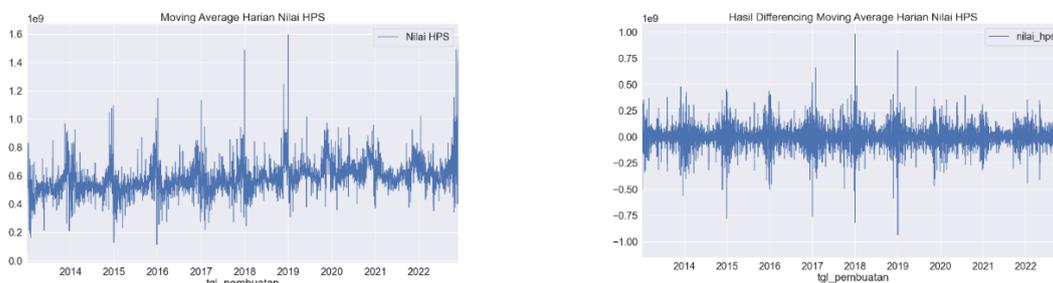
Nilai HPS merupakan variabel bertipe numerikal sehingga penting untuk dilihat dari segi persebaran datanya melalui grafik histogram. Gambar 3 (sebelah kiri) memperlihatkan sebaran nilai HPS nasional yang ada mayoritas pada rentang nilai HPS 200 juta hingga 500 juta yang kuantitasnya mencapai lebih dari 30 ribu tender selama 9 tahun. Grafik ini juga memberikan posibilitas terdapatnya nilai *outlier* suatu tender.



Gambar 3. Grafik sebaran data pada variabel nilai hps

Gambar 3 (sebelah kanan) memvalidasi terdapatnya distribusi data *outlier* menggunakan diagram *box plot*. Gambar 4 melihat sebaran normal nilai HPS berada di rentang nilai 100 juta hingga sekitar 500 juta dan terdapat lonjakan frekuensi tinggi di nilai HPS 2 miliar rupiah. Hasil data yang diakuisisi ini akan tetap digunakan karena bila dilakukan modifikasi data akan menghilangkan aspek natural dari himpunan berderet waktu.

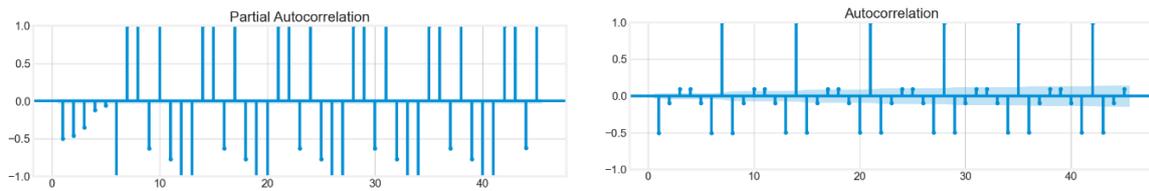
Implementasi model SARIMA memerlukan himpunan data berderet waktu yang memenuhi ketentuan. Langkah pertama adalah memenuhi syarat stasioneritas data menggunakan pengujian teknik ADF dan juga *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS). Didapatkan bahwa data deret waktunya tidak stasioner pada uji KPSS, dengan nilai *p-value* sama dengan 0,01 yang artinya lebih kecil dari nilai 0,05 (5%), sehingga diperlukan transformasi data berupa proses *differencing*. Adapun nilai rata-ran bergerak dari granularitas hariannya yang belum dan yang sudah stasioner dapat dilihat pada gambar 5.



Gambar 5. Perubahan hasil proses *differencing*

Gambar 5 menghasilkan data berderet waktu yang lebih tersentral nilai rata-ratanya dan memiliki kecenderungan pola statis dari waktu ke waktu. Selanjutnya dilakukan pengidentifikasian order SARIMA terhadap data deret waktunya (data tren) lengkap dengan pengidentifikasian order pada data musimannya menggunakan autokorelasi *correlogram* pada data historis. Grafik ACF (sebelah kanan) dan grafik PACF (sebelah kiri) dapat dilihat pada gambar 6. Plot PACF menunjukkan hingga nilai lag ke-4, autokorelasinya cukup signifikan sehingga order untuk *autoregressive* (AR) pada data musimannya adalah empat. Plot ACF hanya lag ke-1 saja yang menunjukkan autokorelasi yang signifikan karena pada lag ke-2

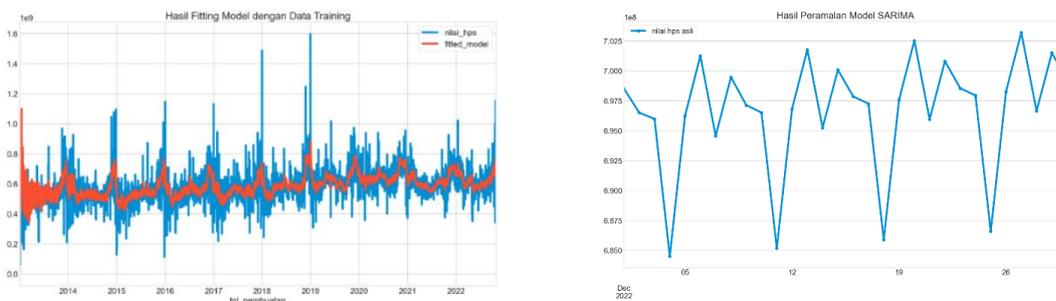
nilainya sudah turun drastis dan besar kemungkinan nilai prediksi akan menjadi sulit ditentukan karena modelnya terlalu kompleks.



Gambar 6. Hasil pacf dan acf untuk menentukan ordo model sarima

Gambar 6 membuktikan bahwa ordo musiman untuk model SARIMA adalah tujuh. Dapat dilihat setiap order kelipatan tujuh nilai dari autokorelasinya sama dengan satu. Untuk order dari model *integrated*, nilainya didapatkan dari berapa banyak proses *differencing* dilakukan artinya nilainya adalah satu. Sehingga model musimannya menghasilkan ARIMA(5,1,1) dan membentuk SARIMA((5,1,1),(4,1,1,7)).

Pengimplementasian model SARIMA dapat menggunakan atribut *fitted_value* yang hasil performansi modelnya pada gambar 7 (sebelah kiri). Garis berwarna merah merupakan seberapa tinggi tingkat model dapat memahami data latih yang tersedia pada garis berwarna biru yang merupakan data latih. Grafik ini melihat seberapa baik performansi model yang sudah dibuat terhadap data latih dan menjadi basis proses prediksi selanjutnya. Dapat dilihat, model berhasil menangkap nilai rata-rata keseluruhan data latihnya dengan baik tanpa ada nilai ekstrem yang dihasilkan. Gambar 7 (sebelah kanan) adalah hasil peramalan nilai HPS untuk 30 hari kedepan yang mengutamakan aspek musiman serta berulang dalam periode waktu kelipatan hari ke-6.



Gambar 7. Fitted value dari model komputasi sarima

Aspek penting peramalan yang digunakan untuk memenangkan lelang tender adalah nilai HPS harian dalam periode mingguan. Nilai HPS dalam lima hari sebelumnya berpengaruh untuk mendapatkan nilai galat dari hasil peramalan. Nilai HPS pada kelipatan tujuh hari di pekan pertama, kedua, ketiga, dan keempat sebelumnya berpengaruh pada nilai fluktuatif nilai HPS di LPSE tender. Pengujian kelayakan model dari nilai galat yang dibuat dapat menggunakan pengujian *Ljung-Box* pada lag ke-20 mendapati nilai *p-value* sama dengan 0.018120. Sedangkan pengujian *Kolmogorov-Smirnov p-value* untuk menerima hipotesis null dan hasilnya mengambil hipotesis alternatif yaitu distribusi residualnya tidak normal yang didapatkan bernilai 0.00099.

Pembahasan

Proses validasi dilakukan dengan membandingkan nilai prediksi dari modelnya sesuai kuantitas data uji yaitu 30 data terakhir (Safira et al., 2022). Hasil model peramalan SARIMA untuk metrik evaluasi MAPE adalah 33,56%. Model ini dikategorikan sebagai model yang masuk akal karena nilai evaluasi MAPE berada di rentang 20% - 50% secara keseluruhan

(Haryadi & Meilidhani, 2023). Nilai evaluasi ini dikarenakan juga menggunakan teknik peramalan selama 30 hari ke depan yang dapat menimbulkan 30 kemungkinan dalam memprediksi nilai HPS, karena himpunan data yang digunakan terlalu fluktuatif. Model yang dibuat melihat pada aspek data nilai HPS harian (selama lima hari) dan nilai HPS kelipatan tujuh hari (pekan satu, dua, tiga, dan empat).

Model SARIMA ini kemudian dapat digunakan oleh para penyedia tender dalam menentukan nilai HPS di pelelangan berikutnya sehingga memiliki nilai yang kompetitif (Innayah, 2024). Secara keseluruhan model SARIMA yang dihasilkan memiliki performansi gabungan dengan model ARIMA yang membuat model peramalan dapat menangkan aspek musiman dalam data berderet waktu untuk nilai HPS di situs LPSE. Hal ini dikarenakan himpunan data yang digunakan berasal langsung dari aktivitas manusia atau *natural resources* sehingga lebih mampu ditangkap polanya menggunakan model SARIMA yang sejalan dengan penelitian sebelumnya pada data runtun waktu oleh (Hasibuan et al., 2023).

Faktor dalam melakukan agregasi data menggunakan interpolasi linear juga mempengaruhi dalam evaluasi model peramalan yang digunakan pada tingkat granularitas harian. Sejalan dengan penelitian yang dilakukan (Ignasius & Lamabelawa, 2018) data berderet waktu nilai HPS menerapkan metode *down sampling* yang mengoptimalkan hasil dari nilai peramalannya. Hasil teknik ini mengelompokkan nilai HPS yang dipublikasi pada tanggal yang sama, sehingga tidak ada data nilai HPS yang diabaikan. Hasil peramalan dari model SARIMA ini memenuhi hasil perhitungan menggunakan persamaan model yang diciptakan. Perhitungan model peramalan berdasarkan kepada aspek nilai HPS pada periode harian dalam penentuan nilai galat dan juga mempertimbangkan nilai HPS periode satu pekan (tujuh hari) sebagai faktor musiman dalam nilai HPS di LPSE Indonesia.

Apabila model SARIMA diterapkan untuk memprediksi data 30 hari kedepannya dengan rentang waktu 1 Desember 2022 hingga 30 Desember 2022, model SARIMA menghasilkan nilai HPS yang logis bagi para penyedia dengan kisaran rentang nilai median dari seluruh nilai HPS yang ada pada proses validasinya. Hasilnya dapat dijadikan landasan yang sesuai dengan penyediaan barang dan jasa di Indonesia serta membantu para penyedia barang / jasa dalam memutuskan nilai penawaran yang tidak terlalu rendah (mengurangi profit) atau terlalu tinggi (kalah bersaing) yang sejalan dengan penelitian dari (Ferdiansyah & Nadjib, 2022) pada strategi kemenangan tender dan memperkuat algoritma SARIMA yang digunakan.

Kelebihan penelitian ini menggunakan data deret waktu yang sesuai dengan kenyataan melalui teknik akuisisi data *web scraping*. Metode SARIMA membuat aspek musiman dipertimbangkan sebagai faktor peramalan model (Ariyanti & Yusnitasari, 2023). Kekurangan dari penggunaan algoritma SARIMA adalah himpunan data berderet waktu yang memiliki nilai *outlier* akan membuat model kesulitan dalam mempelajari pola himpunan datanya. Selain itu, model ini mungkin memiliki keterbatasan jika terdapat peristiwa tak terduga yang menyebabkan perubahan besar dalam nilai HPS, seperti pandemi atau perubahan regulasi. Penelitian ini menganjurkan untuk penelitian selanjutnya memerlukan adaptasi atau kombinasi dengan metode lain untuk meningkatkan akurasi.

Melalui penelitian ini juga dapat berkontribusi untuk memperlengkapi penelitian yang telah dilakukan oleh Noreen et al. (2019) dan Qonita et al. (2017) dalam melakukan peramalan pada himpunan data berderet waktu. Penelitian ini membuktikan signifikansi pada aspek musiman suatu data sangat berpengaruh dalam menentukan nilai peramalan HPS di LPSE Indonesia. Hasil peramalan nilai HPS menunjukkan pola di waktu tertentu yang dapat menjadi basis pengajuan nilai HPS di pelelangan LPSE.

Sebagai proses validasi model SARIMA yang dibuat memenuhi kerangka kerja peramalan, maka diperlukan pengujian pada nilai galat/residualnya. Pengujian *Ljung-Box* pada lag ke-20 mendapati nilai *p-value* sama dengan 0.018120 yang artinya nilai residual yang dihasilkan modelnya merupakan *white noise*. Interpretasi dari pengujian *Kolmogorov-Smirnov*

p-value yang didapatkan adalah menolak hipotesis null distribusi residualnya tidak normal. Hasil penelitian ini menunjukkan model peramalan termasuk dalam kategori yang masuk akal dan dapat dijadikan pendukung strategi pelelangan tender LPSE di Indonesia.

SIMPULAN

Implementasi algoritma SARIMA pertama berhasil diimplementasi terhadap keseluruhan data tender dalam bentuk model komputasional. Model komputasional menggunakan data hasil prapemrosesan yang memenuhi standar metode SARIMA menghasilkan model SARIMA((5,1,1),(4,1,1,7)) yang performanya masuk dalam kategori model peramalan yang masuk akal dengan metrik evaluasi MAPE senilai 33,56%. Melalui model peramalan ini diharapkan dapat menjadi landasan bagi para penyedia untuk mengajukan nilai HPS pada proses pelelangan tender di situs web LPSE pada masing-masing KLPD. Harapannya dapat menjadi panduan bagi penyedia barang/jasa dalam menentukan harga penawaran yang lebih sesuai dengan perkiraan pasar, seperti layaknya peta jalan yang memandu arah dalam kompetisi tender. Penelitian ini juga memberikan kebaruan dari segi penggunaan himpunan data karena berasal dari pengakuisisian berbasis teknik *web scraping* dengan jumlah himpunan data berderet waktu sebanyak 785.361 di 617 KLPD berbeda dari rentang tahun 2013 hingga tahun 2022. Kebermanfaatan dari segi literatur juga akan memperlengkapi penelitian yang sudah ada dalam peramalan menggunakan model ARIMA dengan mempertimbangkan aspek *seasonality* pada data berderet waktunya yang akan meningkatkan performa model peramalannya. Disarankan untuk penelitian selanjutnya dapat mempertimbangkan model peramalan berbasis *deep learning* lebih spesifik lagi menggunakan model LSTM yang kemudian dikombinasikan dengan model SARIMA sebagai upaya dalam meningkatkan indikator kinerja peramalannya.

REFERENSI

- Abimantara, C. G., & Purwito, A. (2019). Analisa Kegagalan Kontraktor Dalam Proses Tender Sistem E-Procurement Pada Proyek Pemkot Surabaya Tahun 2018. *Jurnal Rekayasa Dan Manajemen Konstruksi*, 7(1), 33–41. <https://doi.org/10.30742/axial.v7i1.706>
- Aksan, I., & Nurfadilah, K. (2020). Aplikasi Metode Arima Box-Jenkins Untuk Meramalkan Penggunaan Harian Data Seluler. *JOMTA Journal of Mathematics: Theory and Applications*, 2(1), 1–10. <https://doi.org/10.31605/jomta.v2i1.749>
- Almahdali, H., Ismail, I., & Engka, R. (2022). Analisis Pelaksanaan Pengadaan Barang/Jasa Secara Elektronik (E-Procurement) pada Pemerintah Kabupaten Wajo. *Jurnal Ada Na Gau: Public Administration*, 3(1), 1031–1042.
- Ariyanti, V. P., & Yusnitasari, T. (2023). Comparison of Arima and Sarima for forecasting crude oil prices. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(2), 405–413. <https://doi.org/10.29207/resti.v7i2.4895>
- Arta, A., Almadevi, M., Ardine Puspitasari, P., & Septiani Pontoh, R. (2022). Peramalan Kandungan CO (Karbon Monoksida) DKI JAKARTA dengan Menggunakan Metode Arima. *Seminar Nasional Statistika Aktuaria I*, 1–10.
- Djufri, M. (2020). Penerapan Teknik Web Scraping Untuk Penggalan Potensi Pajak (Studi Kasus Pada Online Market Place Tokopedia, Shopee Dan Bukalapak). *JURNAL BPPK*, 13(5), 65–75. <https://doi.org/10.48108/jurnalbppk.v13i2.636>
- Fauziah, A. N., Sadida, H. Q., Rizki, M. I., & Pontoh, R. S. (2021). Peramalan Harga Beras Premium di Tingkat Penggilingan di Indonesia dengan Metode Pemodelan ARCH-GARCH. *Seminar Nasional Statistika Online*, 10–20.
- Ferdiansyah, A., & Nadjib, M. F. (2022). Rancangan Pre-Procurement, Procurement Process, Dan Pos-Recurement Untuk Memenangkan Tender Di Instansi Pemerintah Atau

- Perusahaan Swasta. *Prosiding The 13th Industrial Research Workshop and National Seminar*, 13–14.
- Haryadi, R., & Meilidhani, R. M. (2023). Penentuan Jumlah Produksi Dalam Industri Digital Printing Yang Berkhusus Pada Percetakan Banner Berdasarkan Data Persediaan Dan Jumlah Permintaan Menggunakan Metode Fuzzy Sugeno. *Journal Computer, Technology, and Informations System*, 2(2), 49–56.
- Hasibuan, L. H., Musthofa, S., Putri, D. M., & Jannah, M. (2023). Comparison Of Seasonal Time Series Forecasting Using Sarima And Holt Winter’S Exponential Smoothing (Case Study: West Sumatra Export Data). *BAREKENG: Journal of Mathematics and Its Applications*, 12(3), 1773–1784. <https://doi.org/10.30598/barekengvol17iss3pp1773-1784>
- Ignasius, M., & Lamabelawa, J. (2018). Perbandingan Interpolasi Dan Ekstrapolasi Newton Untuk Prediksidata Time Series. *HOAQ: Jurnal Teknologi Informasi*, 10(12), 73–80. <https://doi.org/10.52972/hoaq.vol10no2.p73-80>
- Innayah, A. N. (2024). Utilizing The Sarima Model And Support Vector Regression To Forecast Monthly Rainfall In Bandung City. *Jurnal Ilmiah Teknologi Infomasi Terapan*, 10(2), 100–110. <https://doi.org/10.33197/jitter.vol10.iss2.2024.1663>
- Liu, J., Zhao, Z., Zhong, Y., Zhao, C., & Zhang, G. (2022). Prediction of the dissolved gas concentration in power transformer oil based on SARIMA model. *Energy Reports*, 8(10), 1360–1367. <https://doi.org/10.1016/j.egy.2022.03.020>
- Noureen, S., Atique, S., Roy, V., & Bayne, S. (2019). Analysis and application of seasonal ARIMA model in Energy Demand Forecasting: A case study of small scale agricultural load. *2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS)*, 521–524. <https://doi.org/10.1109/MWSCAS.2019.8885349>
- Qonita, A., Pertiwi, A. G., & Widiyaningtyas, T. (2017). Prediction of Rupiah Against US Dollar by Using ARIMA. *4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 1–5, IEEE. <https://doi.org/10.1109/EECSI.2017.8239205>
- Ristijana, D. K., Dirgantoro, B., & Ruriawan, M. F. (2021). Implementasi Metode Proteksi Situs Web Dari Web Scraping. *E-Proceeding of Engineering*, 8(5), 210–217.
- Safira, A. N., Warsito, B., & Rusgiyono, A. (2022). Analisis Support Vector Regression (Svr) Dengan Algoritma Grid Search Time Series Cross Validation Untuk Prediksi Jumlah Kasus Terkonfirmasi Covid-19 Di Indonesia. *Jurnal Gaussian*, 11(4), 512–521. <https://doi.org/10.14710/j.gauss.11.4.512-521>
- Satria, A., Nugraha, P., Zakaria, S., & Lesmana, A. C. (2021). Implementasi E-Procurement Dalam Proses Pengadaan Barang/Jasa Pada Lembaga Layanan Pengadaan Secara Elektronik (LPSE) Kota Bogor Tahun 2020. *Jurnal Administrasi Pemerintahan (Janitra)*, 1(2), 73–87. <https://doi.org/10.24198/janitra.v1i2.38241>
- Sirisha, U. M., Belavagi, M. C., & Attigeri, G. (2022). Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison. *IEEE Access*, 10(1), 124715–124727. <https://doi.org/10.1109/ACCESS.2022.3224938>
- Subagia, I. N., Lestari, I. G. A. A. I., & Praditha, P. F. (2020). Analisis Strategi Penawaran Kontraktor Untuk Memenangkan Tender Online Di Denpasar. *Ganec Swara*, 14(2), 738–745.
- Susanto, V. Y. (2021). Nilai Pengadaan Barang dan Jasa Pemerintah pada 2021 Mencapai Rp 1.214 Triliun. *Jurnal Penelitian Pendidikan Indonesia*, 8(3), 668–675.