

## ANALISIS PENGGUNAAN *MACHINE LEARNING* TERHADAP PREDIKSI JUMLAH NASABAH PADA PRODUK AMANAH DI PT. PEGADAIAN CPS PANGKAJENE

Randy Arya Renaldi<sup>1)</sup>, Muhammad Rais Rahmat Razak<sup>2)</sup>, Haslindah<sup>3)</sup>, Rahman Yakub<sup>4)</sup>

<sup>1,2,3,4)</sup> Fakultas Ekonomi dan Bisnis, Universitas Muhammadiyah Sidenreng Rappang, Indonesia

<sup>1)</sup> [zeronineone8118@email.com](mailto:zeronineone8118@email.com), <sup>2)</sup> [mraisrahmat@gmail.com](mailto:mraisrahmat@gmail.com), <sup>3)</sup> [haslindahpratiwi7@gmail.com](mailto:haslindahpratiwi7@gmail.com),

<sup>4)</sup> [rahman12@gmail.com](mailto:rahman12@gmail.com)

---

### ARTICLE HISTORY

Received:  
November 17, 2025  
Revised  
December 9, 2025  
Accepted:  
December 11, 2025  
Online available:6  
Januari 15, 2026

### Keywords:

Amanah Product, Customer Prediction, Machine learning, Marketing Strategy, PT. Pegadaian, SARIMA

\*Correspondence:

Name: Randy Arya Renaldi  
E-mail:  
[zeronineone8118@email.com](mailto:zeronineone8118@email.com)

### Editorial Office

Ambon State Polytechnic  
Center for Research and  
Community Service  
Ir. M. Putuhena Street, Wailela-  
Rumahtiga, Ambon  
Maluku, Indonesia  
Postal Code: 97234

### ABSTRACT

**Introduction:** This study aims to analyze the use of machine learning in predicting the number of customers for Amanah products at PT. Pegadaian CPS Pangkajene. The issues examined include the application of machine learning, the selection of the most effective algorithm, and factors affecting prediction accuracy.

**Methods:** The study uses a quantitative approach, drawing on historical customer data over five years. The three machine learning models compared are SARIMA, Random Forest, and Linear Regression.

**Results:** The SARIMA model achieved the best performance, with a Mean Absolute Error (MAE) of 2.61%, a Root Mean Square Error (RMSE) of 2.94%, and a Mean Absolute Percentage Error (MAPE) of 8.58%. This model is the most accurate in predicting trends and seasonal patterns in customer numbers. The Linear Regression model showed less accurate results, while Random Forest excelled in reliability but was still less accurate than SARIMA.

**Conclusion and suggestion:** The application of machine learning, especially SARIMA, can be a strategic tool for business planning and decision-making in the Islamic finance sector, providing data-based recommendations to improve the performance of Amanah products.

### PENDAHULUAN

PT. Pegadaian (Persero) merupakan Badan Usaha Milik Negara yang fokus memberikan pinjaman melalui sistem pegadaian. Proses pinjaman dirancang sederhana, cepat, dan aman agar tidak membebani peminjam atau menimbulkan masalah tambahan, sejalan dengan motto pegadaian yaitu "Menyelesaikan masalah tanpa masalah"

(Kusuma, 2024). Pegadaian adalah organisasi atau lembaga yang memberikan jasa peminjaman uang dengan menggunakan barang sebagai jaminan. Pada umumnya, usaha pegadaian menjaminkan barang-barang berharga kepada pihak tertentu dengan sejumlah uang tertentu, dengan syarat barang yang digadaikan ditebus menurut kesepakatan antara nasabah dengan pegadaian (Kusuma, 2024).

Revолюси Industri 4.0 dan konsep *Society 5.0* menggambarkan perkembangan teknologi yang mendorong semua sektor untuk beradaptasi. Transformasi ini menunjukkan pergeseran dari teknologi konvensional ke teknologi modern yang menggunakan internet dan sistem digital, yang menjadi faktor kunci dalam meningkatkan daya saing produk di berbagai sektor, termasuk industri jasa (Mursalat et al., 2025). Kompetensi kewirausahaan mencakup manajemen bisnis yang efektif, identifikasi peluang yang mahir, dan manajemen risiko yang terampil. Secara bersamaan, inovasi produk muncul sebagai kunci untuk menciptakan nilai tambah, meningkatkan daya saing produk, dan merespons kebutuhan pasar yang terus berkembang (Herman et al., 2024).

Prediksi jumlah nasabah menjadi komponen vital dalam perencanaan strategis PT. Pegadaian. Dengan mengetahui estimasi jumlah nasabah di masa depan, perusahaan dapat merancang strategi pemasaran yang efektif, mengalokasikan sumber daya secara optimal, dan meningkatkan kualitas layanan. Penelitian menunjukkan bahwa peningkatan jumlah nasabah dapat dicapai melalui strategi pemasaran yang efektif dan pelayanan yang baik. Seiring dengan kemajuan teknologi, *machine learning* telah menjadi alat yang efektif dalam analisis data dan prediksi di bidang keuangan. Teknologi ini mampu mengolah data dalam jumlah besar dan kompleks, mengidentifikasi pola tersembunyi, serta menghasilkan prediksi dengan tingkat akurasi yang tinggi. Dalam industri keuangan, *machine learning* digunakan untuk berbagai keperluan, seperti analitik prediktif, manajemen risiko, dan optimalisasi portofolio.

Produk Amanah sebagai salah satu layanan berbasis syariah dari PT. Pegadaian memerlukan pengelolaan yang baik agar dapat terus memberikan manfaat maksimal kepada masyarakat sekaligus mendukung keberlanjutan bisnis perusahaan. Pengelolaan yang efektif membutuhkan data prediktif yang akurat untuk mengantisipasi kebutuhan nasabah dan merespons dinamika pasar. Dengan memanfaatkan teknologi *machine learning*, PT. Pegadaian dapat meningkatkan efisiensi operasional dan memberikan solusi keuangan yang lebih relevan bagi masyarakat (Komala & Padilah, 2023). Pada bulan Maret hingga April 2024, dilakukan observasi di PT. Pegadaian CPS Pangkajene untuk mengidentifikasi kendala yang dihadapi terkait produk Amanah. Hasil observasi menunjukkan bahwa jumlah nasabah pada produk ini cenderung rendah dibandingkan dengan target yang telah ditetapkan. Kondisi ini menimbulkan kekhawatiran terkait efektivitas strategi pemasaran dan segmentasi pasar yang digunakan. Oleh karena itu, penelitian ini berfokus pada analisis penggunaan *machine learning* dalam memprediksi jumlah nasabah pada produk Amanah untuk memberikan rekomendasi yang lebih terarah dan berbasis data, sehingga dapat meningkatkan daya tarik produk ini di pasar yang kompetitif. Dengan demikian, penelitian ini penting dilakukan untuk menganalisis bagaimana *machine learning* dapat digunakan dalam memprediksi jumlah nasabah pada produk Amanah di PT. Pegadaian CPS Pangkajene, Kabupaten Sidenreng Rappang. Hasil penelitian diharapkan dapat menjadi dasar bagi pengembangan strategi yang lebih berbasis data untuk meningkatkan daya saing perusahaan.

Beberapa penelitian terdahulu telah membuktikan efektivitas *machine learning* dalam prediksi nasabah di sektor keuangan. (Naufal et al., 2023) menunjukkan bahwa algoritma Decision Tree mampu memprediksi potensi hilangnya nasabah bank dengan akurasi 93,3%. Sementara itu, (Fitrianah et al., 2021) membuktikan bahwa metode Naïve Bayes efektif dalam memprediksi nasabah potensial di institusi keuangan. Namun, penelitian khusus mengenai prediksi nasabah produk syariah di lembaga pegadaian masih sangat terbatas. Urgensi penelitian ini terletak pada kebutuhan PT. Pegadaian CPS Pangkajene untuk meningkatkan performa produk Amanah yang mengalami penurunan jumlah nasabah. Tanpa prediksi yang akurat, perusahaan akan kesulitan dalam merancang strategi pemasaran yang tepat sasaran dan mengalokasikan sumber daya secara efisien. Oleh karena itu, penelitian ini menjadi penting untuk mengisi gap penelitian terdahulu sekaligus memberikan solusi praktis berbasis data bagi pengembangan produk syariah di industri pegadaian.

## TINJAUAN PUSTAKA

### **Machine Learning**

Machine learning atau pembelajaran mesin adalah salah satu cabang kecerdasan buatan yang berfokus pada pembelajaran dari data (*learning from data*), yaitu mengembangkan sistem yang dapat belajar sendiri tanpa perlu pemrograman berulang-ulang oleh manusia (Cholissodin et al., 2020). Teori ini diperkuat oleh pendekatan berbasis data (Sidik & Ansawarman, 2022) yang menekankan pentingnya pemanfaatan data historis untuk melatih model sehingga dapat mencapai kinerja optimal dalam mengekstraksi serta mengolah informasi. Penelitian (Sihombing & Arsani, 2021) lebih lanjut menunjukkan bahwa teknologi machine learning mampu belajar secara mandiri tanpa intervensi langsung dari manusia, dengan dukungan dasar keilmuan lain seperti statistik, matematika, serta teknik penambangan data yang menjadi fondasi utama dalam pengembangannya.

Dalam konteks konseptual, penelitian ini mendefinisikan machine learning sebagai penerapan algoritma dan teknik pembelajaran komputer untuk menganalisis data historis guna mengidentifikasi pola, menghasilkan wawasan, dan membuat prediksi yang akurat secara otomatis. Konsep ini mencakup berbagai metode seperti *Random Forest*, *Decision Tree*, *Support Vector Machine*, *Naive Bayes*, *K-Nearest Neighbor*, dan *Artificial Neural Network* yang dapat digunakan untuk klasifikasi dan prediksi (Sihombing & Arsani, 2021). Dalam konteks prediksi nasabah, machine learning menjadi landasan penting dalam memahami hubungan antara data historis dan proyeksi masa depan. Penggunaan machine learning untuk memprediksi jumlah nasabah dan perilaku mereka telah menjadi fokus penting di bidang penelitian. (Amuda & Adeyemo, 2019) menunjukkan bahwa model *Artificial Neural Network* (ANN) mampu memberikan tingkat akurasi yang tinggi dalam memprediksi churn nasabah dalam institusi keuangan. Pendekatan hibrida yang menggabungkan metode statistik klasik dengan algoritma machine learning juga telah menjadi populer untuk memahami dan memprediksi perilaku pelanggan (Alizamir et al., 2022). Selain itu, *Customer Lifetime Value* (CLV) sebagai metrik penting dalam strategi pemasaran menunjukkan bahwa model berbasis machine learning dapat memberikan estimasi yang lebih akurat dibandingkan pendekatan tradisional (Cowan et al., 2023)

Bukti empiris dari berbagai penelitian lintas negara secara konsisten mendukung efektivitas machine learning dalam prediksi. (Khalijah Zainal, 2024) menemukan bahwa *MLP Regressor* menghasilkan akurasi tertinggi mencapai 97% dalam prediksi harga, sementara *Random Forest Regressor* menunjukkan performa terbaik dengan akurasi 98%. Temuan ini diperkuat oleh (Naufal et al., 2023) yang menunjukkan bahwa *Decision Tree* memberikan performa tertinggi dengan nilai akurasi 93,3% dalam memprediksi potensi hilangnya nasabah bank. Sudriyanto et al. (2023) juga mengungkapkan bahwa *K-Nearest Neighbors* memiliki performa terbaik dengan nilai RMSE terendah dalam prediksi harga saham. Mekanisme pengaruh machine learning terhadap akurasi prediksi nasabah dapat dijelaskan melalui beberapa jalur. Pertama, kemampuan mengidentifikasi pola tersembunyi dalam data historis yang sulit terdeteksi secara manual. Kedua, fleksibilitas dalam menangani data kompleks dan non-linear yang tidak dapat diakomodasi oleh metode statistik tradisional. Ketiga, kemampuan adaptasi terhadap perubahan pola data melalui proses pembelajaran berkelanjutan. Mekanisme ini menunjukkan bahwa machine learning tidak hanya meningkatkan akurasi prediksi tetapi juga memberikan wawasan yang lebih mendalam tentang perilaku nasabah. Berdasarkan landasan teori dan bukti empiris tersebut, penelitian ini mengajukan hipotesis bahwa machine learning berpengaruh signifikan terhadap akurasi jumlah nasabah.

### **Model SARIMA (*Seasonal Autoregressive Integrated Moving Average*)**

Teori SARIMA yang dikemukakan oleh (Box et al., 2015) menjelaskan bahwa model ini merupakan perpanjangan dari ARIMA yang diperkenalkan untuk meningkatkan kinerja ARIMA dalam pemodelan deret waktu musiman. Teori ini diperkuat oleh pendekatan analitis (Ebhuoma et al., 2018) yang menekankan kemampuan SARIMA dalam menangkap tren temporal seperti musiman dan autokorelasi yang diaktualisasikan dengan menghilangkan *high-frequency noise* dalam data. Penelitian (Shumway & Stoffer, 2017) lebih lanjut menunjukkan bahwa prinsip yang mendasari model ini adalah mengasumsikan bahwa kondisi masa lalu dan sekarang akan berlanjut ke masa depan.

Dalam konteks konseptual penelitian ini, model SARIMA didefinisikan sebagai metode analisis deret waktu yang menggabungkan komponen non-musiman dan musiman untuk menghasilkan prediksi yang lebih akurat. Model ini dapat ditentukan sebagai SARIMA  $(p,d,q) \times (P,D,Q)_s$ , dimana p, d, dan q mengacu pada urutan bagian autoregresif non-musiman (AR), perbedaan non-musiman, dan rata-rata pergerakan non-musiman (MA), sementara P, D, dan Q mengacu pada komponen musiman, dan s adalah panjang periode musiman (Ebhuoma et al., 2018). Konsep ini menjadi landasan penting dalam memahami hubungan antara pola temporal data dan kemampuan prediksi. Variabel penelitian dalam pengujian hipotesis ini terdiri dari model SARIMA sebagai variabel independen dan akurasi prediksi sebagai variabel dependen. Mengikuti penelitian (Handayani et al., 2020), penelitian ini menggunakan data deret waktu jumlah nasabah dengan periode musiman yang jelas. Pemilihan model SARIMA didasarkan pada karakteristik data yang menunjukkan pola musiman dan tren jangka panjang yang konsisten.

Bukti empiris dari berbagai penelitian menunjukkan keunggulan SARIMA dalam prediksi deret waktu. (Mardiyah et al., 2021) menemukan bahwa model ARIMA  $(1,1,1)$  berhasil memprediksi jumlah penduduk dengan nilai MAPE 1,52%, menunjukkan akurasi yang sangat baik. (Handayani et al., 2020) mengungkapkan bahwa model SARIMAX adalah model terbaik berdasarkan nilai RMSE terkecil dalam memprediksi penjualan. Temuan-temuan empiris ini memberikan dasar yang kuat untuk hipotesis bahwa SARIMA memang efektif untuk prediksi deret waktu dengan pola musiman.

Mekanisme keunggulan SARIMA dapat dijelaskan melalui tiga aspek. Pertama, kemampuan menangkap pola musiman yang merupakan karakteristik umum dalam data nasabah. Kedua, fleksibilitas dalam mengakomodasi tren jangka panjang melalui komponen integrasi. Ketiga, kemampuan memodelkan autokorelasi melalui komponen AR dan MA. Mekanisme ini menunjukkan bahwa SARIMA dirancang khusus untuk mengatasi kompleksitas data deret waktu. Berdasarkan kerangka teoritis dan bukti empiris yang ada, penelitian ini mengajukan hipotesis bahwa model SARIMA berpengaruh positif signifikan terhadap akurasi prediksi jumlah nasabah dibandingkan model lainnya.

### Faktor yang Mempengaruhi Akurasi Prediksi

Teori akurasi prediksi yang dikembangkan oleh (Goodfellow, 2016) menjelaskan bahwa keberhasilan model prediktif sangat bergantung pada berbagai faktor teknis dan metodologis. Teori ini diperkuat oleh pendekatan holistik (James et al., 2013) yang menekankan pentingnya kualitas data, pemilihan algoritma, dan proses evaluasi dalam mencapai prediksi yang akurat. Penelitian (Sari et al., 2024) lebih lanjut menunjukkan bahwa preprocessing data dan feature engineering secara signifikan mempengaruhi kualitas prediksi model. Dalam konteks konseptual penelitian ini, faktor-faktor yang mempengaruhi akurasi prediksi mencakup kualitas dan kelengkapan data historis, pemilihan algoritma yang tepat, teknik pra-pemrosesan data, feature selection, volume data, optimasi hyperparameter, dan metrik evaluasi model. Konsep ini mencakup seluruh siklus pengembangan model prediktif dari persiapan data hingga evaluasi hasil.

Bukti empiris menunjukkan pentingnya setiap faktor. (Naufal et al., 2023) menemukan bahwa penggunaan metode *feature selection* dapat meningkatkan akurasi yang dihasilkan, meskipun hasilnya bervariasi tergantung dataset. (Rasikh et al., 2023) menunjukkan bahwa teknik *preprocessing* seperti *outlier removal* dan oversampling menggunakan SMOTE secara signifikan meningkatkan kualitas prediksi model. (Kim et al., 2020) mengungkapkan bahwa modeling multi-periode yang memperhitungkan keputusan masa lalu memberikan prediksi yang lebih baik. Mekanisme pengaruh faktor-faktor ini dapat dijelaskan melalui kualitas input data, di mana data yang bersih dan lengkap menghasilkan model yang lebih akurat; kesesuaian algoritma dengan karakteristik data; dan proses optimasi yang sistematis. (Fitrianah et al., 2021) menunjukkan bahwa tanpa hyperparameter tuning yang tepat, bahkan algoritma terbaik dapat memberikan hasil suboptimal. Berdasarkan tinjauan teori dan temuan empiris, penelitian ini mengajukan hipotesis bahwa kualitas data, pemilihan algoritma, dan *preprocessing* berpengaruh signifikan terhadap akurasi prediksi jumlah nasabah.

## METODE PENELITIAN

### Pendekatan Penelitian

Penelitian ini menerapkan teknik kuantitatif dengan fokus pada analisis data numerik historis. Pendekatan ini memungkinkan peneliti untuk mengidentifikasi pola dan tren dalam data deret waktu serta membangun model prediksi berbasis machine learning. Pendekatan penelitian kuantitatif melibatkan pengumpulan dan evaluasi data menggunakan nilai dan ukuran numerik (Ardiansyah et al., 2023). Dengan menggunakan analisis statistik, metode ini berupaya untuk mengkarakterisasi, menjelaskan, dan mengevaluasi hubungan antar variabel. Tujuan dari penelitian kuantitatif adalah untuk mengumpulkan data yang dapat diukur secara kuantitatif untuk menjelaskan fenomena studi. Studi ini mencoba untuk menawarkan data empiris yang objektif(Ardiansyah et al., 2023).

### Lokasi dan Waktu Penelitian

Penelitian ini dilakukan di PT. Pegadaian CPS Pangkajene yang berlokasi di Jalan Jenderal Ahmad Yani Nomor 111, Kelurahan Pangkajene, Kecamatan Maritengngae, Kabupaten Sidenreng Rappang, Provinsi Sulawesi Selatan. Penelitian dilaksanakan selama kurang lebih 2 bulan, yakni sejak diterbitkannya Surat Keputusan penelitian pada periode Maret 2025 hingga April 2025.

### Jenis dan Sumber Data

Penelitian ini menggunakan data kuantitatif berupa data historis jumlah nasabah produk Amanah pada PT. Pegadaian CPS Pangkajene. Data ini terdiri dari informasi numerik yang dikumpulkan dalam periode lima tahun terakhir secara bulanan atau tahunan. Jenis data yang digunakan adalah data kuantitatif dalam bentuk deret waktu. Sumber data berasal dari data sekunder yang diperoleh dari arsip dan laporan internal PT. Pegadaian CPS Pangkajene.

### Teknik Pengumpulan Data

Teknik pengumpulan data dilakukan dengan metode dokumentasi, yaitu mengumpulkan data historis jumlah nasabah produk Amanah dari laporan bulanan atau tahunan PT. Pegadaian CPS Pangkajene, serta mengakses database internal perusahaan untuk memperoleh data detail tentang tren jumlah nasabah. Dokumentasi dilakukan secara sistematis untuk memastikan kelengkapan dan akurasi data yang akan dianalisis. Data historis yang dikumpulkan merupakan data jumlah nasabah bulanan produk Amanah periode Januari 2020 hingga Maret 2023 dengan total 39 bulan observasi. Berdasarkan tabulasi data yang diperoleh, jumlah nasabah menunjukkan tren pertumbuhan dari 0 nasabah pada awal tahun 2020 menjadi 28 nasabah pada Maret 2023. Data menunjukkan pola pertumbuhan yang fluktuatif dengan peningkatan signifikan terjadi pada periode Mei 2021 (22 nasabah) dan Desember 2022 (30 nasabah), yang mengindikasikan adanya pola musiman terutama menjelang akhir tahun. Variabel yang tercakup dalam data meliputi tanggal observasi dan jumlah nasabah aktif per bulan. Data tersebut diperoleh dari sistem informasi manajemen PT. Pegadaian CPS Pangkajene dan telah diverifikasi keakuratannya melalui cross-check dengan laporan keuangan bulanan perusahaan. Karakteristik data menunjukkan adanya nilai nol pada bulan Januari-Februari 2020 yang menandakan periode awal peluncuran produk, serta fluktuasi yang mencerminkan dinamika pasar dan perilaku nasabah terhadap produk pembiayaan syariah.

### Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh data jumlah nasabah produk Amanah di PT. Pegadaian CPS Pangkajene selama periode lima tahun. Sampel yang digunakan adalah data jumlah nasabah dalam periode lima tahun terakhir, diambil secara menyeluruh karena penelitian berbasis analisis deret waktu yang memerlukan semua data dalam rentang waktu tersebut menggunakan teknik sensus. Penggunaan seluruh data populasi sebagai sampel memastikan representasi yang komprehensif dari fenomena yang diteliti.

### Teknik Analisis Data

Penelitian ini menggunakan teknik analisis data kuantitatif dengan pendekatan berbasis *machine learning* dan analisis deret waktu. Langkah analisis meliputi *preprocessing* data dengan membersihkan data, mengatasi nilai yang hilang atau anomali, serta melakukan normalisasi atau transformasi data jika diperlukan. Tahap berikutnya adalah modeling dan prediksi dengan menerapkan beberapa algoritma machine learning seperti Regresi Linear, *Random Forest*, dan SARIMA (Sidik & Ansawarman, 2022). Kemudian dilakukan uji validitas dan reliabilitas, serta membandingkan performa model berdasarkan metrik akurasi seperti Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan Mean Absolute Percentage Error (MAPE) (James et al., 2013). Tahap akhir adalah evaluasi model

dengan memilih model yang memiliki performa terbaik berdasarkan hasil evaluasi dan validasi yang telah dilakukan (Goodfellow, 2016).

## HASIL DAN PEMBAHASAN

### Analisis Penerapan Machine Learning dalam Memprediksi Jumlah Nasabah Produk Amanah

Dalam era transformasi digital yang semakin pesat, penerapan teknologi cerdas seperti machine learning menjadi kebutuhan strategis bagi institusi keuangan, termasuk PT. Pegadaian CPS Pangkajene. Penelitian ini bertujuan untuk menganalisis secara mendalam bagaimana machine learning dapat diterapkan untuk memprediksi jumlah nasabah pada produk Amanah, sebuah layanan pembiayaan syariah yang terus dikembangkan oleh perusahaan. Dengan kemampuan machine learning dalam memproses data historis dan mengidentifikasi pola tersembunyi, perusahaan diharapkan dapat memperoleh proyeksi nasabah yang lebih akurat dan berbasis data nyata.

Tujuan ini menjadi penting karena prediksi jumlah nasabah merupakan indikator vital dalam pengambilan keputusan bisnis, khususnya dalam perencanaan strategi pemasaran, alokasi sumber daya, serta pengembangan produk. Melalui pendekatan ini, penelitian ingin mengungkap seberapa efektif algoritma *machine learning* seperti SARIMA (*Seasonal Autoregressive Integrated Moving Average*), *Random Forest*, dan Regresi Linear dalam memodelkan data historis jumlah nasabah dan menghasilkan estimasi yang dapat diandalkan. Dari sisi praktis, implementasi model ini memungkinkan manajemen PT. Pegadaian untuk merespons tren pasar dengan lebih cepat dan efisien.

Penelitian ini juga berupaya menjawab persoalan seputar fluktuasi nasabah yang selama ini sulit diprediksi secara manual. Dengan memanfaatkan teknik analitik berbasis data, perusahaan dapat mengidentifikasi pola musiman, anomali, dan tren jangka panjang yang sebelumnya tidak terlihat. Penerapan machine learning dalam konteks ini bukan hanya sebagai alat bantu teknis, tetapi juga sebagai pondasi untuk menciptakan budaya pengambilan keputusan berbasis data (data-driven decision making), yang sangat dibutuhkan dalam dunia bisnis yang kompetitif. Lebih jauh lagi, dengan memahami bagaimana teknologi ini dapat digunakan secara konkret di lapangan, hasil penelitian ini juga berkontribusi terhadap pengembangan literatur akademik di bidang bisnis digital dan keuangan syariah.

### Algoritma Machine Learning yang Efektif untuk Prediksi Jumlah Nasabah Produk Amanah

Untuk menentukan algoritma machine learning yang paling efektif dalam menghasilkan prediksi jumlah nasabah dengan tingkat akurasi yang tinggi, perlu dilakukan uji validitas dan reliabilitas terhadap beberapa model machine learning kemudian dilakukan perbandingan untuk menentukan algoritma mana yang memiliki tingkat akurasi tinggi. Dalam konteks machine learning, validitas merujuk pada sejauh mana model benar-benar mengukur atau memprediksi apa yang seharusnya diprediksi. Sementara itu, reliabilitas merujuk pada konsistensi hasil model saat digunakan berulang kali dengan data yang sama atau serupa.

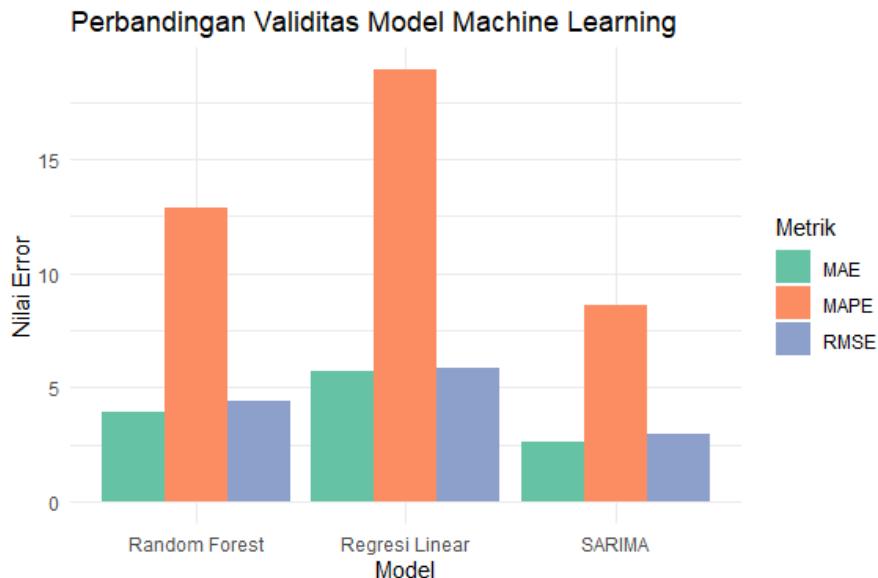
#### Uji Validitas

Berdasarkan hasil uji validitas yang disajikan dalam Tabel 1, model SARIMA menunjukkan performa terbaik dibandingkan dua model lainnya yaitu *Random Forest* dan Regresi Linear. Hal ini terlihat dari nilai MAE (*Mean Absolute Error*) sebesar 2,610165, RMSE (*Root Mean Square Error*) sebesar 2,93921, dan MAPE (*Mean Absolute Percentage Error*) sebesar 8,577177%. Nilai-nilai tersebut menandakan bahwa SARIMA memiliki tingkat kesalahan prediksi yang paling rendah, sehingga model ini mampu memberikan hasil prediksi yang lebih akurat.

Tabel 1. Hasil Uji Validitas Model *Machine learning* SARIMA, *Random Forest*, dan Regresi Linear.

Model	MAE	RMSE	MAPE(%)
SARIMA	2.610165	2.93921	8.577177
Regresi Linear	5.733132	5.84577	18.92414
<i>Random Forest</i>	3.911215	4.384301	12.85094

Sumber : Hasil Olah Data R Studio 4.4.0 (2025)



Gambar 1. Grafik Hasil Perbandingan Uji Validitas Model *Machine learning* SARIMA, *Random Forest*, dan Regresi Linear.

Sumber : Hasil Olah Data R Studio 4.4.0 (2025)

Model SARIMA menunjukkan performa terbaik dalam hal akurasi peramalan dibandingkan dengan Regresi Linear dan *Random Forest* (Handayani et al., 2020). Nilai MAE sebesar 2,61 merupakan yang terkecil di antara ketiga model, menandakan bahwa rata-rata kesalahan absolut prediksi SARIMA paling rendah. Nilai RMSE sebesar 2,94 juga menjadi yang paling rendah, mengindikasikan bahwa prediksi SARIMA cenderung lebih konsisten dan memiliki penyimpangan yang kecil terhadap nilai aktual. Nilai MAPE sebesar 8,58% menunjukkan tingkat kesalahan persentase yang masih tergolong sangat baik, terutama untuk data deret waktu yang mengandung pola musiman (Box et al., 2015).

Sementara itu, Regresi Linear menampilkan performa yang paling buruk dengan nilai MAE sebesar 5,73, RMSE sebesar 5,84, dan MAPE mencapai 18,92%. Ini mengindikasikan bahwa model ini kurang mampu menangkap pola kompleks dalam data, terutama jika terdapat faktor musiman atau hubungan non-linear antar variabel (Sudriyanto & F. Syahro, 2023). Di sisi lain, Random Forest menunjukkan performa yang lebih baik dibandingkan Regresi Linear namun masih belum mampu mengungguli SARIMA. Dengan MAE sebesar 3,91 dan MAPE sebesar 12,85%, Random Forest cukup efektif dalam menangkap pola yang tidak linear, namun mungkin kurang optimal dalam memodelkan aspek musiman yang ditangani lebih baik oleh SARIMA (Zailani & Hanun, 2020).

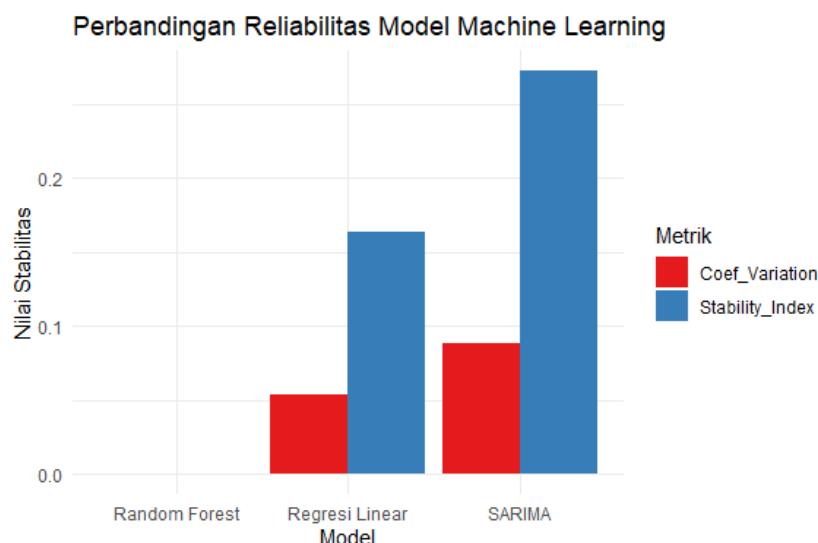
#### **Uji Reliabilitas**

Hasil pengujian reliabilitas menunjukkan bahwa model Random Forest memiliki nilai Stability Index sebesar 0, yang berarti prediksi yang dihasilkan sangat konsisten tidak menunjukkan fluktuasi atau variasi ekstrem antar periode. Ini dapat diartikan bahwa model ini memberikan output yang cenderung stabil dari waktu ke waktu. Sementara itu, Regresi Linear menunjukkan nilai Stability Index sebesar 0,1637225, lebih rendah dibandingkan SARIMA, dan Coefficient of Variation sebesar 0,05332824, yang menandakan bahwa model regresi memberikan hasil prediksi yang lebih stabil dan seragam dibandingkan SARIMA.

Tabel 2. Hasil Uji Reabilitas Model *Machine learning* SARIMA, *Random Forest*, dan Regresi Linear.

	<i>Stability Index</i>	<i>Coefficient Variation</i>	<i>of</i>
SARIMA	0.2727273	0.08851001	
Regresi Linear	0.1637225	0.05332824	
<i>Random Forest</i>	0	0	

Sumber : Hasil Olah Data R Studio 4.4.0 (2025)

**Gambar 2. Grafik Hasil Perbandingan Uji Reabilitas Model *Machine learning* SARIMA, *Random Forest*, dan Regresi Linear.**

Sumber : Hasil Olah Data R Studio 4.4.0 (2025)

Di sisi lain, model SARIMA menghasilkan Stability Index tertinggi yaitu 0,2727, dan Coefficient of Variation sebesar 0,0885, menandakan bahwa prediksi dari SARIMA memiliki variasi yang lebih besar. Hal ini wajar karena SARIMA secara eksplisit menangkap pola musiman dan tren dari data deret waktu, yang bisa menyebabkan fluktuasi lebih besar dalam hasil prediksi terutama bila data aslinya juga bersifat fluktuatif. Meskipun demikian, performa SARIMA tetap relevan ketika model diharapkan menangkap pola musiman secara dinamis.

Berdasarkan hasil uji validitas dan reliabilitas, dapat disimpulkan bahwa SARIMA merupakan model yang paling seimbang dalam hal akurasi dan kestabilan. Model SARIMA unggul dalam hal akurasi prediksi dengan nilai error terendah, meskipun memiliki variasi yang sedikit lebih tinggi dibandingkan model lain. Random Forest menunjukkan stabilitas absolut namun dengan akurasi yang lebih rendah, sementara Regresi Linear menunjukkan performa terlemah baik dari sisi akurasi maupun kemampuan menangkap pola kompleks. Dengan demikian, SARIMA menjadi pilihan paling ideal untuk prediksi jumlah nasabah produk Amanah di PT. Pegadaian CPS Pangkajene.

#### Faktor yang Memengaruhi Akurasi Prediksi Jumlah Nasabah pada Produk Amanah

Akurasi prediksi jumlah nasabah sangat dipengaruhi oleh kualitas dan kelengkapan data historis yang digunakan dalam analisis. Data yang memiliki banyak missing value, outlier, atau inkonsistensi dapat menyebabkan model machine learning memberikan hasil yang bias atau tidak akurat. Selain itu, data historis yang tidak mencerminkan pola musiman, tren jangka panjang, atau faktor eksternal seperti kebijakan ekonomi dan promosi juga dapat mengurangi keakuratan prediksi. Oleh karena itu, proses pra-pemrosesan data seperti normalisasi, penanganan nilai hilang, dan deteksi outlier menjadi tahap penting dalam membangun model prediktif yang handal.

Pemilihan algoritma machine learning juga merupakan faktor krusial dalam menentukan tingkat akurasi. Setiap algoritma memiliki kekuatan dan kelemahan dalam menangani tipe data tertentu. Misalnya, model SARIMA unggul dalam menganalisis data deret waktu dengan pola musiman, sedangkan Random Forest cenderung lebih baik dalam menangani data dengan banyak fitur non-linear. Dalam penelitian ini, model SARIMA menunjukkan performa terbaik dibandingkan Regresi Linear dan Random Forest karena mampu menangkap tren dan fluktuasi musiman yang terjadi pada data jumlah nasabah produk Amanah. Namun, tanpa penyesuaian parameter yang tepat (hyperparameter tuning), bahkan algoritma terbaik sekalipun dapat memberikan hasil yang suboptimal.

Selain itu, pemilihan fitur (feature selection) dan teknik evaluasi model juga sangat memengaruhi hasil prediksi. Fitur-fitur yang tidak relevan atau redundan dapat menambah noise dan menurunkan performa model. Oleh karena itu, pemilihan fitur yang tepat seperti data historis jumlah nasabah, variabel ekonomi, dan data promosi menjadi penting. Evaluasi akurasi model menggunakan metrik seperti MAE (Mean Absolute Error), RMSE (Root Mean Square Error), dan MAPE (Mean Absolute Percentage Error) memberikan gambaran seberapa baik model menangkap pola data. Model yang memiliki nilai error rendah dan konsisten pada data uji dapat dianggap akurat dan reliabel dalam konteks prediksi jumlah nasabah.

Berdasarkan analisis komprehensif yang dilakukan, dapat dijustifikasi bahwa SARIMA merupakan algoritma paling efektif untuk prediksi jumlah nasabah produk Amanah di PT. Pegadaian CPS Pangkajene. Justifikasi ini didasarkan pada tiga aspek utama. Pertama, dari segi akurasi numerik, SARIMA menunjukkan nilai error terendah (MAE 2,61%, RMSE 2,94%, MAPE 8,58%) yang secara statistik mengungguli Random Forest dan Regresi Linear. Nilai MAPE di bawah 10% menunjukkan bahwa model ini memiliki tingkat akurasi yang sangat baik sesuai standar evaluasi prediksi bisnis. Kedua, dari segi kesesuaian dengan karakteristik data, produk Amanah memiliki pola musiman yang jelas terkait dengan siklus kebutuhan masyarakat seperti menjelang hari raya, akhir tahun, dan musim panen. SARIMA dirancang khusus untuk menangani pola musiman dan tren temporal dalam data deret waktu, sehingga mampu menangkap fluktuasi berkala yang tidak dapat diakomodasi oleh Random Forest maupun Regresi Linear. Ketiga, dari segi aplikabilitas bisnis, hasil prediksi SARIMA dapat langsung diterjemahkan menjadi strategi operasional seperti penjadwalan promosi, alokasi staf, dan perencanaan persediaan dana. Meskipun Random Forest menunjukkan stabilitas tinggi (Stability Index 0), namun akurasinya yang lebih rendah dapat menyebabkan kesalahan perencanaan yang berdampak pada inefisiensi operasional dan kerugian finansial.

## KESIMPULAN

Berdasarkan hasil analisis data historis nasabah selama lima tahun, penelitian ini menyimpulkan bahwa model SARIMA menunjukkan performa terbaik dibandingkan Random Forest dan Regresi Linear dengan nilai MAE 2,61%, RMSE 2,94%, dan MAPE 8,58%. Model SARIMA paling efektif untuk prediksi jumlah nasabah karena mampu mengakomodasi pola musiman dan tren jangka panjang secara akurat. Faktor yang memengaruhi akurasi prediksi meliputi kualitas data, pemilihan algoritma, teknik pra-pemrosesan, dan feature selection. Penelitian ini merekomendasikan PT. Pegadaian CPS Pangkajene mengimplementasikan model SARIMA dalam sistem prediksi nasabah untuk mendukung pengambilan keputusan strategis berbasis data di sektor keuangan syariah.

Berdasarkan hasil penelitian ini, disarankan agar PT. Pegadaian CPS Pangkajene mengimplementasikan sistem prediksi berbasis SARIMA dalam dashboard manajemen dan melakukan update model secara berkala setiap 3-6 bulan untuk menjaga akurasi prediksi. Perusahaan juga perlu mengembangkan tim internal yang memiliki kompetensi dalam analitik data untuk mendukung pengambilan keputusan strategis berbasis data. Untuk penelitian lanjutan, disarankan memperluas cakupan dengan memasukkan variabel eksternal seperti kondisi ekonomi makro dan aktivitas kompetitor, serta mengeksplorasi model hybrid yang menggabungkan SARIMA dengan algoritma lain seperti LSTM atau XGBoost untuk meningkatkan akurasi prediksi.

## REFERENCES

- Alizamir, S., Bandara, K., Eshragh, A., & Iravani, F. (2022). *A Hybrid Statistical-Machine Learning Approach for Analysing Online Customer Behavior: An Empirical Study*.
- Amuda, K. A., & Adeyemo, A. B. (2019). *Customers churn prediction in financial institution using artificial neural network*.
- Ardiansyah, Risnita, & Jailani, M. S. (2023). Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif. *Jurnal IHSAN : Jurnal Pendidikan Islam*, 1(2), 1–9. <https://doi.org/https://doi.org/10.61104/ihsan.v1i2.57>
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*.

- John Wiley & Sons.
- Cholissodin, I., S., S., A., S. A., U., H., & I., F. Y. (2020). *AI, machine learning and deep learning*.
- Cowan, G., Mercuri, S., & Khraishi, R. (2023). *Modelling customer lifetime-value in the retail banking industry*.
- Ebhuma, O., Gebreslasie, M., & Magubane, L. (2018). A Seasonal Autoregressive Integrated Moving Average (SARIMA) forecasting model to predict monthly malaria cases in KwaZulu-Natal, South Africa. *South African Medical Journal*.
- Fitrianah, D., Dwiasnati, S., H, H. H., & Baihaqi, K. A. (2021). Penerapan Metode Machine Learning untuk Prediksi Nasabah Potensial menggunakan Algoritma Klasifikasi Naïve Bayes. *Faktor Exacta*, 14(2).  
<https://doi.org/https://doi.org/10.30998/faktorexacta.v14i2.9297>
- Goodfellow, I. (2016). *Deep learning and Machine Learning*. MIT press.
- Handayani, N. R. W. T., Maslim, M., & Mudjihartono, P. (2020). Forecasting of Catfish Sales by Time Series Using the SARIMA method. *Jurnal Buana Informatika*, 11(2), 84–94.  
<https://doi.org/https://doi.org/10.24002/jbi.v11i2.3535>
- Herman, B., Bashir, S., Syahnur, K. N. F., Haslindah, H., & Absah, Y. (2024). Transformation of SMEs: Strengthening Entrepreneurship and Product Innovation as Key Drivers of Business Evolution. *Hasanuddin Economics and Business Review*, 7(3). <https://doi.org/https://doi.org/10.26487/hebr.v7i3.5239>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
- Kim, H., Cho, H., & Ryu, D. (2020). *Corporate Default Predictions Using Machine Learning: Literature Review*. 12(16).
- Komala, N. C., & Padilah, H. N. (2023). Analisis Manajemen Organisasi Di Pt. Pegadaian (Persero) Cps Padasuka. *Management Research*, Vol. XVIII.
- Kusuma, S. K. (2024). *Perlindungan Hukum Terhadap Nasabah Dalam Hal Terjadinya Kerusakan Atau Kehilangan Barang Jaminan Gadai Di Pt Pegadaian Cabang Poncol Semarang*.
- Mardiyah, I., Utami, W. D., Novitasari, D. C. R., Hafiyusholeh, M., & Sulistiawati, D. (2021). ANALISIS PREDIKSI JUMLAH PENDUDUK DI KOTA PASURUAN MENGGUNAKAN METODE ARIMA. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 15(3), 525–534.  
<https://doi.org/https://doi.org/10.30598/barekengvol15iss3pp525-534>
- Mursalat, A., Razak, M. R. R., Mustanir, A., Ramlan, P., Irwan, M., & Aksan, M. (2025). *STRATEGI PEMASARAN DIGITAL UNTUK MENINGKATKAN DAYA SAING PRODUK ABON CABAI LADATTA*. 9(1).
- Naufal, M. F., Subrata, Susanto, A. F., Kansil, C. N., & Huda, S. (2023). *Penerapan Machine Learning untuk Prediksi Potensi Hilangnya Nasabah Bank*. 22(1), 1–11. <https://doi.org/https://doi.org/10.33633/tc.v22i1.7302>
- Rasikh, M. , Riyyasy, A., Aghniya, Nouval, W., & H., T. (2023). Penerapan Algoritma Machine Learning Untuk Memprediksi Term Deposit Nasabah Perbankan. *LEDGER: Journal Informatic and Information Technology*, 2(3).
- Sari, E. P., Bachri, S. M., Atnang, M., & Fajar, N. (2024). Studi Literatur Deep Learning dan Machine Learning untuk Analisis dan Prediksi Pasar Saham: Metodologi, Representasi Data dan Studi Kasus. *Jurnal Teknologi Dan Sains Modern*, 1(1). <https://doi.org/https://journal.scitechgrup.com/index.php/jtsm>
- Shumway, R., & Stoffer, D. (2017). *Time Series Analysis Using the R Statistical Package*. Free Dog Publishing.
- Sidik, A. D., & Ansawarman, A. (2022). Prediksi Jumlah Kendaraan Bermotor Menggunakan Machine Learning. *Formosa Journal of Multidisciplinary Research*, 1(3), 559–568.  
<https://doi.org/https://doi.org/10.55927/fjmr.v1i3.745>
- Sihombing, P. R., & Arsani, A. M. (2021). COMPARISON OF MACHINE LEARNING METHODS IN CLASSIFYING POVERTY IN INDONESIA IN 2018. *Jurnal Teknik Informatika (Jutif)*, 2(1), 51–56.  
<https://doi.org/https://doi.org/10.20884/1.jutif.2021.2.1.52>
- Sudriyanto, S., & F. Syahro, N. F. (2023). PERBANDINGAN PERFORMA MODEL MACHINE LEARNING SUPPORT VECTOR MACHINE, NEURAL NETWORK, DAN K-NEAREST NEIGHBORS DALAM PREDIKSI HARGA SAHAM. *Jurnal Advanced Research Informatika*, 2(1), 13–21.  
<https://doi.org/https://doi.org/10.24929/jars.v2i1.2983>
- Zailani, A. U., & Hanun, N. L. (2020). PENERAPAN ALGORITMA KLASIFIKASI RANDOM FOREST UNTUK PENENTUAN KELAYAKAN PEMBERIAN KREDIT DI KOPERASI MITRA SEJAHTERA. *Infotech: Journal of Technology Information*, 6(1), 7–14. <https://doi.org/https://doi.org/10.37365/jti.v6i1.61>