KAMPUS AKADEMIK PUBLISING
Jurnal Sains Student Research (JSSR)
Vol.4, No.6 Desember 2026
e-ISSN: 3025-9851, p-ISSN : 3025-986X, Hal 478-493
DOI : https://doi.org/10.61722/jssr.v4i6.12938

Optimasi Prediksi Harga Kartu Game Berdasarkan Fitur Ekosistem Pasar Lokal Menggunakan Support Vector Regression Berbasis Kernel dan Linear Regression

M Arief Sugondo Nst[1], Irvan[2*]
[1,2] Program Studi Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia
*Penulis Korespondensi: ariefsugondo04@gmail.com, *irvan@umsu.ac.id

Abstract. *The local trading card game (TCG) marketplace in Indonesia has grown rapidly, yet card prices are still determined subjectively through individual negotiation, which is prone to information asymmetry among market participants. This study builds a card price prediction model based on local market ecosystem features (rarity, condition, card age since release, card name, and card set) by comparing Support Vector Regression (SVR) with a Radial Basis Function (RBF) kernel against Linear Regression as a baseline. The dataset consists of 60,530 rows of monthly price data for 1,000 unique cards observed from 2020 to mid-2026. The data were processed through ordinal and binary encoding, card-age feature engineering, one-hot encoding, and standardization before being split into training and test sets (80:20). The SVR hyperparameters (RBF kernel, C=500, gamma=0.5) were obtained through GridSearchCV with 3-fold cross-validation testing linear, RBF, and polynomial kernels. The RBF-kernel SVR model consistently outperformed Linear Regression across all metrics, achieving R^2 of 0.604 versus 0.181, MAE of IDR 11,667,484 versus IDR 18,290,123 (a 36.2% reduction), and RMSE of IDR 32,183,966 versus IDR 46,258,319 (a 30.4% reduction). SVR performance was strongest for high-rarity cards (Ultra Rare, R^2=0.736) compared to Common cards (R^2=0.388). The nonlinear kernel-based SVR approach is better suited to modeling the nonlinear pattern of card prices and can serve as a baseline price-estimation tool for collectors, sellers, and card-trading platforms.*

Keywords: *trading card game, price prediction, Support Vector Regression, Radial Basis Function kernel, Linear Regression, local market ecosystem*

Abstrak. Ekosistem jual beli kartu game koleksi (trading card game/TCG) di pasar lokal Indonesia berkembang pesat, namun harga kartu masih ditentukan secara subjektif melalui negosiasi individual sehingga rawan terhadap informasi asimetris antarpelaku pasar. Penelitian ini bertujuan membangun model prediksi harga kartu game berdasarkan fitur ekosistem pasar lokal, yaitu rarity, condition, usia kartu sejak rilis, nama kartu, dan set penerbitan, dengan membandingkan performa Support Vector Regression (SVR) berkernel Radial Basis Function (RBF) terhadap Linear Regression sebagai baseline. Dataset yang digunakan terdiri dari 60.530 baris data harga bulanan dari 1.000 kartu unik pada periode 2020 hingga pertengahan 2026. Data diproses melalui encoding ordinal dan biner, rekayasa fitur usia kartu, one-hot encoding variabel kategorik, serta standardisasi fitur numerik, sebelum dibagi menjadi data latih dan data uji dengan proporsi 80:20. Hyperparameter SVR (kernel RBF, C=500, gamma=0,5) diperoleh melalui GridSearchCV dengan skema 3-fold cross-validation yang menguji kernel linear, RBF, dan polynomial. Hasil pengujian menunjukkan bahwa SVR kernel RBF secara konsisten mengungguli Linear Regression pada seluruh metrik evaluasi, dengan R^2 sebesar 0,604 berbanding 0,181, MAE sebesar Rp11.667.484 berbanding Rp18.290.123 (turun 36,2%), dan RMSE sebesar Rp32.183.966 berbanding Rp46.258.319 (turun 30,4%). Performa SVR juga paling optimal pada kartu bertingkat rarity tinggi (Ultra Rare, R^2=0,736) dibandingkan kartu Common (R^2=0,388). Penelitian ini menyimpulkan bahwa pendekatan SVR berbasis kernel nonlinier lebih sesuai untuk memodelkan pola harga kartu game yang bersifat nonlinier dan dapat dimanfaatkan sebagai alat bantu estimasi harga awal bagi kolektor, penjual, maupun platform jual beli kartu di ekosistem pasar lokal.

Kata kunci: *kartu game, prediksi harga, support vector regression, kernel radial basis function, linear regression, ekosistem pasar lokal*

Naskah Masuk:01 Oktober 2026; Revisi: 02 Oktober 2026; Diterima: 04 Oktober 2026; Terbit: 07 Oktober 2026

1. LATAR BELAKANG

Industri permainan kartu koleksi atau Trading Card Game (TCG) berkembang pesat di Indonesia dalam satu dekade terakhir. Kartu game yang semula hanya media hiburan kini menjadi barang koleksi dengan nilai ekonomi yang cukup tinggi, didorong oleh komunitas pemain dan kolektor yang aktif bertransaksi melalui grup Facebook, server Discord, dan marketplace daring.

Berdasarkan data harga historis kartu Pokémon versi bahasa Indonesia yang dikumpulkan dalam penelitian ini, tercatat 60.530 data harga bulanan dari 1.000 kartu yang berasal dari 11 set resmi dalam rentang tahun 2020–2026. Kartu dengan tingkat kelangkaan tertinggi (Ultra Rare) dapat dihargai puluhan hingga ratusan kali lipat dibandingkan kartu Common dari set yang sama, yang menunjukkan bahwa pasar kartu game lokal memiliki dinamika harga yang kompleks dan sulit diprediksi hanya dengan perkiraan sederhana. Harga kartu dipengaruhi banyak faktor sekaligus, bukan hanya kondisi fisik: tingkat kelangkaan (rarity), usia kartu sejak rilis, set penerbitan, hingga pengaruh komunitas terbukti membentuk pola harga yang berulang. Aldiansyah et al. (2025) menemukan bahwa brand awareness dan gaya hidup komunitas memengaruhi keputusan pembelian kartu Pokémon, sementara Mekacher et al. (2022) membuktikan bahwa pola kelangkaan yang heterogen menjadi penggerak utama dinamika harga pada aset koleksi digital, temuan yang relevan pula bagi kartu fisik.

Permasalahan utama pelaku pasar saat ini adalah belum adanya sistem penilaian harga yang objektif dan berbasis data; penentuan harga masih dilakukan secara manual dengan melihat transaksi terakhir atau perkiraan pribadi, yang rentan terhadap kesalahan penilaian dan lambat merespons lonjakan harga mendadak (price spikes) akibat tren turnamen lokal maupun hype komunitas. Pendekatan machine learning terbukti mampu mengatasi kelemahan metode manual pada berbagai domain prediksi harga: Wood et al. (2024) mendemonstrasikan estimasi harga yang lebih akurat pada lebih dari 1,2 juta data lelang kartu Pokémon, sementara Aulia et al. (2022) dan Hidayatulloh & Yuwono (2025) membuktikan bahwa Support Vector Regression (SVR) memprediksi harga emas dengan kesalahan lebih rendah dibandingkan metode statistik tradisional.

Meskipun penelitian prediksi harga berbasis machine learning cukup banyak, kajian yang secara khusus membandingkan performa SVR berkernel dengan Linear Regression

pada domain kartu game koleksi masih terbatas. Penelitian terdahulu pada kartu koleksi (Wood et al., 2024; Mekacher et al., 2022) lebih banyak menggunakan deep learning atau analisis statistik sederhana, sehingga perbandingan sistematis SVR kernel linear, SVR kernel RBF, dan Linear Regression pada domain ini masih jarang dilakukan. Untuk menjawab kesenjangan tersebut, penelitian ini membangun dan membandingkan tiga varian model — SVR kernel linear, SVR kernel RBF, dan Linear Regression — dengan proses optimasi hyperparameter melalui GridSearchCV dan tahap preprocessing (standardisasi dan seleksi fitur) yang diterapkan secara konsisten agar perbandingan antarmodel bersifat adil (Gananta et al., 2023; Hidayatulloh & Yuwono, 2025).

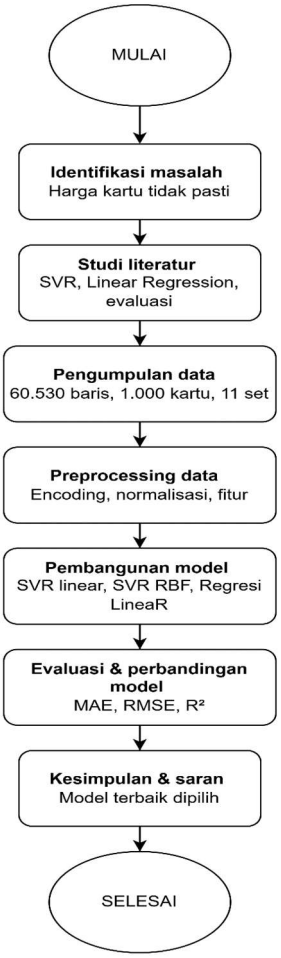
Penelitian ini penting karena pasar kartu game lokal Indonesia terus tumbuh namun belum memiliki mekanisme penilaian harga berbasis data, sehingga pelaku pasar tetap bergantung pada penilaian subjektif yang rentan menimbulkan kerugian finansial. Berdasarkan uraian tersebut, penelitian ini dirancang untuk membangun, mengoptimalkan melalui Grid Search, dan membandingkan model prediksi harga kartu game menggunakan SVR berbasis kernel dan Linear Regression pada dataset 60.530 baris data historis pasar kartu game lokal Indonesia, dengan kontribusi utama berupa (1) perbandingan sistematis SVR kernel linear, RBF, dan Linear Regression pada domain kartu koleksi fisik, (2) analisis performa model pada berbagai segmen rarity dan rentang harga, serta (3) identifikasi keterbatasan dan arah pengembangan lanjutan bagi riset prediksi harga aset ekonomi alternatif.

2. METODE PENELITIAN

2.1 Jenis Penelitian

Penelitian ini merupakan penelitian terapan (applied research) dengan pendekatan kuantitatif, karena bertujuan mengimplementasikan metode machine learning untuk menyelesaikan permasalahan nyata berupa prediksi harga kartu Pokémon versi Indonesia berdasarkan data numerik yang terstruktur. Penelitian mengikuti alur eksperimental komparatif: membangun, melatih, dan mengevaluasi lebih dari satu model prediksi pada kondisi data yang sama, kemudian membandingkan performanya secara objektif menggunakan metrik evaluasi yang terukur. Tahapan penelitian secara garis besar mencakup pengumpulan dan pemahaman data, preprocessing, rekayasa fitur, pembagian

data latih-uji, pembangunan dan pelatihan model, evaluasi dan perbandingan model, hingga penarikan kesimpulan, sebagaimana diilustrasikan pada Gambar 1.



Gambar 1. *Diagram Alur Penelitian*

2.2 Sumber dan Deskripsi Data

Data yang digunakan merupakan data sekunder yang direkonstruksi dari catatan harga transaksi dan listing kartu Pokémon versi bahasa Indonesia pada marketplace daring dan grup komunitas jual-beli kartu, disusun ulang secara bulanan dalam format Microsoft Excel. Validitas harga dijaga dengan mengambil harga transaksi/penawaran yang representatif untuk setiap kombinasi kartu, kondisi, dan bulan pencatatan sehingga outlier akibat kesalahan input dapat diminimalkan. Dataset final berisi 60.530 baris data harga bulanan dari 1.000 kartu (20 nama karakter berbeda) yang berasal dari 11 set resmi (Base Set, Jungle, Fossil, Neo Genesis, EX Ruby & Sapphire, Diamond & Pearl, Platinum, Black & White, XY, Sun & Moon, dan Sword & Shield) pada periode 2020

hingga pertengahan 2026, tanpa nilai yang hilang (missing values) pada seluruh kolom. Deskripsi variabel dataset disajikan pada Tabel 1.

No	Nama Kolom	Tipe Data	Deskripsi
1	Id Kartu	Teks	Kode unik pengenal kartu (C0001–C1000)
2	Nama Kartu	Teks	Nama karakter Pokémon pada kartu (20 kategori)
3	Set Kartu	Kategorikal	Nama set/ekspansi resmi (11 set)
4	Waktu Rilis	Tahun-Bulan	Periode resmi kartu mulai beredar (format YYYY-MM)
5	Rarity	Kategorikal	Tingkat kelangkaan: Common, Uncommon, Rare, Ultra Rare
6	Condition	Kategorikal	Kondisi fisik: NM (Near Mint) atau LP (Lightly Played)
7	Tahun	Numerik	Tahun harga tercatat (2020–2026)
8	Bulan	Numerik	Bulan harga tercatat (1–12)
9	Harga	Numerik	Harga kartu dalam Rupiah (variabel target)

Tabel 1. *Deskripsi Variabel Dataset*

2.3 Preprocessing dan Rekayasa Fitur

Preprocessing dilakukan agar data mentah siap digunakan oleh algoritma machine learning yang sensitif terhadap skala dan format data, mencakup tiga langkah utama. Pertama, pemeriksaan data meliputi pengecekan nilai hilang, tipe data, dan duplikasi; hasil pemeriksaan menunjukkan seluruh 60.530 baris data tidak memiliki nilai hilang. Kedua, encoding variabel kategorikal: Label Encoding ordinal diterapkan pada kolom Rarity (Common=1, Uncommon=2, Rare=3, Ultra Rare=4) dan encoding biner pada kolom Condition (LP=0, NM=1) karena kedua variabel memiliki hierarki yang bermakna,

sedangkan One-Hot Encoding diterapkan pada kolom Set Kartu (11 kategori) dan Nama Kartu (20 kategori) karena tidak memiliki urutan. Ketiga, rekayasa fitur (feature engineering) berupa Usia Kartu (dalam bulan), dihitung sebagai selisih antara periode observasi harga dan waktu rilis kartu: Usia Kartu = (Tahun_observasi × 12 + Bulan_observasi) − (Tahun_rilis × 12 + Bulan_rilis). Fitur ini penting karena kartu yang lebih tua cenderung memiliki dinamika harga berbeda terkait kelangkaan stok dan nostalgia premium.

Seluruh fitur numerik selanjutnya distandardisasi menggunakan StandardScaler (rata-rata 0, standar deviasi 1) karena SVR berbasis kernel RBF sensitif terhadap perbedaan skala antarfitur; parameter standardisasi dihitung dari data latih saja lalu diterapkan pada data uji untuk mencegah data leakage. Kolom Id Kartu tidak diikutsertakan sebagai fitur input karena bersifat pengenal unik, sedangkan kolom Waktu Rilis tidak digunakan langsung karena informasinya telah diturunkan ke fitur Usia Kartu. Fitur input final dan variabel target dirangkum pada Tabel 2. Dataset kemudian dibagi menjadi data latih dan data uji dengan proporsi 80:20 menggunakan skema acak (random split), menghasilkan 48.424 baris data latih dan 12.106 baris data uji, sehingga kedua model dapat diuji pada kondisi pembagian data yang identik.

No	Fitur	Jenis
1	Rarity (encoded ordinal 1–4)	Input
2	Condition (encoded biner 0/1)	Input
3	Usia Kartu dalam bulan (rekayasa fitur)	Input
4	Tahun dan Bulan pencatatan harga	Input
5	Set Kartu (one-hot, 11 kolom)	Input
6	Nama Kartu (one-hot, 20 kolom)	Input
7	Harga	Target

Tabel 2. *Fitur Input dan Variabel Target Model*

2.4 Pembangunan dan Optimasi Model

Dua model regresi dibangun dan dibandingkan performanya: Support Vector Regression (SVR) dengan kernel Radial Basis Function (RBF) sebagai model utama, dan Linear Regression sebagai model acuan (baseline) berupa Ordinary Least Squares standar tanpa regularisasi tambahan. SVR merupakan pengembangan Support Vector Machine untuk tugas regresi yang mencari fungsi $f(x) = w \cdot \varphi(x) + b$ dengan deviasi tidak melebihi toleransi ε dari nilai aktual (ε-insensitive loss), sehingga hanya titik-titik pada batas atau di luar tube (support vectors) yang menentukan fungsi prediksi akhir (Du et al., 2024). Kernel RBF, $K(x_i,x_j) = \exp(-\gamma\|x_i - x_j\|^2)$, dipilih karena mampu memetakan data ke ruang fitur berdimensi tinggi sehingga efektif menangkap hubungan nonlinear yang kompleks, sebagaimana ditunjukkan oleh Yang et al. (2022) dan Hidayatulloh & Yuwono (2025) pada data harga yang fluktuatif.

Pemilihan kernel terbaik beserta optimasi hyperparameter dilakukan melalui GridSearchCV dengan skema 3-fold cross-validation, menguji tiga jenis kernel (linear, RBF, polynomial) pada ruang pencarian $C \in \{100; 300; 500\}$ dan $\gamma \in \{0{,}5; 1\}$ (khusus kernel RBF dan polynomial), menggunakan rata-rata R^2 validasi silang sebagai skor acuan. Dengan pendekatan ini, istilah "optimasi" pada penelitian merujuk secara eksplisit pada proses tuning melalui GridSearchCV, bukan sekadar nilai default pustaka. Seluruh model diimplementasikan menggunakan Python dan pustaka scikit-learn, pandas, dan NumPy, dengan visualisasi menggunakan Matplotlib/Seaborn.

2.5 Evaluasi Model

Performa model dievaluasi pada data uji menggunakan tiga metrik: Mean Absolute Error ($MAE = (1/n)\Sigma|y_i - \hat{y}_i|$), Root Mean Squared Error ($RMSE = \sqrt{[(1/n)\Sigma(y_i - \hat{y}_i)^2]}$), dan koefisien determinasi ($R^2 = 1 - \Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2$). MAE dan RMSE dinyatakan dalam satuan Rupiah dengan nilai lebih kecil menunjukkan performa lebih baik, sedangkan R^2 bersifat scale-free dengan nilai mendekati 1 menunjukkan model mampu menjelaskan variasi harga secara lebih baik (Hodson, 2022). Ketiga metrik digunakan bersamaan agar penilaian performa model bersifat komprehensif.

3. HASIL DAN PEMBAHASAN

3.1 Karakteristik Data

Dataset penelitian terdiri dari 60.530 baris observasi harga bulanan dari 1.000 kartu unik, mencakup 20 nama karakter yang tersebar pada 11 set penerbitan, dengan periode pengamatan 2020 hingga pertengahan 2026. Statistik deskriptif variabel harga (Tabel 3) menunjukkan rata-rata harga sebesar Rp20.179.790 dengan standar deviasi Rp47.195.150 yang jauh lebih besar dari rata-rata, serta rentang antara median (Rp7.538.768) dan nilai maksimum (Rp1.733.743.000) yang sangat lebar. Kondisi ini mengindikasikan distribusi harga yang right-skewed dengan sejumlah kartu bernilai sangat tinggi (kartu langka edisi awal). Eksplorasi lanjutan menunjukkan bahwa rata-rata harga meningkat sejalan dengan tingkat kelangkaan (rarity), meskipun tidak selalu berbanding lurus sempurna, serta tren harga terhadap usia kartu (bulan sejak rilis) bersifat nonlinier dan fluktuatif menyerupai siklus pasar — pola yang menjadi salah satu dasar penggunaan SVR berkernel nonlinear.

Statistik	**Nilai (Rupiah)**
Jumlah Data	60.530
Rata-rata (Mean)	20.179.790
Standar Deviasi	47.195.150
Minimum	17.967
Kuartil 1 (25%)	3.475.893
Median (50%)	7.538.768
Kuartil 3 (75%)	19.181.310
Maksimum	1.733.743.000

Tabel 3. *Statistik Deskriptif Variabel Harga*

3.2 Hasil Optimasi Hyperparameter

Hasil GridSearchCV menunjukkan bahwa kernel RBF secara konsisten mengungguli kernel linear dan polynomial pada seluruh kombinasi yang diuji, dengan kombinasi terbaik C=500 dan γ=0,5 menghasilkan rata-rata R^2 validasi silang sebesar 0,605 (Tabel 4), mengonfirmasi bahwa hubungan antara fitur ekosistem pasar dan harga

kartu bersifat nonlinear. Mengingat kompleksitas komputasi kernel RBF yang berada pada rentang $O(n^2)$ hingga $O(n^3)$ terhadap jumlah data latih, pelatihan model final SVR dilakukan menggunakan sampel acak representatif sebanyak 8.000 dari 48.424 data latih — strategi yang lazim diterapkan pada penelitian dengan algoritma kernel nonlinear berskala besar — sementara evaluasi tetap dilakukan pada keseluruhan 12.106 data uji. Model Linear Regression sebagai baseline dilatih menggunakan seluruh 48.424 data latih.

Peringkat	Kernel	C	Gamma	Rata-rata R^2 (CV)
1	RBF	500	0,5	0,605
2	RBF	300	0,5	0,568
3	RBF	500	1	0,545
4	RBF	300	1	0,529
5	RBF	100	0,5	0,434

Tabel 4. *Lima Kombinasi Hyperparameter Terbaik Hasil Grid Search*

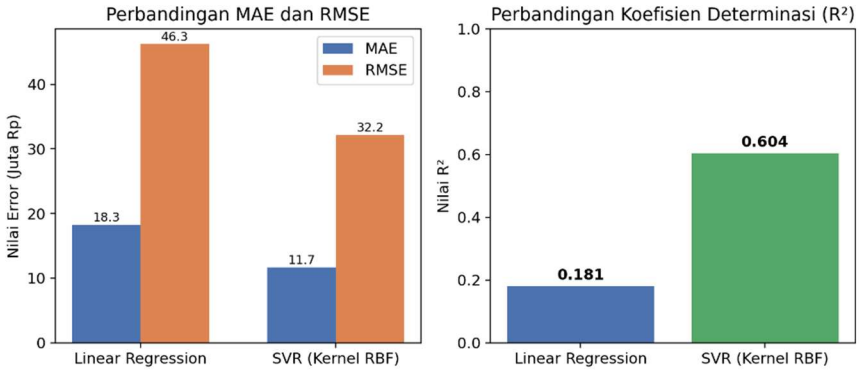
3.3 Perbandingan Kinerja Model

Perbandingan kinerja akhir kedua model pada data uji disajikan pada Tabel 5 dan Gambar 2. Model SVR kernel RBF menunjukkan performa yang jauh lebih baik dibandingkan Linear Regression pada seluruh metrik evaluasi: nilai MAE SVR (Rp11.667.484) lebih rendah 36,2% dibandingkan Linear Regression (Rp18.290.123); RMSE SVR (Rp32.183.966) lebih rendah 30,4% dibandingkan Linear Regression (Rp46.258.319); dan R^2 meningkat dari 0,181 menjadi 0,604, atau lebih dari tiga kali lipat kemampuan menjelaskan variasi harga. Perbandingan median absolute error turut memperkuat temuan ini: median kesalahan absolut SVR (Rp4.751.914) jauh lebih rendah dibandingkan Linear Regression (Rp10.896.801), atau berkurang sekitar 56%, yang mengindikasikan bahwa untuk sebagian besar kartu (kasus tipikal), presisi prediksi SVR jauh lebih baik daripada yang tercermin dari MAE keseluruhan yang cenderung dipengaruhi sejumlah kecil kartu bernilai ekstrem pada kedua model.

Keunggulan SVR ini secara teoretis dijelaskan melalui prinsip kernel trick, yaitu memetakan data masukan ke ruang fitur berdimensi lebih tinggi tanpa perlu menghitung transformasi secara eksplisit, sehingga model mampu menemukan fungsi regresi yang lebih fleksibel dan mengikuti kelengkungan hubungan antara fitur ekosistem pasar dan harga kartu. Sebaliknya, Linear Regression membatasi ruang hipotesis pada fungsi linear, sehingga ketika pola data sesungguhnya bersifat nonlinear model mengalami underfitting (bias tinggi) dan tidak mampu menangkap dinamika harga secara memadai, sebagaimana tercermin dari R^2-nya yang hanya 0,181.

Model	MAE (Rp)	RMSE (Rp)	R^2
Linear Regression	18.290.123	46.258.319	0,181
SVR (Kernel RBF)	11.667.484	32.183.966	0,604

Tabel 5. *Perbandingan Kinerja Model pada Data Uji*



Gambar 2. *Perbandingan Metrik Evaluasi Kedua Model*

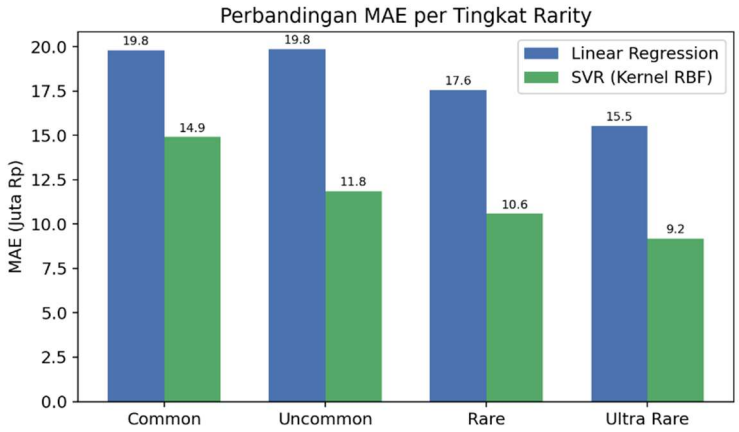
3.4 Kinerja Model Berdasarkan Tingkat Rarity

Untuk memahami segmen data mana yang paling optimal diprediksi, kinerja model dievaluasi secara terpisah berdasarkan tingkat rarity (Tabel 6 dan Gambar 3). Performa SVR pada kartu Common tercatat paling lemah (R^2=0,388), sementara performa terbaik dicapai pada kartu Ultra Rare (R^2=0,736). Peningkatan performa antartingkat rarity tidak sepenuhnya monoton — R^2 pada kartu Uncommon (0,704) sedikit lebih tinggi dibandingkan Rare (0,661) — sehingga pola yang lebih tepat digambarkan adalah performa SVR meningkat tajam begitu rarity naik dari Common ke tingkat di atasnya.

Hal ini mengindikasikan bahwa harga kartu langka mengikuti pola pertumbuhan harga yang lebih sistematis dan dapat diprediksi, kemungkinan karena permintaan pasar terhadap kartu langka lebih didorong faktor kelangkaan itu sendiri dan lebih sedikit dipengaruhi spekulasi jangka pendek, sedangkan harga kartu Common lebih fluktuatif dan dipengaruhi faktor di luar variabel dataset (mis. sentimen komunitas atau tren meta permainan). Meskipun demikian, pada seluruh segmen rarity, SVR tetap konsisten mengungguli Linear Regression dengan margin yang besar.

Rarity	n (uji)	MAE SVR (Rp)	MAE LR (Rp)	R^2 SVR	R^2 LR
Common	2.905	14.903.837	19.799.293	0,388	0,211
Uncommon	3.417	11.845.012	19.847.803	0,704	0,142
Rare	3.098	10.578.985	17.555.852	0,661	0,172
Ultra Rare	2.686	9.196.877	15.523.164	0,736	0,248

Tabel 6. *Kinerja Model per Tingkat Rarity pada Data Uji*



Gambar 3. *Perbandingan MAE per Tingkat Rarity*

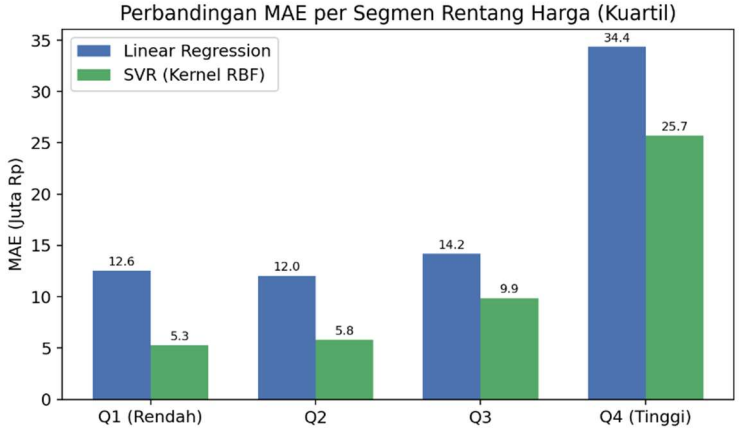
3.5 Kinerja Model Berdasarkan Segmen Rentang Harga

Evaluasi lanjutan dilakukan dengan mengelompokkan data uji ke dalam empat segmen rentang harga berdasarkan kuartil (Tabel 7 dan Gambar 4). SVR secara konsisten menghasilkan MAE yang jauh lebih rendah dibandingkan Linear Regression pada keempat segmen, dengan perbaikan MAE terbesar pada segmen harga rendah (Q1), yaitu

berkurang hingga 58,0%. Secara absolut, nilai MAE meningkat seiring naiknya segmen harga — wajar karena kartu bernilai tinggi (Q4) memiliki rentang nilai yang jauh lebih lebar. Perlu dicatat bahwa nilai R^2 yang dihitung pada masing-masing segmen kuartil justru bernilai negatif dan sangat rendah pada segmen Q1–Q3, meskipun MAE pada segmen tersebut relatif kecil. Fenomena ini merupakan artefak statistik yang umum terjadi ketika R^2 dihitung pada subset data dengan rentang nilai yang dipersempit (restricted range): variasi harga aktual dalam satu segmen kuartil sangat kecil sehingga penyebut pada formula R^2 menjadi sangat kecil, membuat R^2 sangat sensitif dan dapat bernilai sangat negatif walaupun kesalahan absolut prediksi tetap kecil. Oleh karena itu, MAE merupakan metrik yang lebih tepat untuk membandingkan kinerja model antarsegmen harga, sedangkan R^2 lebih tepat diinterpretasikan pada keseluruhan rentang data.

Segmen	**Rerata Harga Aktual (Rp)**	**MAE SVR (Rp)**	**MAE LR (Rp)**	**Selisih MAE**
Q1 (Rendah)	1.813.447	5.276.788	12.555.761	-58,0%
Q2	5.327.021	5.805.314	12.037.480	-51,8%
Q3	12.032.379	9.866.518	14.200.158	-30,5%
Q4 (Tinggi)	63.553.661	25.718.783	34.363.645	-25,2%

Tabel 7. *Perbandingan MAE per Segmen Rentang Harga (Kuartil)*



Gambar 4. *Perbandingan MAE per Segmen Rentang Harga*

3.6 Pembahasan

Sebagaimana disinggung pada bagian hasil pemodelan, model SVR menunjukkan R^2 sebesar 0,970 pada data latih (sampel 8.000 data) tetapi menurun menjadi 0,604 pada data uji, atau terdapat kesenjangan generalisasi (generalization gap) sebesar 0,366. Dalam kerangka bias-variance tradeoff, kesenjangan ini mengindikasikan model berada pada kondisi variance yang relatif tinggi — cukup fleksibel menyesuaikan diri terhadap data latih, namun kemampuan generalisasinya terhadap data baru menjadi terbatas. Kondisi ini konsisten dengan temuan bahwa sekitar 53% data latih menjadi support vector pada hyperparameter terbaik (C=500). Beberapa faktor yang diduga berkontribusi terhadap kesenjangan ini antara lain ukuran sampel pelatihan akhir yang relatif terbatas (8.000 dari 48.424 data) akibat kompleksitas komputasi kernel RBF, sebaran harga yang sangat lebar dan menceng dengan outlier bernilai ekstrem, serta representasi fitur yang belum sepenuhnya menangkap seluruh faktor harga di pasar riil. Meskipun demikian, performa SVR pada data uji tetap jauh lebih unggul dibandingkan Linear Regression pada seluruh metrik dan segmen evaluasi, sehingga kesenjangan generalisasi ini tidak mengurangi kesimpulan utama penelitian.

Temuan bahwa fitur rarity dan usia kartu berkontribusi signifikan terhadap harga sejalan dengan prinsip dasar ekonomi barang koleksi (collectible economics), yang menyatakan nilai suatu aset koleksi banyak dipengaruhi oleh tingkat kelangkaan (scarcity) dan faktor waktu (aging/maturity value), sebagaimana umum ditemukan pada studi harga kartu olahraga, perangko, maupun barang seni langka. Keunggulan SVR kernel RBF dibandingkan model linear pada data dengan hubungan nonlinear juga konsisten dengan penelitian terdahulu pada domain prediksi harga aset dan komoditas (Pangestu et al., 2024; Gananta et al., 2023; Yang et al., 2022), yang umumnya melaporkan bahwa algoritma berbasis kernel atau ensemble nonlinear cenderung mengungguli regresi linear ketika data memiliki pola musiman, siklus, atau pertumbuhan nonproporsional terhadap waktu.

Secara praktis, model SVR yang dihasilkan paling andal digunakan sebagai alat bantu estimasi harga pada kartu dengan rarity tinggi (Rare dan Ultra Rare) serta pada rentang harga menengah ke bawah, di mana tingkat kesalahan prediksi relatif kecil dan

proporsional terhadap nilai kartu. Bagi kolektor, penjual, maupun platform jual-beli kartu, model ini berpotensi dimanfaatkan sebagai referensi awal (baseline estimation) dalam menentukan harga wajar, membantu mendeteksi harga yang jauh di luar kewajaran, serta mendukung transparansi harga di pasar yang sebelumnya sangat bergantung pada negosiasi individual. Untuk kartu bernilai sangat tinggi (segmen Q4) atau kartu Common dengan volatilitas tinggi, hasil prediksi sebaiknya digunakan sebagai pertimbangan pendukung, bukan acuan tunggal.

Penelitian ini memiliki sejumlah keterbatasan: (1) pelatihan akhir SVR menggunakan sampel representatif 8.000 dari 48.424 data latih akibat kompleksitas komputasi kernel RBF; (2) dataset hanya mencakup fitur rarity, kondisi, usia kartu, nama karakter, dan set penerbitan, tanpa variabel eksternal seperti volume transaksi, sentimen komunitas, atau kondisi makroekonomi; (3) evaluasi menggunakan skema pembagian data acak (random split), bukan time-series split berbasis kronologi waktu, sehingga idealnya perlu diuji lebih lanjut dengan skema berbasis waktu untuk mengukur kemampuan prediksi di masa depan secara lebih realistis; dan (4) interpretasi R^2 pada subset data bersegmen sempit memiliki keterbatasan sebagaimana dibahas sebelumnya. Pengembangan lanjutan disarankan mencakup penerapan teknik aproksimasi kernel (Nyström method, Random Fourier Features) atau pustaka SVR yang dioptimalkan untuk data besar (mis. ThunderSVM dengan akselerasi GPU), penambahan variabel eksternal ekosistem pasar, penerapan skema validasi berbasis waktu (time-series cross-validation), perbandingan dengan algoritma nonlinear lain (Random Forest, Gradient Boosting, LSTM), serta pengembangan model bertingkat (hierarchical/segmented model) per tingkat rarity atau segmen harga.

4. KESIMPULAN DAN SARAN

Fitur ekosistem pasar lokal yang digunakan — usia kartu sejak rilis, tingkat rarity, kondisi fisik, nama kartu, dan set penerbitan — terbukti berkontribusi terhadap variasi harga kartu game, dengan hubungan harga terhadap usia kartu yang bersifat nonlinear dan harga rata-rata yang meningkat sejalan dengan tingkat kelangkaan. Model Linear Regression sebagai baseline hanya mampu menjelaskan variasi harga secara terbatas (R^2=0,181; MAE=Rp18.290.123; RMSE=Rp46.258.319), menegaskan bahwa pendekatan linear kurang memadai untuk pola harga yang nonlinear. Model SVR kernel

RBF dengan hyperparameter optimal (C=500, γ=0,5) hasil GridSearchCV dan cross-validation memberikan performa jauh lebih baik (R^2=0,604; MAE=Rp11.667.484; RMSE=Rp32.183.966), atau penurunan MAE 36,2% dan RMSE 30,4% serta peningkatan R^2 lebih dari tiga kali lipat dibandingkan Linear Regression. Kinerja SVR tidak seragam pada seluruh segmen: performa terbaik pada kartu Ultra Rare (R^2=0,736) dan terendah pada kartu Common (R^2=0,388), yang mengindikasikan bahwa harga kartu langka mengikuti pola yang lebih sistematis dan dapat diprediksi. Dengan demikian, optimasi prediksi harga kartu game menggunakan SVR berbasis kernel nonlinear terbukti signifikan lebih unggul dibandingkan Linear Regression dan dapat dimanfaatkan sebagai alat bantu estimasi harga awal bagi kolektor, penjual, maupun platform jual-beli kartu di ekosistem pasar lokal. Penelitian selanjutnya disarankan menerapkan teknik aproksimasi kernel untuk data berskala besar, menambahkan variabel eksternal ekosistem pasar, menerapkan skema validasi berbasis waktu, membandingkan dengan algoritma nonlinear lain, serta mengembangkan model bertingkat per segmen rarity atau harga guna meningkatkan akurasi dan generalisasi model.

UCAPAN TERIMA KASIH

Artikel ini merupakan bagian dari hasil skripsi penulis pertama pada Program Studi Teknologi Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, di bawah bimbingan penulis kedua.

DAFTAR REFERENSI

Aldiansyah, M., Muzakir, M., Syarifuddin, U., & Lamusa, F. (2025). Brand awareness dan lifestyle pada purchase decision kartu Pokémon melalui customer satisfaction. Studi Akuntansi, Keuangan, Dan Manajemen, 5(2), 571–585. https://doi.org/10.35912/sakman.v5i2.5494

Alon, I., Bretas, V. P. G., & Katrih, V. (2023). Predictors of NFT prices: An automated machine learning approach. Journal of Global Information Management, 31(1), 1–18. https://doi.org/10.4018/JGIM.317097

Asyraq, C., & Hartati, E. (2025). Perangkat lunak untuk memprediksi harga cryptocurrency menggunakan algoritma Support Vector Regression. Jurnal Teknik Informatika, 5(1).

Aulia, A., Aprianti, B., Supriyanto, Y., & Rozikin, C. (2022). Prediksi harga emas dengan menggunakan algoritma SVR dan Linear Regression. Jurnal Ilmiah Wahana Pendidikan, 8(5). https://doi.org/10.5281/zenodo.6408864

Devanda, M., Kusuma, H., & Hidayat, S. (2024). Penerapan model regresi linier dalam prediksi harga mobil bekas di India dan visualisasi dengan menggunakan Power BI. Jurnal Ilmiah, 5(2), 1097–1110.

Du, K., Jiang, B., Lu, J., & Hua, J. (2024). Exploring kernel machines and support vector machines: Principles, techniques, and future directions. arXiv preprint, 1–58.

Gananta, I. M., Purnama, I. N., & Fredlina, K. Q. (2023). Optimasi prediksi harga emas dengan metode Support Vector Regression (SVR) menggunakan algoritma Grid Search. Jurnal Ilmu Komputer, 7(6), 3160–3165.

Hidayatulloh, M. R., & Yuwono, D. P. (2025). Implementasi algoritma Support Vector Regression untuk prediksi harga emas berdasarkan data historis. Jurnal Informatika: Jurnal Pengembangan IT, 10(4), 1106–1114. https://doi.org/10.30591/jpit.v10i4.9233

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. Geoscientific Model Development, 15(14), 5481–5487.

Jiang, H. (2025). Machine learning models for predicting second-hand house prices: A comparative study. Proceedings of the 2025 International Conference on Big Data, Artificial Intelligence and Digital Economy, 1(1), 99–106. https://doi.org/10.1145/3767052.3767068

Mekacher, A., Bracci, A., Nadini, M., Martino, M., Alessandretti, L., Aiello, L. M., & Baronchelli, A. (2022). Heterogeneous rarity patterns drive price dynamics in NFT collections. Scientific Reports, 12, 1–9. https://doi.org/10.1038/s41598-022-17922-5

Pangestu, R. A., Vitianingsih, A. V., Kacung, S., Maukar, A. L., & Noertjahyana, A. (2024). Comparative analysis of Support Vector Regression and Linear Regression models to predict Apple Inc. share prices. Indonesian Journal of Artificial Intelligence and Data Mining, 7(1), 148. https://doi.org/10.24014/ijaidm.v7i1.28594

Shobayo, O., Adeyemi-Longe, S., Popoola, O., & Okoyeigbo, O. (2025). A comparative analysis of machine learning and deep learning techniques for accurate market price forecasting. Analytics, 4(1). https://doi.org/10.3390/analytics4010005

Sonkavde, G., Dharrao, D. S., Bongale, A. M., & Deokate, S. T. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. International Journal of Financial Studies, 11(3), 94.

Urrochman, M. Y., & Hizham, F. A. (2025). Performance comparison of Random Forest Regression and SVR in stock price prediction. Jurnal Pilar Nusa Mandiri, 21(1), 16–23. https://doi.org/10.33480/pilar.v21i1.6072

Widiyatmoko, A. T., Butsianto, S., & Nugroho, A. (2025). Application of machine learning for premium rice price increase prediction using Linear Regression algorithm. Jurnal Teknologi Informasi, 5(2), 1125–1132.

Wood, D., Sufan, U., Villamil, C., & Wood, D. R. (2024). Neural networks for automated valuation prediction for collectible cards. International Journal of Scientific Research and Management, 12(10), 1496–1510. https://doi.org/10.18535/ijsrm/v12i10.ec04

Yang, X., Mi, J. L., Jiang, J., Li, J. W., Zhang, Q. S., & Geng, M. M. (2022). Carbon sink price prediction based on radial basis kernel function support vector machine regression model. International Journal of Low-Carbon Technologies, 17, 1075–1084. https://doi.org/10.1093/ijlct/ctac088