

PREDIKSI KLASTER KONSUMEN BERDASARKAN PERILAKU BELANJA MUSIMAN MENGUNAKAN METODE SARIMA

Elisabeth Monalisa Sawe¹, I Wayan Sudiarsa², Maria Yulita Runu Ke'e³, Anastasia Deveni Putri⁴, Maria Dewi Sartika⁵

Bisnis Digital¹³⁴⁵, Sitem Komputer², Institut Bisnis dan Teknologi Indonesia

E-mail: *sudiarsa@instiki.ac.id

ABSTRAK

Perilaku belanja konsumen cenderung mengalami fluktuasi yang dipengaruhi oleh faktor musiman seperti periode waktu tertentu, kebiasaan pembelian, dan tren konsumsi. Penelitian ini bertujuan untuk memprediksi pola belanja konsumen musiman sebagai dasar pembentukan kluster konsumen menggunakan pendekatan Artificial Intelligence. Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) digunakan untuk memodelkan data deret waktu penjualan yang bersumber dari dataset *shopping_trends.csv*. Tahapan penelitian meliputi pemahaman data, pra-pemrosesan data, uji stasioneritas menggunakan Augmented Dickey-Fuller (ADF), pemodelan SARIMA, serta evaluasi model menggunakan Root Mean Squared Error (RMSE) dan Mean Absolute Error (MAE). Hasil penelitian menunjukkan bahwa model SARIMA dengan parameter $(0,0,0)(2,1,0,7)$ mampu menangkap pola musiman mingguan dengan nilai RMSE sebesar 28,34 dan MAE sebesar 23,20. Hasil prediksi ini dapat dimanfaatkan sebagai dasar pengelompokan (klasterisasi) konsumen berdasarkan pola dan intensitas belanja musiman.

Kata kunci

Artificial Intelligence, SARIMA, Perilaku Belanja, Pola Musiman, Kluster Konsumen

ABSTRACT

Consumer shopping behavior tends to fluctuate due to seasonal factors such as time periods, purchasing habits, and consumption trends. This study aims to predict seasonal consumer shopping patterns as a basis for consumer clustering using an Artificial Intelligence approach. The Seasonal Autoregressive Integrated Moving Average (SARIMA) method is applied to model sales time-series data derived from the shopping_trends.csv dataset. The research stages include data understanding, data preprocessing, stationarity testing using the Augmented Dickey-Fuller (ADF) test, SARIMA modeling, and model evaluation using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). The results show that the SARIMA $(0,0,0)(2,1,0,7)$ model is able to capture weekly seasonal patterns with an RMSE value of 28.34 and an MAE of 23.20. These prediction results can be used as a basis for clustering consumers based on seasonal shopping behavior.

Keywords

Artificial Intelligence, SARIMA, Consumer Behavior, Seasonality, Consumer Clustering

1. PENDAHULUAN

Kecerdasan Buatan (AI) telah mengambil peran penting dalam melakukan analisis data transaksi konsumen dalam jumlah besar, yang pada gilirannya mendukung pengambilan keputusan berbasis data di berbagai sektor. Dengan kemampuan untuk memproses dan menganalisis data yang sangat besar, AI memungkinkan perusahaan untuk menggali wawasan yang lebih mendalam tentang perilaku konsumen dan merespons kebutuhan mereka dengan lebih efektif (Duan et al., 2019); (Enholm et al., 2021).

Pola belanja musiman dapat dipengaruhi oleh berbagai faktor, antara lain perubahan musim, hari besar nasional atau keagamaan, promosi berkala, serta kebiasaan konsumsi konsumen. Pemahaman terhadap pola ini menjadi sangat penting bagi perusahaan untuk menyusun strategi pemasaran, pengelolaan persediaan, serta segmentasi atau klusterisasi konsumen secara lebih efektif. Dengan mengetahui pola belanja yang bersifat musiman, perusahaan dapat mengelompokkan konsumen berdasarkan karakteristik perilaku belanjanya sehingga strategi yang diterapkan menjadi lebih tepat sasaran.

Perkembangan teknologi digital dan meningkatnya aktivitas transaksi berbasis data telah mendorong kebutuhan akan metode analisis yang mampu menangani kompleksitas dan dinamika perilaku konsumen. Data transaksi konsumen yang tersimpan dalam jumlah besar tidak hanya mewakili nilai penjualan, tetapi juga mengandung informasi pola perilaku yang dapat dijelaskan secara temporal. Oleh karena itu, pendekatan analitik berbasis deret waktu menjadi relevan untuk mengidentifikasi pola berulang dan kecenderungan musiman yang muncul dalam aktivitas belanja konsumen.

Berbagai pendekatan telah digunakan untuk menganalisis perilaku belanja konsumen, mulai dari metode statistik hingga pendekatan berbasis Artificial Intelligence. Salah satu pendekatan yang banyak digunakan untuk menganalisis data yang memiliki komponen musiman adalah metode peramalan deret waktu. Seasonal Autoregressive Integrated Moving Average (SARIMA) merupakan pengembangan dari metode ARIMA yang dirancang khusus untuk menangani data deret waktu dengan pola musiman. Metode ini mampu menangkap komponen tren, musiman, serta fluktuasi acak dalam data secara simultan.

Selain itu, pemanfaatan hasil peramalan dalam konteks klusterisasi konsumen memberikan nilai tambah strategi bagi pengambilan keputusan bisnis. Dengan mengombinasikan hasil prediksi pola musiman dan analisis perilaku belanja, perusahaan dapat mengembangkan segmentasi konsumen yang lebih adaptif terhadap perubahan waktu dan musim.

Meskipun metode SARIMA telah banyak digunakan dalam peramalan deret waktu, penerapannya sebagai dasar analisis klaster konsumen berdasarkan perilaku belanja musiman masih relatif terbatas, khususnya pada data transaksi konsumen. Oleh karena itu, penelitian ini bertujuan untuk memodelkan dan memprediksi perilaku belanja konsumen musiman menggunakan metode SARIMA. Hasil prediksi tersebut diharapkan dapat memberikan gambaran pola belanja konsumen yang bersifat musiman dan menjadi dasar dalam pembentukan klaster konsumen berdasarkan intensitas serta kecenderungan belanja pada periode tertentu Tang, C., Wang, L., Wei, Y., Wu, P., & Wei, H. (2023). Time-Frequency Domain Variation Analysis and LSTM

Forecasting of Regional Visibility in the China Region Based on GSOD Station Data. Atmosphere, 14(7), 1072. b

Secara khusus, tujuan penelitian ini adalah: (1) membentuk model peramalan perilaku belanja konsumen musiman menggunakan metode SARIMA, (2) mengevaluasi kinerja model berdasarkan metrik kesalahan prediksi, dan (3) menganalisis potensi pemanfaatan hasil peramalan sebagai dasar klasterisasi konsumen. Penelitian ini diharapkan dapat memberikan kontribusi dalam penerapan Artificial Intelligence dan analitik deret waktu untuk analisis perilaku konsumen.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam studi ini mengikuti tahapan analisis data berbasis Artificial Intelligence dan peramalan deret waktu. Tahapan penelitian meliputi pemahaman data, pra-pemrosesan data, analisis stasioneritas, pemodelan SARIMA, serta evaluasi kinerja model.

Pemilihan metode SARIMA dalam penelitian ini didasarkan pada kemampuannya dalam menangani data deret waktu yang memiliki komponen musiman secara eksplisit. Berbeda dengan metode peramalan non-musiman, SARIMA memungkinkan integrasi parameter musiman yang dapat menangkap pola berulang dalam periode tertentu. Hal ini menjadikan SARIMA mampu menganalisis perilaku belanja konsumen yang dipengaruhi oleh siklus mingguan maupun musiman lainnya.

Seluruh proses pemodelan dan evaluasi dilakukan secara berurutan untuk memastikan tidak terjadi kebocoran data antara data latih dan data uji. Dengan demikian, hasil evaluasi yang diperoleh mencerminkan kemampuan model dalam melakukan generalisasi terhadap data yang belum pernah dilihat sebelumnya. Pendekatan ini diharapkan mampu menghasilkan model peramalan yang lebih reliabel dan dapat diterapkan dalam konteks analisis perilaku konsumen berbasis waktu.

2.1 Fase Pemahaman Bisnis (*Business Understanding Phase*)

Dataset yang digunakan dalam penelitian ini adalah *shopping_trends.csv* yang berisi data transaksi konsumen. Dataset ini mencakup berbagai atribut yang merepresentasikan karakteristik konsumen dan transaksi, seperti identitas pelanggan, usia, jenis kelamin, kategori produk, musim pembelian, lokasi, serta nilai transaksi (*Purchase Amount (USD)*).

	Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Payment Method	Shipping Type	Discount Applied	Promo Code Used	Previous Purchases	Preferred Payment Method	Frequency of Purchases
0	1	55	Male	Blouse	Clothing	53	Kentucky	L	Gray	Winter	3.1	Yes	Credit Card	Express	Yes	Yes	14	Venmo	Fortnightly
1	2	19	Male	Sweater	Clothing	64	Maine	L	Maroon	Winter	3.1	Yes	Bank Transfer	Express	Yes	Yes	2	Cash	Fortnightly
2	3	50	Male	Jeans	Clothing	73	Massachusetts	S	Maroon	Spring	3.1	Yes	Cash	Free Shipping	Yes	Yes	23	Credit Card	Weekly
3	4	21	Male	Sandals	Footwear	90	Rhode Island	M	Maroon	Spring	3.5	Yes	PayPal	Next Day Air	Yes	Yes	49	PayPal	Weekly
4	5	45	Male	Blouse	Clothing	49	Oregon	M	Turquoise	Spring	2.7	Yes	Cash	Free Shipping	Yes	Yes	31	PayPal	Annually

Tabel 1. Deskripsi Atribut Dataset *shopping_trends.csv*

Tabel ini menyajikan informasi mengenai atribut-atribut utama dalam dataset beserta deskripsinya. Atribut *Purchase Amount (USD)* dipilih sebagai variabel utama karena merepresentasikan intensitas belanja konsumen yang akan dianalisis sebagai deret waktu penjualan.

2.2 Fase Pemahaman Data (*Data Understanding Phase*)

Tahap pra-pemrosesan bertujuan untuk mengubah data transaksi mentah menjadi data deret waktu yang siap digunakan dalam pemodelan SARIMA. Karena dataset tidak memiliki kolom tanggal transaksi aktual, maka dilakukan pembuatan tanggal sintetis dengan interval harian yang berurutan. Selanjutnya, data transaksi diagregasi berdasarkan tanggal sintetis untuk memperoleh nilai penjualan harian.

Proses selanjutnya adalah melakukan resampling data ke frekuensi harian untuk memastikan konsistensi interval waktu. Apabila terdapat nilai yang hilang akibat proses agregasi atau resampling, maka digunakan metode interpolasi linier untuk mengisi nilai tersebut. Tahapan ini penting untuk menjaga kontinuitas deret waktu dan meningkatkan kualitas model peramalan.

```
Tanggal
0 2023-01-01
1 2023-01-02
2 Januari 2023
3 2023-01-04
4 2023-01-05

Header DataFrame Penjualan Harian:
      Tanggal  Jumlah Pembelian (USD)
0 2023-01-01    53
1 2023-01-02    64
2 2023-01-03    73
3 2023-01-04    90
4 2023-01-05    49

Informasi DataFrame Penjualan Harian:
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 3900 entri, 2023-01-01 hingga 2033-09-04
Kolom data (total 1 kolom):
# Kolom Jumlah Non-Null Dtype
-----
0 Jumlah Pembelian (USD) 3900 non-null int64
tipe data: int64(1)
Penggunaan memori: 60,9 KB
```

Gambar 1. Alur Pra-pemrosesan Data Deret Waktu

Gambar ini menunjukkan tahapan pra-pemrosesan data mulai dari pembuatan tanggal sintetis, agregasi penjualan harian, resampling data, hingga interpolasi nilai yang hilang.

Diagram alur pada Gambar ini memberikan gambaran sistematis mengenai tahapan pra-pemrosesan data deret waktu. Setiap tahapan dirancang untuk memastikan data memenuhi persyaratan pemodelan SARIMA, khususnya terkait konsistensi interval waktu dan kelengkapan data. Proses ini menjadi fondasi penting dalam menghasilkan model peramalan yang akurat dan stabil.

2.3 Fase Pengolahan Data (*Data Preparation Phase*)

Salah satu asumsi penting dalam pemodelan SARIMA adalah bahwa data deret waktu bersifat stasioner, yaitu memiliki nilai rata-rata dan varians yang konstan sepanjang waktu. Untuk menguji stasioneritas data, digunakan metode Augmented Dickey-Fuller (ADF).

Gambar ini menampilkan nilai statistik ADF, p-value, serta keputusan uji. Berdasarkan

Tes Dickey-Fuller yang Diperluas: Seri Asli

Statistik ADF: -61,0876

nilai p: 0,0000

#Lag yang Digunakan: 0.0000

Jumlah Pengamatan yang Digunakan: 3899.0000

Bukti kuat yang menentang hipotesis nol, tolak H_0 . Data tidak memiliki akar unit dan bersifat stasioner.

Tes Dickey-Fuller yang Diperluas: Seri yang Dibedakan

Statistik ADF: -18,9285

nilai p: 0,0000

#Lag yang Digunakan: 30.0000

Jumlah Pengamatan yang Digunakan: 3868.0000

Bukti kuat yang menentang hipotesis nol, tolak H_0 . Data tidak memiliki akar unit dan bersifat stasioner.

Gambar 2. Hasil Uji Augmented Dickey-Fuller (ADF)

hasil pengujian, diperoleh nilai p-value $< 0,05$ yang menunjukkan bahwa data penjualan harian telah bersifat stasioner dan memenuhi syarat untuk pemodelan SARIMA.

Visualisasi hasil uji ADF pada Gambar 2 berperan sebagai validasi statistik terhadap sifat stasioner data. Keputusan stasioneritas yang diperoleh menjadi dasar dalam menentukan perbedaan parameter pada model SARIMA, sehingga risiko kesalahan spesifikasi model dapat diminimalkan.

2.4 Fase Pemodelan Data (Modelling Phase)

Data deret waktu yang telah dipra-pemrosesan dibagi menjadi dua bagian secara berurutan berdasarkan waktu. Sebanyak 80% data digunakan sebagai data latih (training data), sedangkan 20% sisanya digunakan sebagai data uji (testing data). Pembagian ini bertujuan untuk mengevaluasi kemampuan model dalam memprediksi data pada periode waktu yang belum pernah digunakan dalam proses pelatihan.

2.5 Fase Evaluasi (Evaluation Phase)

Evaluasi kinerja model dilakukan untuk mengukur tingkat akurasi hasil peramalan. Metrik evaluasi yang digunakan dalam penelitian ini adalah Root Mean Squared Error (RMSE) dan Mean Absolute Error (MAE), yang mengukur selisih antara nilai aktual dan nilai hasil prediksi.

Kesalahan Kuadrat Rata-rata (RMSE): 28,34

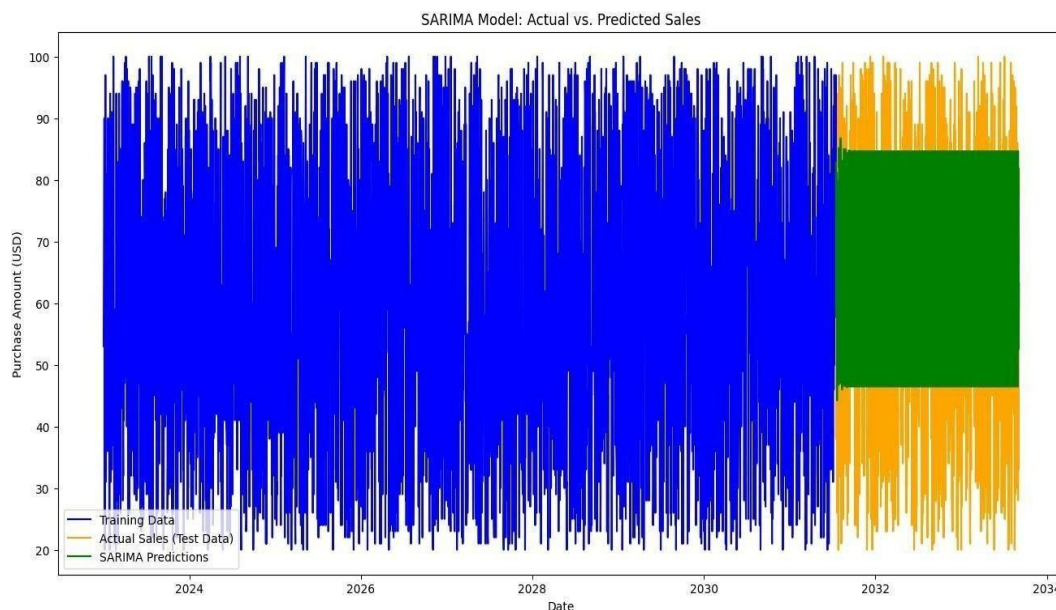
Kesalahan Absolut Rata-rata (MAE): 23,20

Gambar 3. Hasil Evaluasi Model SARIMA

Nilai RMSE dan MAE yang diperoleh menunjukkan bahwa model SARIMA memiliki tingkat kesalahan prediksi yang relatif rendah dan mampu merepresentasikan pola belanja musiman dengan cukup baik.

3. HASIL DAN PEMBAHASAN

Bagian ini membahas hasil pemodelan dan peramalan perilaku belanja konsumen menggunakan metode SARIMA serta interpretasi terhadap hasil yang diperoleh. Analisis dilakukan berdasarkan hasil evaluasi kuantitatif dan visualisasi data deret waktu



Gambar 4. Perbandingan Data Aktual dan Hasil Prediksi SARIMA

Gambar 4 menunjukkan perbandingan antara data penjualan aktual pada data uji dan hasil prediksi yang dihasilkan oleh model SARIMA. Secara visual, terlihat bahwa pola prediksi mengikuti fluktuasi data aktual dengan cukup baik, khususnya pada pola musiman mingguan. Hal ini menunjukkan bahwa model mampu menangkap karakteristik periodik yang terdapat dalam data perilaku belanja konsumen.

Hasil evaluasi model menunjukkan nilai Root Mean Squared Error (RMSE) sebesar 28,34 dan Mean Absolute Error (MAE) sebesar 23,20. Nilai ini mengindikasikan bahwa selisih antara nilai prediksi dan nilai aktual relatif moderat, sehingga model SARIMA dapat dikategorikan memiliki kinerja yang cukup baik dalam melakukan peramalan jangka pendek. Kesalahan prediksi yang masih muncul terutama disebabkan oleh fluktuasi ekstrem pada beberapa periode tertentu yang sulit ditangkap sepenuhnya oleh model statistik.

Parameter musiman yang digunakan dalam model, yaitu periode 7, mengindikasikan adanya pola belanja mingguan pada data transaksi konsumen. Hal ini sejalan dengan karakteristik perilaku belanja masyarakat yang umumnya mengalami peningkatan atau penurunan pada hari-hari tertentu dalam satu minggu. Komponen autoregressive musiman pada lag ke-7 dan ke-14 menunjukkan bahwa nilai penjualan pada periode sebelumnya memiliki pengaruh signifikan terhadap nilai penjualan pada periode berikutnya.

Hasil peramalan ini memiliki implikasi penting dalam analisis kluster konsumen. Pola musiman yang berhasil diprediksi dapat digunakan sebagai dasar pengelompokan konsumen berdasarkan intensitas dan konsistensi belanja pada periode tertentu. Konsumen dengan pola belanja tinggi pada periode musiman tertentu dapat dikelompokkan ke dalam kluster yang berbeda dengan konsumen yang memiliki pola belanja stabil atau rendah. Gambar ini tidak hanya menunjukkan kedekatan

antara nilai aktual dan nilai prediksi, tetapi juga mengilustrasikan kemampuan model dalam mengikuti dinamika data. Perbedaan yang muncul pada beberapa titik ekstrem menunjukkan adanya volatilitas yang bersifat acak, yang secara umum sulit dimodelkan secara sempurna oleh pendekatan statistik.

Temuan penelitian ini sejalan dengan penelitian-penelitian sebelumnya yang menyatakan bahwa metode SARIMA efektif dalam menangkap pola musiman pada data deret waktu, khususnya untuk peramalan jangka pendek. Namun demikian, keterbatasan penelitian ini terletak pada penggunaan tanggal sintetis yang belum sepenuhnya merepresentasikan kondisi transaksi riil. Oleh karena itu, hasil penelitian ini masih dapat dikembangkan lebih lanjut dengan menggunakan data transaksi aktual dan dikombinasikan dengan metode klasterisasi eksplisit berbasis Artificial Intelligence.

4. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, berikut adalah beberapa kesimpulan utama:

- a. Metode SARIMA berhasil memodelkan dan memprediksi pola belanja konsumen musiman berdasarkan data deret waktu penjualan yang bersumber dari dataset *shopping_trends.csv*. Model SARIMA (0,0,0)(2,1,0,7) mampu menangkap pola musiman mingguan yang mencerminkan fluktuasi perilaku belanja konsumen secara periodik.
- b. Hasil evaluasi model menunjukkan tingkat akurasi yang cukup baik, dengan nilai Root Mean Squared Error (RMSE) sebesar 28,34 dan Mean Absolute Error (MAE) sebesar 23,20. Nilai tersebut menunjukkan bahwa selisih antara hasil prediksi dan data aktual relatif moderat, sehingga model layak digunakan untuk peramalan jangka pendek.
- c. Pola musiman yang dihasilkan dari model SARIMA dapat dimanfaatkan sebagai dasar pembentukan klaster konsumen, khususnya dalam mengelompokkan konsumen berdasarkan intensitas dan konsistensi belanja pada periode waktu tertentu. Hal ini menunjukkan bahwa pendekatan Artificial Intelligence berbasis peramalan deret waktu memiliki potensi untuk mendukung strategi segmentasi konsumen dan pengambilan keputusan bisnis berbasis data

5. DAFTAR PUSTAKA

- A data-driven clustering method for time course gene expression data. *nucleic acids research*, 34(4), 1261-1269. <https://doi.org/10.1093/nar/gkl013>
- Cong, J., Ren, M., Xie, S., & Wang, P. (2019). Predicting Seasonal Influenza Based on SARIMA Model, in Mainland China from 2005 to 2018. *International Journal of Environmental Research and Public Health*, 16(23), 4760. <https://doi.org/10.3390/ijerph16234760>
- Development of a scale to measure the perceived benefits and risks of online shopping. *journal of interactivemarketing* 20(2), 5575. <https://doi.org/10.1002/dir.20061>
- Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*, 48, 63-71. <https://doi.org/10.1016/j.ijinfomgt.2019.01.021>

- Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2021). Artificial Intelligence and Business Value: a Literature Review. *Information Systems Frontiers*, 24(5), 1709-1734. <https://doi.org/10.1007/s10796-02110186-w>
- Fildes, R., & Ma, S. (2015). Peramalan dan riset operasional: Sebuah tinjauan. *Jurnal Masyarakat Riset Operasional*, 66(12), 2042–2053. <https://doi.org/10.1057/jors.2015.28>
- Going shopping: key determinants of shopping behaviors and motivations. *international journal of retail & distribution management*, 27(4), 154-165. <https://doi.org/10.1108/09590559910268499>
- Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. *bmc medical research methodology*, 21(1). <https://doi.org/10.1186/s12874-02101235-8>
- Kumar, S. V. and Vanajakshi, L. (2015). Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review*, 7(3). <https://doi.org/10.1007/s12544-015-0170-8>
- Madrid, E. A. and António, N. (2021). Short-Term Electricity Load Forecasting with Machine Learning. *Information*, 12(2), 50. <https://doi.org/10.3390/info12020050>
- Measuring festival quality and value affecting visitors' satisfaction and loyalty using a structural approach. *international journal of hospitality management*, 29(2), 335-342. <https://doi.org/10.1016/j.ijhm.2009.10.002>
- Multichannel customer management: Understanding the research-shopper phenomenon. *international journal of research in marketing*, 24(2), 129-148. <https://doi.org/10.1016/j.ijresmar.2006.11.002>
- Predicting Potential Evapotranspiration for Kalaburagi District using a Seasonal Arima Model. *international journal of environment and climate change*, 13(11), 2073-2082. <https://doi.org/10.9734/ijecc/2023/v13i113367>
- Purchasing, Pricing, and Quick Response in the Presence of Strategic Consumers. *management science*, 55(3), 497-511. <https://doi.org/10.1287/mnsc.1080.0948>
- Rotella, P., Salomon, F. L. R., & Pamplona, E. d. O. (2014). ARIMA: An Applied Time Series Forecasting Model for the Bovespa Stock Index. *Applied Mathematics*, 05(21), 3383-3391. <https://doi.org/10.4236/am.2014.521315>
- Scatter in fatigue life due to effects of porosity in cast A356-T6 aluminum-silicon alloys. *metallurgical and materials transactions a*, 34(9), 1879-1890. <https://doi.org/10.1007/s11661-003-0153-6>
- Tang, C., Wang, L., Wei, Y., Wu, P., & Wei, H. (2023). Time-Frequency Domain Variation Analysis and LSTM Forecasting of Regional Visibility in the China Region Based on GSOD Station Data. *Atmosphere*, 14(7), 1072. <https://doi.org/10.3390/atmos14071072>
- The impact of website quality on customer satisfaction and purchase intentions: Evidence from Chinese online visitors. *international journal of hospitality management*, 27(3), 391-402. <https://doi.org/10.1016/j.ijhm.2007.10.008>
- Tsai, CF, & Chiu, CC (2004). Metodologi segmentasi pasar berbasis pembelian. *Expert Systems with Applications*, 27(2), 265–276. <https://doi.org/10.1016/j.eswa.2004.03.006>

- Wang, K., Deng, C., Li, J. P., Zhang, Y. Y., Li, X., & Wu, M. (2017). Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiology and Infection*, 145(6), 1118- 1129.
<https://doi.org/10.1017/s0950268816003216>
- Wang, H., Tian, C., Wang, W. M., & Luo, X. (2018). Time-series analysis of tuberculosis from 2005 to 2017 in China. *Epidemiology and Infection*, 146(8), 935-939. <https://doi.org/10.1017/s0950268818001111>