

Peramalan Penerimaan Pajak Daerah Menggunakan ARIMA, LSTM, dan XGBoost: Studi Kasus Kabupaten Bantul

Brian Rizadhani Latuconsina^{1*}, Achmad Nizar Hidayanto²

^{1,2}Program Magister Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Indonesia, Depok, Indonesia.

ARTICLE INFO

Article History:

Received Jun 25th, 2026

Revised July 20th, 2026

Accepted Aug 10th, 2026

Keywords:

Pendapatan Asli Daerah,
Peramalan,
ARIMA,
XGBoost,
Tapping Box.

ABSTRAKSI

The Bantul Regency's own-source revenue (PAD), derived from hotel, restaurant and parking taxes, has been collected electronically via tapping boxes. However, the transaction data collected over the years has not been utilised optimally to support local revenue forecasting. Interviews with the Regional Financial, Revenue and Asset Management Agency (BPKPAD) of Bantul Regency revealed that the revenue estimation process has so far been based on intuition and simple trend analysis using spreadsheet applications. This study aims to identify the most accurate forecasting model for hotel, restaurant and car park tax revenue by comparing the ARIMA and SARIMA methods as statistical approaches, Long Short-Term Memory (LSTM) as a deep learning approach, and XGBoost as a machine learning approach based on the CRISP-DM framework. The data used consists of approximately 4.8 million tapping box transactions for the period 2020–2025, which were processed into time series data and evaluated using the rolling-origin backtest method with the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) indicators. The research results indicate that no single model consistently delivers the best accuracy across all sectors and forecasting horizons. XGBoost produced the highest accuracy in the restaurant sector with the lowest MAPE value of 1.86 per cent, whilst ARIMA(1,1,1) was the best model

Pendapatan Asli Daerah (PAD) Kabupaten Bantul yang bersumber dari pajak hotel, restoran, dan parkir telah dipungut secara elektronik melalui perangkat *tapping box*. Namun, data transaksi yang telah terkumpul selama bertahun-tahun belum dimanfaatkan secara optimal untuk mendukung peramalan penerimaan daerah. Hasil wawancara dengan Badan Pengelolaan Keuangan, Pendapatan, dan Aset Daerah (BPKPAD) Kabupaten Bantul menunjukkan bahwa proses estimasi penerimaan selama ini masih dilakukan berdasarkan intuisi dan analisis tren sederhana menggunakan aplikasi spreadsheet. Penelitian ini bertujuan untuk mengidentifikasi model peramalan yang paling akurat untuk penerimaan pajak hotel, restoran, dan parkir dengan membandingkan metode ARIMA dan SARIMA sebagai pendekatan statistik, *Long Short-Term Memory* (LSTM) sebagai pendekatan *deep learning*, serta XGBoost sebagai pendekatan *machine learning* berdasarkan kerangka kerja CRISP-DM. Data yang digunakan terdiri atas sekitar 4,8 juta transaksi *tapping box* periode 2020–2025 yang diolah menjadi data deret waktu dan dievaluasi menggunakan metode *rolling-origin backtest* dengan indikator *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). Hasil penelitian menunjukkan bahwa tidak terdapat satu model yang secara konsisten memberikan akurasi terbaik pada seluruh sektor dan horizon peramalan. XGBoost menghasilkan akurasi tertinggi pada sektor restoran dengan nilai MAPE terendah sebesar 1,86%, sedangkan ARIMA(1,1,1) menjadi model terbaik pada sektor parkir dengan nilai MAPE berkisar antara 7% hingga 9%. Pada sektor hotel yang memiliki tingkat volatilitas tinggi, model *Seasonal Naïve* memberikan hasil terbaik dengan nilai MAPE antara 10% hingga 14%. Temuan ini menunjukkan bahwa pemilihan model peramalan harus disesuaikan dengan karakteristik masing-masing sektor serta menegaskan pentingnya kualitas data *tapping box* dalam mendukung perencanaan penerimaan daerah yang lebih akurat dan berbasis data.



© 2026 The Author(s). Published by Universitas Indraprasta PGRI, Jakarta, Indonesia. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

Corresponding Author:

Brian Rizadhani Latuconsina

Email: brian.rizadhani@ui.ac.id

How to Cite: Latuconsina, B. R., Hidayanto, A. N. (2026). Peramalan Penerimaan Pajak Daerah Menggunakan ARIMA, LSTM, dan Xgboost: Studi Kasus Kabupaten Bantul. *Sosio e-Kons*, 18 (2), 194-205

PENDAHULUAN

Transformasi digital telah menjadi agenda strategis dalam tata kelola sektor publik di berbagai negara karena kemampuannya meningkatkan efektivitas manajemen berbasis data dan kualitas pengambilan keputusan. Pemanfaatan teknologi digital seperti *big data*, *artificial intelligence* (AI), dan *Internet of Things* (IoT) memungkinkan pemerintah mengelola informasi secara *real-time*, meningkatkan efisiensi operasional, memperkuat transparansi, serta mempercepat penyediaan layanan publik yang responsif terhadap kebutuhan masyarakat. Berbagai penelitian menunjukkan bahwa integrasi teknologi digital tidak hanya mendukung modernisasi administrasi publik, tetapi juga mendorong terbentuknya model tata kelola yang lebih adaptif, akuntabel, dan berorientasi pada data dalam menghadapi tantangan pembangunan yang semakin kompleks (Abdullah Sani & Jaafar, 2026; Valle-Cruz & García-Contreras, 2025; Yukhno, 2024).

Digitalisasi administrasi perpajakan semakin diakui sebagai strategi penting untuk meningkatkan transparansi, efisiensi, dan optimalisasi penerimaan pemerintah. Berbagai negara telah memanfaatkan teknologi digital untuk memperkuat pengawasan kepatuhan wajib pajak, memperluas basis pajak, serta meningkatkan efektivitas pengumpulan pendapatan negara dan daerah (He & Yi, 2023; Ongo Nkoa & Song, 2022). Namun, keberhasilan digitalisasi perpajakan tidak hanya ditentukan oleh ketersediaan teknologi, tetapi juga oleh kemampuan pemerintah dalam memanfaatkan data yang dihasilkan untuk mendukung pengambilan keputusan berbasis bukti (Lee, 2025).

Di Indonesia, modernisasi administrasi pajak daerah melalui pemanfaatan teknologi digital terus didorong untuk meningkatkan kualitas pelayanan publik dan optimalisasi penerimaan daerah. Transformasi ini sejalan dengan perkembangan *financial technology* (FinTech) dan digitalisasi sektor publik yang memungkinkan proses pengelolaan pajak menjadi lebih efektif, transparan, dan berbasis data. Namun, keberhasilan implementasinya sangat bergantung pada kesiapan infrastruktur, kapasitas kelembagaan, serta kemampuan pemerintah daerah dalam mengelola dan memanfaatkan data digital untuk mendukung pengambilan keputusan yang lebih akurat (David et al., 2023; Lubis et al., 2021).

Salah satu bentuk implementasi digitalisasi perpajakan daerah adalah penggunaan *tapping box*, yaitu perangkat pencatat transaksi yang terhubung langsung dengan wajib pajak dan mengirimkan data transaksi secara *real-time* kepada pemerintah daerah. Kehadiran teknologi ini tidak hanya meningkatkan transparansi dan pengawasan penerimaan pajak, tetapi juga menghasilkan data transaksi yang dapat dimanfaatkan untuk mendukung perencanaan fiskal, penyusunan target pendapatan, serta pengambilan kebijakan berbasis data. Dengan ketersediaan informasi yang lebih akurat dan berkelanjutan, pemerintah daerah memiliki peluang untuk mengembangkan strategi pengelolaan pendapatan yang lebih efektif dan responsif terhadap perubahan kondisi ekonomi (Chen & He, 2024; W. Li & Li, 2024).

Meskipun data transaksi pajak digital terus meningkat seiring implementasi *tapping box*, pemanfaatan data tersebut untuk mendukung perencanaan dan prediksi penerimaan daerah masih relatif terbatas. Banyak pemerintah daerah masih mengandalkan pendekatan manual, pengalaman praktisi, atau analisis sederhana dalam menetapkan target pendapatan, sehingga berpotensi menghasilkan estimasi yang kurang akurat. Padahal, peramalan penerimaan yang tepat sangat penting untuk mendukung penyusunan anggaran, pengelolaan fiskal, dan pengambilan kebijakan yang lebih efektif. Kondisi ini menunjukkan perlunya pemanfaatan metode analisis data yang lebih sistematis untuk mengubah data transaksi pajak menjadi informasi yang bernilai bagi perencanaan pendapatan daerah.

Dalam konteks pengelolaan keuangan daerah, kemampuan melakukan peramalan penerimaan pajak secara akurat menjadi faktor penting dalam mendukung penyusunan target pendapatan, perencanaan anggaran, dan pengambilan kebijakan fiskal. Informasi mengenai potensi penerimaan pada periode mendatang memungkinkan pemerintah daerah menyusun strategi yang lebih tepat dalam mengelola sumber pendapatan serta mengantisipasi perubahan kondisi ekonomi. Oleh karena itu, pemanfaatan data transaksi pajak yang tersedia melalui sistem digital perlu diikuti dengan penerapan metode peramalan yang mampu menghasilkan estimasi penerimaan secara lebih akurat dan objektif.

Secara konseptual, transformasi digital dalam pengelolaan pajak daerah membentuk suatu rantai nilai yang saling terkait, mulai dari digitalisasi proses administrasi, akumulasi data transaksi, pemanfaatan data melalui metode peramalan, hingga pengambilan keputusan fiskal berbasis bukti. Digitalisasi menghasilkan data transaksi yang masif dan berkelanjutan, namun data tersebut baru memberikan nilai apabila diolah menjadi informasi prediktif yang relevan bagi perencanaan. Dalam kerangka pengambilan keputusan berbasis data, metode peramalan berperan sebagai penghubung antara ketersediaan data digital dan kualitas keputusan fiskal, khususnya dalam penetapan target penerimaan dan penyusunan anggaran daerah. Dengan demikian, keberhasilan digitalisasi perpajakan tidak berhenti pada aspek pencatatan, melainkan ditentukan oleh kemampuan mengubah data menjadi dasar pengambilan keputusan yang lebih akurat dan objektif.

Berbagai penelitian telah mengembangkan metode peramalan untuk meningkatkan akurasi prediksi penerimaan pajak dan pendapatan daerah. Pendekatan statistik dan ekonometrika, seperti *Support Vector Regression* (SVR), terbukti mampu menghasilkan prediksi yang akurat dengan memanfaatkan karakteristik data historis dan indikator ekonomi yang relevan (Yu, 2024). Penelitian lain menunjukkan bahwa penggunaan indikator ekonomi regional dapat meningkatkan ketepatan estimasi basis pajak dibandingkan pendekatan konvensional yang hanya mengandalkan data historis penerimaan (Rich et al., 2005). Temuan-temuan tersebut menunjukkan bahwa pemilihan metode peramalan yang tepat menjadi faktor penting dalam mendukung perencanaan fiskal dan pengelolaan pendapatan daerah yang lebih efektif.

Perkembangan metode *forecasting* menunjukkan pergeseran dari pendekatan statistik tradisional menuju pemanfaatan *machine learning* dan *deep learning*. Metode statistik seperti *Autoregressive Integrated Moving Average* (ARIMA) telah lama digunakan untuk memodelkan pola linier pada data deret waktu, namun memiliki keterbatasan dalam menangkap hubungan nonlinier yang kompleks. Seiring berkembangnya teknologi komputasi, model *deep learning* seperti *Long Short-Term Memory* (LSTM) mampu memberikan akurasi yang lebih baik dalam mengidentifikasi pola, tren, dan musiman pada data deret waktu yang kompleks (X. Li et al., 2023; Song et al., 2024). Selain itu, integrasi *machine learning* dan *AutoML* semakin memperluas peluang pengembangan model peramalan yang lebih adaptif dan efisien untuk berbagai kebutuhan prediksi (Alsharef et al., 2022).

Metode *Autoregressive Integrated Moving Average* (ARIMA) merupakan salah satu pendekatan yang paling banyak digunakan dalam peramalan deret waktu, termasuk untuk memprediksi pendapatan daerah dan indikator ekonomi. Keunggulan utama ARIMA terletak pada kemampuannya memodelkan pola historis dan tren linier secara relatif sederhana namun efektif ((Kumbasaroglu, 2026; Zhu et al., 2018). Meskipun demikian, beberapa penelitian menunjukkan bahwa ARIMA memiliki keterbatasan dalam menangkap hubungan nonlinier dan sensitif terhadap data yang mengandung fluktuasi tinggi atau *noise*, sehingga berbagai pengembangan dan kombinasi metode terus dilakukan untuk meningkatkan akurasi prediksi (Y. Li et al., 2023). Temuan tersebut menunjukkan bahwa evaluasi dan perbandingan dengan metode yang lebih modern menjadi penting untuk memperoleh model peramalan yang optimal.

Model *Long Short-Term Memory* (LSTM) semakin banyak digunakan dalam peramalan deret waktu karena kemampuannya menangkap hubungan nonlinier dan ketergantungan jangka panjang yang sulit dimodelkan oleh metode statistik tradisional. Beberapa penelitian menunjukkan bahwa LSTM mampu menghasilkan akurasi prediksi yang lebih baik dibandingkan ARIMA dan SARIMA pada data keuangan maupun penerimaan pemerintah (Jothi, 2025; Mahmud et al., 2024). Namun, temuan lain menunjukkan bahwa kinerja LSTM tidak selalu unggul pada seluruh kasus, terutama ketika karakteristik data lebih sesuai dimodelkan menggunakan pendekatan hibrida atau metode alternatif lainnya (Scherer, 2025; Thayyib et al., 2023). Hal ini mengindikasikan bahwa efektivitas LSTM sangat dipengaruhi oleh karakteristik data dan konteks peramalan yang digunakan.

Berbagai penelitian menunjukkan bahwa tidak terdapat satu model peramalan yang secara konsisten memberikan kinerja terbaik pada seluruh jenis data dan konteks aplikasi. Akurasi model sangat dipengaruhi oleh karakteristik data, pola musiman, tingkat volatilitas, serta tujuan peramalan yang ingin dicapai (Zhao et al., 2021). Temuan tersebut mengindikasikan bahwa pemilihan model peramalan tidak

dapat dilakukan secara universal, melainkan harus disesuaikan dengan karakteristik kasus yang dianalisis. Oleh karena itu, evaluasi komparatif terhadap berbagai metode peramalan menjadi pendekatan yang penting untuk mengidentifikasi model yang paling sesuai dan menghasilkan tingkat akurasi yang optimal.

Perkembangan metode peramalan deret waktu memperlihatkan evolusi yang bertahap seiring meningkatnya kompleksitas data dan kapasitas komputasi. Pendekatan statistik klasik seperti ARIMA berkembang menjadi SARIMA untuk menangkap komponen musiman, kemudian bergeser ke metode machine learning seperti XGBoost yang mampu memodelkan hubungan nonlinier melalui rekayasa fitur, dan berlanjut ke metode deep learning seperti LSTM yang dirancang untuk menangkap ketergantungan jangka panjang. Meskipun masing-masing pendekatan telah banyak diteliti secara terpisah, perbandingan menyeluruh yang mencakup keempat kelompok metode tersebut dalam satu kerangka evaluasi masih jarang dilakukan, terlebih pada data penerimaan pajak daerah di Indonesia. Kondisi ini menegaskan perlunya studi komparatif yang mengevaluasi seluruh spektrum metode, dari statistik konvensional hingga deep learning, pada karakteristik data pajak daerah yang nyata.

Meskipun berbagai penelitian telah menerapkan metode ARIMA, LSTM, dan XGBoost untuk berbagai kebutuhan peramalan, masih terdapat kesenjangan penelitian yang perlu diperhatikan. Sebagian besar studi hanya berfokus pada penerapan satu metode tertentu atau membandingkan metode pada konteks yang berbeda, seperti pariwisata, permintaan pasar, dan sektor ekonomi lainnya. Hingga saat ini, penelitian yang secara komprehensif membandingkan kinerja ARIMA, LSTM, dan XGBoost dalam memprediksi penerimaan pajak hotel, restoran, dan parkir pada tingkat kabupaten atau kota di Indonesia masih sangat terbatas. Akibatnya, belum tersedia bukti empiris yang memadai mengenai model peramalan yang paling sesuai untuk karakteristik data penerimaan pajak daerah pada sektor-sektor tersebut.

Kesenjangan penelitian tersebut dapat ditinjau dari beberapa aspek. Dari aspek literatur, kajian yang membandingkan metode statistik, machine learning, dan deep learning secara simultan pada penerimaan pajak daerah masih terbatas. Dari aspek metodologis, sebagian besar studi hanya menguji satu horizon peramalan atau satu tingkat agregasi data, sehingga belum memberikan gambaran menyeluruh mengenai kestabilan kinerja model pada horizon yang berbeda. Dari aspek kontekstual, penelitian yang menggunakan data transaksi aktual berbasis tapping box pada sektor pajak hotel, restoran, dan parkir di tingkat kabupaten di Indonesia masih sangat jarang ditemukan. Dari aspek praktis, belum tersedia rekomendasi model yang spesifik untuk masing-masing sektor pajak daerah sebagai dasar penyusunan target penerimaan yang lebih objektif.

Berdasarkan kesenjangan tersebut, kebaruan penelitian ini dapat ditegaskan dalam tiga aspek. Pertama, dari aspek data, penelitian ini memanfaatkan data transaksi tapping box aktual pada sektor pajak hotel, restoran, dan parkir di Kabupaten Bantul yang bersifat masif dan berkelanjutan. Kedua, dari aspek metode, penelitian ini membandingkan secara simultan metode statistik (ARIMA dan SARIMA), machine learning (XGBoost), dan deep learning (LSTM) dalam satu kerangka evaluasi yang konsisten pada beberapa horizon peramalan. Ketiga, dari aspek kontribusi praktis, penelitian ini menghasilkan rekomendasi model terbaik untuk masing-masing sektor pajak daerah sebagai dasar pengambilan keputusan fiskal yang lebih objektif dan berbasis data.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan, mengevaluasi, dan membandingkan kinerja metode ARIMA, LSTM, dan XGBoost dalam memprediksi penerimaan pajak hotel, restoran, dan parkir di Kabupaten Bantul guna memperoleh model peramalan yang paling akurat sebagai dasar pengambilan keputusan fiskal berbasis data.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan desain studi kasus tunggal (*single-case study*) untuk membandingkan kinerja beberapa model peramalan dalam memprediksi penerimaan pajak hotel, restoran, dan parkir di Kabupaten Bantul. Penelitian mengikuti kerangka *Cross-Industry Standard*

Process for Data Mining (CRISP-DM) yang meliputi tahapan *business understanding*, *data understanding*, *data preparation*, *modelling*, dan *evaluation*.

Deskripsi Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa realisasi penerimaan pajak harian sektor hotel, restoran, dan parkir yang diperoleh dari Badan Pengelolaan Keuangan, Pendapatan, dan Aset Daerah (BPKPAD) Kabupaten Bantul periode Januari 2020 hingga Agustus 2025. Dataset terdiri atas sekitar 4,8 juta transaksi yang bersumber dari sistem *tapping box*, mencakup informasi waktu transaksi, nilai transaksi bruto, nilai pajak, identitas wajib pajak, dan jenis operator. Setelah dilakukan proses agregasi harian dan penghapusan data duplikat, diperoleh tiga seri data penerimaan pajak yang digunakan dalam penelitian ini sebagaimana ditunjukkan pada Tabel 1. Sektor parkir memiliki jumlah observasi paling sedikit dibandingkan sektor lainnya karena implementasi *tapping box* pada sektor tersebut baru dimulai pada tahun 2024.

Tabel 1.
Karakteristik Data Penelitian

Sektor	Jumlah Transaksi	Tanggal Awal	Tanggal Akhir	Jumlah Hari
Hotel	7,791	20 Mar 2020	21 Aug 2025	1,837
Restoran	4,743,868	30 Jan 2020	22 Aug 2025	2,026
Parkir	50,329	1 Jan 2024	18 Aug 2025	574

Sumber: BPKPAD Kabupaten Bantul (2025), diolah.

Persiapan Data

Tahap persiapan data dilakukan untuk memastikan kualitas data sebelum digunakan dalam proses pemodelan. Tahapan ini diawali dengan analisis eksploratif untuk mengidentifikasi karakteristik data, meliputi distribusi data, rentang nilai, kelengkapan pencatatan, dan keberadaan nilai pencilan (*outlier*). Selanjutnya dilakukan proses pembersihan data yang mencakup penghapusan data duplikat, koreksi anomali nilai transaksi, serta penyesuaian indeks waktu agar membentuk deret waktu yang berkelanjutan. Data yang hilang akibat gangguan pencatatan sistem *tapping box* ditangani menggunakan metode interpolasi, sedangkan nilai pencilan ditangani menggunakan metode *Hampel Filter*.

Setelah proses pembersihan selesai, data harian sektor hotel dan restoran diagregasi menjadi data bulanan, sedangkan data sektor parkir diagregasi menjadi data dua mingguan karena memiliki periode observasi yang lebih pendek. Pengujian stasioneritas dilakukan menggunakan uji *Augmented Dickey-Fuller* (ADF) dan *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS) untuk menentukan kebutuhan diferensiasi pada model ARIMA dan SARIMA. Selanjutnya dilakukan transformasi data menggunakan *log1p* pada model ARIMA, SARIMA, dan XGBoost, sedangkan normalisasi *Min-Max Scaling* diterapkan pada model LSTM. Selain itu, fitur temporal seperti *lag*, *rolling statistics*, dan komponen kalender digunakan sebagai variabel masukan pada model XGBoost.

Tabel 2.
Granularitas Data dan Horizon Peramalan

Sektor	Granularitas Data	Horizon Peramalan
Hotel	Bulanan	3 bulan, 6 bulan, 12 bulan
Restoran	Bulanan	3 bulan, 6 bulan, 12 bulan
Parkir	Dua Mingguan	1 bulan, 3 bulan, 6 bulan

Sumber: Hasil pengolahan data penelitian (2025).

Model dan Pengaturan Eksperimen

Penelitian ini membandingkan beberapa model peramalan yang terdiri atas model statistik, *machine learning*, dan *deep learning*. Model statistik yang digunakan adalah *Autoregressive Integrated Moving Average* (ARIMA) dan *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Model *deep learning* direpresentasikan oleh *Long Short-Term Memory* (LSTM), sedangkan model *machine*

learning menggunakan *Extreme Gradient Boosting* (XGBoost). Selain itu, model *Seasonal Naïve* digunakan sebagai model dasar (*baseline*) untuk membandingkan kinerja seluruh model yang diuji.

Setiap model diterapkan secara terpisah pada data penerimaan pajak hotel, restoran, dan parkir. Penentuan parameter ARIMA dan SARIMA dilakukan secara otomatis menggunakan pendekatan *auto_arima*. Pada model LSTM digunakan beberapa variasi kedalaman lapisan jaringan, sedangkan pada model XGBoost digunakan variasi kedalaman pohon keputusan untuk memperoleh konfigurasi model yang optimal.

Proses evaluasi model dilakukan menggunakan metode *rolling-origin backtest* (*walk-forward validation*) yang mempertahankan urutan kronologis data. Pada setiap iterasi, data historis digunakan sebagai data pelatihan, sedangkan periode berikutnya digunakan sebagai data pengujian. Pendekatan ini dipilih untuk menghindari terjadinya *data leakage* dan menghasilkan evaluasi yang lebih realistis terhadap kemampuan model dalam melakukan peramalan.

Kinerja model dievaluasi menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). Nilai RMSE digunakan untuk mengukur besarnya kesalahan prediksi dalam satuan rupiah, sedangkan MAPE digunakan untuk mengukur tingkat kesalahan relatif sehingga memungkinkan perbandingan antar sektor dan horizon peramalan. Model dengan nilai kesalahan terendah ditetapkan sebagai model terbaik pada masing-masing sektor pajak dan horizon peramalan.

Tabel 3.
Model Peramalan yang Digunakan

Kelompok Model	Model
Baseline	Seasonal Naïve
Statistik	ARIMA
Statistik	SARIMA
Deep Learning	LSTM
Machine Learning	XGBoost

Sumber: Hasil pengolahan peneliti (2025).

HASIL DAN PEMBAHASAN

Hasil

Bagian ini menyajikan hasil pengujian model peramalan penerimaan pajak hotel, restoran, dan parkir di Kabupaten Bantul. Evaluasi dilakukan menggunakan nilai *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) untuk menentukan model terbaik pada setiap sektor dan horizon peramalan.

Tabel 4.
Model Terbaik Berdasarkan Sektor dan Horizon Peramalan

Sektor	Horizon Peramalan	Model Terbaik	Total RMSE (Rp)	Total MAPE (%)
Hotel	3 Bulan	<i>Seasonal Naïve</i>	252.000.439	13,66
Hotel	6 Bulan	<i>Seasonal Naïve</i>	396.171.650	9,85
Hotel	1 Tahun	<i>Seasonal Naïve</i>	764.833.167	10,94
Restoran	3 Bulan	<i>Seasonal Naïve</i>	209.298.321	3,78
Restoran	6 Bulan	XGBoost (<i>depth</i> 5)	295.132.483	2,76
Restoran	1 Tahun	XGBoost (<i>depth</i> 2)	379.728.290	1,86
Parkir	1 Bulan	ARIMA(1,1,1)	112.623	9,09
Parkir	3 Bulan	ARIMA(1,1,1)	340.441	7,84
Parkir	6 Bulan	ARIMA(1,1,1)	770.077	7,97

Sumber: Hasil pengolahan data penelitian (2025).

Berdasarkan Tabel 4, terdapat perbedaan model terbaik pada masing-masing sektor pajak dan horizon peramalan. Pada sektor hotel, model *Seasonal Naïve* secara konsisten menghasilkan nilai RMSE terendah pada seluruh horizon peramalan. Nilai MAPE yang diperoleh berkisar antara 9,85% hingga 13,66%, menunjukkan bahwa pola penerimaan pajak hotel masih didominasi oleh pola musiman yang

berulang sehingga pendekatan berbasis pola historis sederhana mampu menghasilkan prediksi yang relatif akurat.

Pada sektor restoran, model terbaik berbeda menurut horizon peramalan yang digunakan. Untuk horizon tiga bulan, model *Seasonal Naïve* menghasilkan tingkat kesalahan terendah dengan nilai MAPE sebesar 3,78%. Namun, pada horizon enam bulan dan satu tahun, model XGBoost menunjukkan kinerja yang lebih baik dengan nilai MAPE masing-masing sebesar 2,76% dan 1,86%. Hasil ini menunjukkan bahwa penerimaan pajak restoran memiliki pola yang relatif stabil sehingga dapat dimanfaatkan secara optimal oleh algoritma *machine learning* dalam menghasilkan prediksi jangka menengah dan jangka panjang.

Sementara itu, pada sektor parkir, model ARIMA(1,1,1) secara konsisten menjadi model terbaik pada seluruh horizon peramalan. Nilai MAPE yang dihasilkan berada pada rentang 7,84% hingga 9,09%, yang mengindikasikan bahwa pola penerimaan pajak parkir masih dapat dijelaskan dengan baik melalui hubungan historis antarperiode. Secara keseluruhan, temuan ini menunjukkan bahwa tidak terdapat satu model yang unggul untuk seluruh sektor pajak, melainkan efektivitas model sangat dipengaruhi oleh karakteristik data dan horizon peramalan yang digunakan.

Tabel 5.
Perbandingan Nilai MAPE (%) Antar Model Peramalan pada Setiap Sektor dan Horizon Peramalan

Sektor dan Horizon Peramalan	Baseline (<i>Seasonal Naïve</i>)	ARIMA	SARIMA	LSTM	XGBoost
Hotel, 3 Bulan	13,66	17,00	12,67	23,40	14,64
Hotel, 6 Bulan	9,85	21,19	19,14	23,79	18,73
Hotel, 1 Tahun	10,94	21,15	28,96	25,50	28,56
Restoran, 3 Bulan	3,78	10,85	8,92	7,52	5,29
Restoran, 6 Bulan	4,82	10,04	16,59	7,46	2,76
Restoran, 1 Tahun	5,88	9,09	30,09	7,77	1,86
Parkir, 1 Bulan	58,76	9,09	9,78	12,68	10,25
Parkir, 3 Bulan	57,91	7,84	11,85	17,92	12,41
Parkir, 6 Bulan	45,34	7,97	15,81	31,58	18,84

Sumber: Hasil pengolahan data penelitian (2025).

Berdasarkan Tabel 5, terlihat bahwa tingkat akurasi model peramalan berbeda pada setiap sektor dan horizon peramalan. Pada sektor hotel, model *Seasonal Naïve* menghasilkan nilai MAPE terendah pada horizon enam bulan dan satu tahun, sedangkan pada horizon tiga bulan model SARIMA menunjukkan kinerja yang sedikit lebih baik dengan nilai MAPE sebesar 12,67%. Temuan ini menunjukkan bahwa pola penerimaan pajak hotel cenderung dipengaruhi oleh faktor musiman yang kuat sehingga model berbasis musiman mampu memberikan hasil yang lebih baik dibandingkan model lainnya.

Pada sektor restoran, model XGBoost menghasilkan tingkat akurasi tertinggi pada horizon enam bulan dan satu tahun dengan nilai MAPE masing-masing sebesar 2,76% dan 1,86%. Sementara itu, pada horizon tiga bulan model *Seasonal Naïve* masih menunjukkan performa terbaik dengan nilai MAPE sebesar 3,78%. Hasil ini mengindikasikan bahwa pola penerimaan pajak restoran relatif stabil dan dapat dimodelkan secara efektif menggunakan pendekatan *machine learning*.

Pada sektor parkir, model ARIMA secara konsisten menghasilkan nilai MAPE terendah pada seluruh horizon peramalan. Nilai MAPE yang berada pada rentang 7,84% hingga 9,09% menunjukkan bahwa model ARIMA mampu menangkap pola historis penerimaan pajak parkir dengan baik. Sebaliknya, model *Seasonal Naïve* menghasilkan tingkat kesalahan yang jauh lebih tinggi pada sektor ini, dengan nilai MAPE mencapai lebih dari 45%. Temuan tersebut memperkuat hasil pada Tabel 4 bahwa karakteristik data menjadi faktor utama dalam menentukan model peramalan yang paling sesuai.

Hasil Peramalan Penerimaan Pajak Daerah

Setelah model terbaik ditentukan pada setiap sektor dan horizon peramalan, model tersebut dilatih kembali menggunakan seluruh data yang tersedia untuk menghasilkan prediksi penerimaan pajak pada

periode mendatang. Hasil peramalan menunjukkan bahwa nilai prediksi penerimaan berbeda pada masing-masing sektor sesuai dengan karakteristik dan volume transaksi yang dimiliki.

Pada sektor hotel, total penerimaan pajak yang diprediksi mencapai sekitar Rp1,8 miliar untuk horizon tiga bulan dan meningkat menjadi sekitar Rp6,6 miliar pada horizon satu tahun. Sementara itu, sektor restoran menunjukkan potensi penerimaan yang lebih besar dibandingkan sektor lainnya. Hasil peramalan satu tahun menunjukkan bahwa penerimaan pajak restoran diperkirakan mencapai sekitar Rp18,1 miliar. Tingginya nilai tersebut mencerminkan besarnya aktivitas transaksi pada sektor restoran serta pola penerimaan yang relatif stabil selama periode pengamatan.

Berbeda dengan sektor hotel dan restoran, sektor parkir menghasilkan nilai prediksi yang relatif lebih kecil, yaitu berkisar antara Rp1,2 juta hingga Rp7,3 juta. Kondisi ini dipengaruhi oleh jumlah transaksi yang lebih rendah serta periode implementasi sistem *tapping box* yang masih lebih pendek dibandingkan dua sektor lainnya. Meskipun demikian, hasil peramalan tetap memberikan gambaran yang berguna mengenai potensi penerimaan pajak parkir pada periode mendatang.

Secara keseluruhan, hasil peramalan ini dapat menjadi salah satu dasar dalam penyusunan target penerimaan daerah yang lebih objektif dan berbasis data. Informasi mengenai potensi penerimaan pada berbagai horizon waktu juga dapat membantu pemerintah daerah dalam melakukan perencanaan anggaran dan pengambilan keputusan fiskal secara lebih terukur.

Pembahasan

Hasil penelitian menunjukkan bahwa tidak terdapat satu model peramalan yang secara konsisten menghasilkan tingkat akurasi terbaik pada seluruh sektor pajak dan horizon peramalan. Temuan ini mengindikasikan bahwa pemilihan model peramalan sangat dipengaruhi oleh karakteristik data yang dianalisis, termasuk tingkat volatilitas, pola musiman, panjang observasi, dan stabilitas data historis. Hasil penelitian ini sejalan dengan Zhao et al. (2021) serta Kim dan Schwartz (2013) yang menyatakan bahwa tidak ada model peramalan yang bersifat universal untuk seluruh jenis data dan konteks aplikasi. Oleh karena itu, pemilihan model harus disesuaikan dengan karakteristik data dan tujuan peramalan yang ingin dicapai.

Pada sektor restoran, model XGBoost menghasilkan tingkat akurasi tertinggi terutama pada horizon enam bulan dan satu tahun. Hasil ini menunjukkan bahwa penerimaan pajak restoran memiliki pola yang relatif stabil sehingga dapat dimodelkan dengan baik menggunakan pendekatan *machine learning*. Kemampuan XGBoost dalam memanfaatkan hubungan nonlinier, pola historis, serta fitur temporal seperti *lag*, *rolling statistics*, dan komponen kalender memungkinkan model ini menghasilkan prediksi yang lebih akurat dibandingkan model statistik tradisional. Temuan ini mendukung penelitian Yu (2024) dan Wang dan Su (2025) yang menunjukkan bahwa algoritma berbasis *machine learning* memiliki kemampuan yang baik dalam menangkap pola kompleks pada data deret waktu dan menghasilkan tingkat akurasi yang tinggi untuk peramalan jangka menengah maupun jangka panjang.

Berbeda dengan sektor restoran, model ARIMA memberikan hasil terbaik pada sektor parkir untuk seluruh horizon peramalan. Nilai MAPE yang relatif rendah menunjukkan bahwa pola penerimaan pajak parkir masih didominasi oleh hubungan historis yang sederhana dan tidak memiliki pola nonlinier yang terlalu kompleks. Dalam kondisi seperti ini, pendekatan statistik konvensional masih mampu memberikan hasil yang optimal. Temuan ini sejalan dengan penelitian Zhu et al. (2018) dan Kumbasaroglu (2026) yang menunjukkan bahwa model ARIMA tetap relevan dan efektif untuk memodelkan data deret waktu yang memiliki pola historis yang stabil dan dapat diprediksi.

Sementara itu, sektor hotel menunjukkan karakteristik yang berbeda dibandingkan dua sektor lainnya. Pada seluruh horizon peramalan, model *Seasonal Naïve* menghasilkan tingkat kesalahan paling rendah dibandingkan model lainnya. Kondisi ini menunjukkan bahwa penerimaan pajak hotel memiliki pola musiman yang kuat dan berulang dari tahun ke tahun. Tingginya ketergantungan Kabupaten Bantul terhadap sektor pariwisata menyebabkan penerimaan pajak hotel sangat dipengaruhi oleh musim liburan, kondisi cuaca, dan jumlah kunjungan wisatawan. Oleh karena itu, pendekatan sederhana yang memanfaatkan pola musiman periode sebelumnya justru mampu menghasilkan prediksi yang lebih baik

dibandingkan model yang lebih kompleks. Temuan ini memperkuat pendapat Williams dan Kavanagh (2016)) bahwa model yang paling kompleks tidak selalu menjadi model yang paling akurat apabila karakteristik data telah didominasi oleh pola musiman yang jelas.

Hasil yang menunjukkan bahwa model LSTM tidak menjadi model terbaik pada seluruh sektor dapat dijelaskan dari karakteristik data yang tersedia. Model deep learning seperti LSTM umumnya membutuhkan volume data pelatihan yang besar untuk dapat mengestimasi parameter jaringan secara stabil, sehingga pada seri data bulanan maupun dua mingguan yang relatif pendek model ini cenderung mengalami overfitting dan menghasilkan kesalahan yang lebih tinggi. Sebaliknya, keunggulan XGBoost pada sektor restoran menunjukkan bahwa metode machine learning berbasis pohon lebih efisien dalam memanfaatkan fitur temporal pada data dengan jumlah observasi terbatas. Fenomena ini konsisten dengan prinsip bias-variance trade-off, yakni model yang lebih kompleks tidak selalu menghasilkan akurasi yang lebih baik apabila jumlah data tidak memadai atau pola data telah didominasi oleh struktur musiman yang jelas. Dengan demikian, keunggulan atau kelemahan suatu model tidak dapat dilepaskan dari kesesuaian antara kompleksitas model, karakteristik data, dan panjang periode observasi.

Dari perspektif praktis, hasil penelitian ini memberikan implikasi penting bagi BPKPAD Kabupaten Bantul dalam proses penyusunan target penerimaan daerah. Selama ini penetapan target penerimaan sering kali masih dipengaruhi oleh pengalaman praktis dan pertimbangan subjektif. Hasil penelitian menunjukkan bahwa pendekatan berbasis data dapat memberikan estimasi yang lebih objektif dan terukur. Dengan memanfaatkan model terbaik pada masing-masing sektor pajak, pemerintah daerah dapat meningkatkan kualitas perencanaan fiskal sekaligus mengurangi risiko terjadinya selisih yang besar antara target dan realisasi penerimaan.

Selain itu, sektor restoran yang memiliki tingkat akurasi prediksi tertinggi dapat dijadikan sebagai salah satu dasar utama dalam penyusunan proyeksi pendapatan daerah. Sebaliknya, sektor hotel yang memiliki tingkat volatilitas tinggi memerlukan pendekatan yang lebih fleksibel, misalnya melalui penyusunan target dalam bentuk rentang skenario optimistis dan pesimistis. Di samping itu, peningkatan kualitas data melalui perluasan penggunaan *tapping box*, pengurangan perangkat yang tidak aktif, serta rekonsiliasi data secara berkala perlu terus dilakukan agar akurasi peramalan dapat semakin meningkat pada masa mendatang. Dengan demikian, pemanfaatan teknologi digital tidak hanya berfungsi sebagai sarana pencatatan transaksi, tetapi juga sebagai dasar pengambilan keputusan fiskal yang lebih efektif dan berbasis data.

Kontribusi Penelitian

Penelitian ini memberikan tiga kontribusi utama. Secara teoretis, penelitian ini memperkaya kajian peramalan penerimaan pajak daerah dengan menyediakan bukti empiris perbandingan kinerja metode statistik, machine learning, dan deep learning pada data transaksi digital yang nyata. Secara metodologis, penelitian ini menunjukkan pentingnya evaluasi lintas horizon dan lintas tingkat agregasi data dalam menentukan model peramalan yang paling sesuai untuk setiap karakteristik sektor. Secara praktis, penelitian ini menghasilkan rekomendasi pemilihan model spesifik per sektor yang dapat langsung dimanfaatkan oleh BPKPAD Kabupaten Bantul dalam menyusun target penerimaan daerah yang lebih objektif dan terukur.

Keterbatasan Penelitian dan Arah Penelitian Selanjutnya

Penelitian ini memiliki sejumlah keterbatasan yang perlu diperhatikan. Pertama, periode observasi pada sektor parkir masih relatif pendek karena implementasi *tapping box* baru dimulai pada tahun 2024, sehingga hasil peramalan pada sektor tersebut belum sepenuhnya mencerminkan pola musiman jangka panjang. Kedua, penelitian ini hanya menggunakan variabel internal berupa data historis transaksi tanpa memasukkan variabel eksternal seperti jumlah kunjungan wisatawan, hari libur nasional, kondisi cuaca, maupun indikator ekonomi makro yang berpotensi memengaruhi penerimaan pajak. Ketiga, cakupan penelitian terbatas pada tiga sektor pajak di satu kabupaten sehingga generalisasi hasil ke daerah lain perlu dilakukan secara hati-hati. Berdasarkan keterbatasan tersebut, penelitian selanjutnya disarankan

untuk mengintegrasikan variabel eksternal, memperluas cakupan sektor dan wilayah, serta mengeksplorasi pendekatan model hibrida yang menggabungkan keunggulan metode statistik dan machine learning guna meningkatkan akurasi dan generalisasi model peramalan.

SIMPULAN DAN SARAN

Simpulan

Berdasarkan hasil penelitian dan pembahasan, dapat disimpulkan bahwa:

1. Tidak terdapat satu model peramalan yang secara konsisten memberikan tingkat akurasi terbaik pada seluruh sektor pajak dan horizon peramalan. Kinerja model sangat dipengaruhi oleh karakteristik data dan pola penerimaan pada masing-masing sektor.
2. Pada sektor restoran, model XGBoost menghasilkan tingkat akurasi tertinggi terutama pada horizon enam bulan dan satu tahun dengan nilai MAPE masing-masing sebesar 2,76% dan 1,86%. Hasil ini menunjukkan bahwa pendekatan machine learning mampu memanfaatkan pola historis dan karakteristik data yang relatif stabil untuk menghasilkan prediksi yang lebih akurat.
3. Pada sektor parkir, model ARIMA(1,1,1) secara konsisten menjadi model terbaik pada seluruh horizon peramalan dengan nilai MAPE berkisar antara 7,84% hingga 9,09%. Temuan ini menunjukkan bahwa pola penerimaan pajak parkir masih dapat dimodelkan secara efektif menggunakan pendekatan statistik deret waktu konvensional.
4. Pada sektor hotel, model Seasonal Naïve menghasilkan tingkat kesalahan paling rendah dibandingkan model lainnya. Hal ini menunjukkan bahwa penerimaan pajak hotel memiliki pola musiman yang kuat sehingga pendekatan berbasis pengulangan pola musiman mampu memberikan hasil prediksi yang lebih baik.
5. Hasil penelitian membuktikan bahwa pemanfaatan data transaksi yang berasal dari sistem tapping box dapat digunakan sebagai dasar dalam pengembangan model peramalan penerimaan pajak daerah yang lebih objektif, terukur, dan berbasis data untuk mendukung proses perencanaan fiskal daerah.

Saran

1. BPKPAD Kabupaten Bantul disarankan untuk memanfaatkan model peramalan yang berbeda sesuai dengan karakteristik masing-masing sektor pajak dalam proses penyusunan target penerimaan daerah.
2. Pada sektor hotel yang memiliki tingkat volatilitas tinggi, penyusunan target penerimaan sebaiknya menggunakan beberapa skenario peramalan, seperti skenario optimistis, moderat, dan pesimistis, guna mengurangi risiko ketidakpastian.
3. Pemerintah daerah perlu meningkatkan kualitas dan cakupan implementasi sistem tapping box melalui perluasan perangkat, pemeliharaan sistem secara berkala, dan rekonsiliasi data yang berkesinambungan agar akurasi peramalan dapat terus ditingkatkan.
4. Penelitian selanjutnya dapat mengembangkan model peramalan dengan memanfaatkan variabel eksternal seperti jumlah wisatawan, kondisi cuaca, hari libur nasional, dan indikator ekonomi daerah untuk meningkatkan akurasi prediksi penerimaan pajak daerah.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada Badan Pengelolaan Keuangan, Pendapatan, dan Aset Daerah (BPKPAD) Kabupaten Bantul atas izin penggunaan data *tapping box* serta dukungan yang diberikan selama proses pengumpulan data penelitian. Penulis juga mengucapkan terima kasih kepada Fakultas Ilmu Komputer Universitas Indonesia atas dukungan akademik yang diberikan selama pelaksanaan penelitian ini.

REFERENCES

- Abdullah Sani, H. A., & Jaafar, N. I. (2026). Exploring the impact of IoT on governance and public service transformation: evidence from Malaysia's public sector. *Smart and Sustainable Built Environment*, 15(5), 1991–2014. <https://doi.org/10.1108/SASBE-10-2024-0453>
- Alsharef, A., Aggarwal, K., Sonia, Kumar, M., & Mishra, A. (2022). Review of ML and AutoML Solutions to Forecast Time-Series Data. *Archives of Computational Methods in Engineering*, 29(7), 5297–5311. <https://doi.org/10.1007/s11831-022-09765-0>
- Chen, L., & He, R. (2024). Does digital tax enforcement drive corporate digitalization? Evidence from the Golden Tax Project III in China: A pre-registered report. *Pacific-Basin Finance Journal*, 83, 102242. <https://doi.org/10.1016/j.pacfin.2023.102242>
- David, A., Yigitcanlar, T., Li, R. Y. M., Corchado, J. M., Cheong, P. H., Mossberger, K., & Mehmood, R. (2023). Understanding Local Government Digital Technology Adoption Strategies: A PRISMA Review. *Sustainability*, 15(12), 9645. <https://doi.org/10.3390/su15129645>
- He, Y., & Yi, Y. (2023). Digitalization of tax administration and corporate performance: Evidence from China. *International Review of Financial Analysis*, 90, 102859. <https://doi.org/10.1016/j.irfa.2023.102859>
- Jothi, L. S. (2025). *Time Series Analysis Using LSTM Networks and Its Application to Financial Forecasting* (pp. 19–58). <https://doi.org/10.4018/979-8-3373-1200-2.ch002>
- Kim, N., & Schwartz, Z. (2013). The accuracy of tourism forecasting and data characteristics: a meta-analytical approach. *Journal of Hospitality Marketing & Management*, 22(4), 349–374.
- Kumbasaroglu, H. (2026). Global Tea Production Forecasting Using ARIMA Models: A Multi-Country Time-Series Analysis (1961–2028). *Sustainability*, 18(2), 1005. <https://doi.org/10.3390/su18021005>
- Lee, S. (2025). Digitalization and Domestic Revenue Mobilization: Governmental Responsiveness and Capabilities Matter. *Governance*, 38(4). <https://doi.org/10.1111/gove.70076>
- Li, W., & Li, D. (2024). How does technological innovation affect the fiscal decentralization of local governments? Evidence from China. *Heliyon*, 10(9), e30132. <https://doi.org/10.1016/j.heliyon.2024.e30132>
- Li, X., Li, K., Shen, S., & Tian, Y. (2023). Exploring Time Series Models for Wind Speed Forecasting: A Comparative Analysis. *Energies*, 16(23), 7785. <https://doi.org/10.3390/en16237785>
- Li, Y., Wu, K., & Liu, J. (2023). Self-paced ARIMA for robust time series prediction. *Knowledge-Based Systems*, 269, 110489. <https://doi.org/10.1016/j.knosys.2023.110489>
- Lubis, M., Saputra, M., & Agustika Nurtrisha, W. (2021). Financial Technology Development Framework for Prosperity of the Nation and Potential Direction. *Proceedings of the 9th International Conference on Computer and Communications Management*, 212–218. <https://doi.org/10.1145/3479162.3479194>
- Mahmud, Windu Gata, Hafifah Bella Novitasari, Sigit Kurniawan, & Dedi Dwi Saputra. (2024). Indonesian Government Revenue Prediction Using Long Short-Term Memory. *Inspiration: Jurnal Teknologi Informasi Dan Komunikasi*, 14(1), 111–124. <https://doi.org/10.35585/inspir.v14i1.67>
- Ongo Nkoa, B. E., & Song, J. S. (2022). Les canaux de transmission des effets des TIC sur la mobilisation des recettes fiscales en Afrique. *African Development Review*, 34(S1). <https://doi.org/10.1111/1467-8268.12650>
- Rich, R., Bram, J., Haughwout, A., Orr, J., Rosen, R., & Sela, R. (2005). Using Regional Economic Indexes to Forecast Tax Bases: Evidence from New York. *Review of Economics and Statistics*, 87(4), 627–634. <https://doi.org/10.1162/003465305775098215>
- Scherer, M. (2025). Forecasting Public Expenditure Using LSTMs: A Multi-Task Learning Approach for Volatile Data. *2025 30th International Conference on Automation and Computing (ICAC)*, 1–6. <https://doi.org/10.1109/ICAC65379.2025.11196732>
- Song, X., Deng, L., Wang, H., Zhang, Y., He, Y., & Cao, W. (2024). Deep learning-based time series forecasting. *Artificial Intelligence Review*, 58(1), 23. <https://doi.org/10.1007/s10462-024-10989-8>

- Thayyib, P. V., Thorakkattle, M. N., Usmani, F., Yahya, A. T., & Farhan, N. H. S. (2023). Forecasting Indian Goods and Services Tax revenue using TBATS, ETS, Neural Networks, and hybrid time series models. *Cogent Economics & Finance*, 11(2). <https://doi.org/10.1080/23322039.2023.2285649>
- Valle-Cruz, D., & García-Contreras, R. (2025). Towards AI-driven transformation and smart data management: Emerging technological change in the public sector value chain. *Public Policy and Administration*, 40(2), 254–275. <https://doi.org/10.1177/09520767231188401>
- Wang, X., & Su, C. (2025). Deep learning for cancer detection based on genomic and imaging data: A comprehensive review. *Cancer Management and Research*, 2089–2125.
- Williams, D. W., & Kavanagh, S. C. (2016). Local government revenue forecasting methods: Competition and comparison. *Journal of Public Budgeting, Accounting & Financial Management*, 28(4), 488–526.
- Yu, X. (2024). Construction of Local Fiscal Revenue Forecasting Model Based on SVR Model. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-0784>
- Yukhno, A. (2024). Digital Transformation: Exploring big data Governance in Public Administration. *Public Organization Review*, 24(1), 335–349. <https://doi.org/10.1007/s11115-022-00694-x>
- Zhao, Y., Wang, Y., Liu, J., Xia, H., Xu, Z., Hong, Q., Zhou, Z., & Petzold, L. (2021). Empirical quantitative analysis of COVID-19 forecasting models. *2021 International Conference on Data Mining Workshops (ICDMW)*, 517–526.
- Zhu, Y., Wang, Y., Liu, T., & Sui, Q. (2018). Assessing macroeconomic recovery after a natural hazard based on ARIMA—a case study of the 2008 Wenchuan earthquake in China. *Natural Hazards*, 91(3), 1025–1038. <https://doi.org/10.1007/s11069-017-3163-1>