

Sistem Informasi PPDB Berbasis Web untuk Prediksi Jumlah Pendaftar Menggunakan Regresi Linier

Web-Based New Student Admission Information System for Applicant Prediction Using Linear Regression

Nasrul Abdul Karim

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan,
15417, Indonesia

¹nazrudian13@gmail.com

Received: July 14, 2026 | Revised: August 18, 2026 | Accepted: September 1, 2026 | Published: September 16, 2026



COSIE is licensed a Creative Commons Attribution-Share Alike 4.0 International License

Abstrak

Proses Penerimaan Peserta Didik Baru (PPDB) di SMA Islam Hidayatul Ahkam saat ini masih dikerjakan secara manual. Kondisi ini membuat pemeriksaan berkas memakan waktu lama, membuka peluang kekeliruan saat memasukkan data, serta menyulitkan pihak sekolah dalam memperkirakan berapa banyak calon siswa yang akan mendaftar pada periode berikutnya untuk keperluan perencanaan kapasitas kelas. Riset ini bertujuan merancang dan membangun sebuah sistem informasi PPDB berbasis web yang dilengkapi modul prediksi menggunakan Regresi Linier Sederhana, dengan SMA Islam Hidayatul Ahkam sebagai lokasi studi kasus. Pembangunan sistem mengikuti tahapan Rapid Application Development (RAD) dan diimplementasikan memakai PHP native, PDO, serta basis data MySQL. Data jumlah pendaftar selama empat tahun ajaran terakhir diolah dengan pendekatan kuadrat terkecil (least squares), sehingga diperoleh persamaan regresi $\hat{Y} = 36 + 10,9X$. Berdasarkan persamaan itu, sistem memproyeksikan jumlah pendaftar pada tahun ajaran 2026/2027 mencapai 91 siswa, dengan rentang operasional 77 sampai 105 (margin $\pm 15\%$, bukan interval statistik formal), sementara nilai koefisien determinasi (R^2) sebesar 47,65% (in-sample) dan MAPE sebesar 15,14%. Serangkaian pengujian black box dan white box membuktikan seluruh fungsi sistem berjalan sebagaimana mestinya dan hasil perhitungan otomatis sistem sama persis dengan hasil perhitungan manual, yang menegaskan konsistensi implementasi perangkat lunak, bukan validitas statistik model. Temuan ini memperlihatkan bahwa sistem PPDB berbasis web yang dibangun tidak sekadar berperan sebagai perangkat administrasi, tetapi juga sebagai purwarupa alat bantu analisis prediktif yang masih memerlukan validasi lebih lanjut, untuk mendukung perencanaan sekaligus pengambilan keputusan penerimaan siswa baru di SMA Islam Hidayatul Ahkam.

Kata kunci: Sistem Informasi PPDB; Regresi Linier Sederhana; Prediksi; Rapid Application Development; Website

Abstract

At present, SMA Islam Hidayatul Ahkam still handles its New Student Admission (Penerimaan Peserta Didik Baru/PPDB) procedure by hand. Because of this, checking submitted documents takes a long time, data entry mistakes are common, and school staff struggle to estimate how many applicants to expect in the following intake for capacity planning purposes. This study sets out to design and build a web-based PPDB information system equipped with a Simple Linear Regression prediction module, taking SMA Islam Hidayatul Ahkam as the case study. The system was constructed following the Rapid Application Development (RAD) approach and implemented with native PHP, PDO, and a MySQL database. Four years of historical applicant records were processed using the least squares technique, yielding the regression equation $\hat{Y} = 36 + 10.9X$. From this equation, the system forecasts 91 applicants for the 2026/2027 academic year, with an operational range of 77 to 105 (a $\pm 15\%$ margin, not a formal statistical interval), a Mean Absolute Percentage Error (MAPE) of 15.14%, and an R^2 value of 47.65% (in-sample). Black box and white box testing confirmed that every system function operated correctly and that the automatically generated calculations matched the manual ones exactly, confirming software implementation consistency rather than the statistical validity of the forecast. These findings indicate that the resulting web-based PPDB system serves not merely as an administrative tool but also as a prototype

predictive-analysis aid that still requires further validation, for planning and decision-making in new student admissions at SMA Islam Hidayatul Ahkam.

Keywords: PPDB Information System; Simple Linear Regression; Prediction; Rapid Application Development; Website

1. PENDAHULUAN

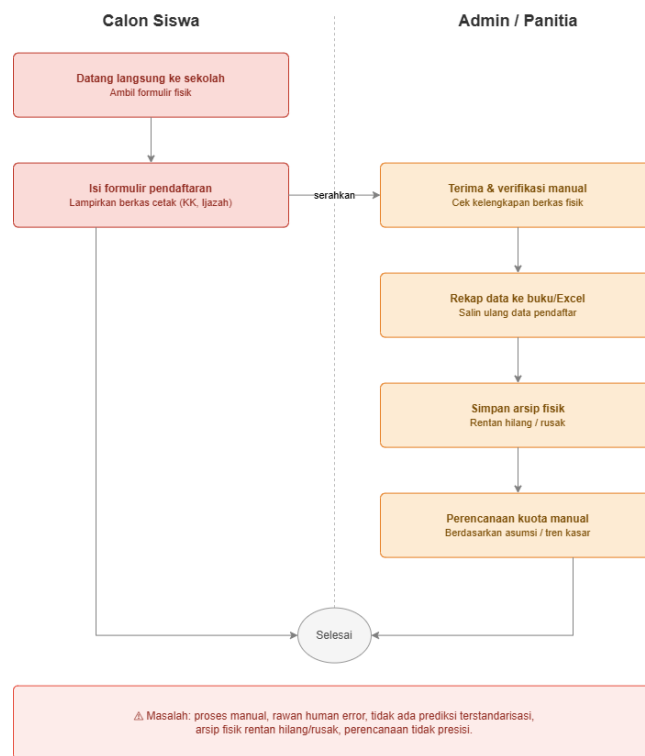
Pengelolaan pendidikan di Indonesia kini banyak bertumpu pada digitalisasi, mulai dari tata kelola administrasi sekolah [1] hingga kemudahan akses layanan penerimaan siswa baru [2]. Data Badan Pusat Statistik mencatat jumlah siswa pada tahun ajaran 2024/2025 sebanyak 52.913.427 orang, turun 0,56% dari tahun ajaran sebelumnya [3]. Angka yang naik-turun seperti ini mengharuskan setiap satuan pendidikan mengantisipasinya lebih awal. Di lapangan, kebanyakan sekolah swasta masih mengandalkan buku catatan dan berkas fisik untuk memproses penerimaan siswa baru, yang membuat pengolahan datanya lambat dan rawan keliru [4]. Kondisi ini persis yang terjadi di SMA Islam Hidayatul Ahkam. Proses PPDB di sekolah tersebut masih dijalankan secara manual: formulir dicetak dan dikumpulkan satu per satu, dokumen diperiksa secara fisik, lalu datanya disalin ulang ke rekapitulasi. Alur kerja seperti ini membuat verifikasi mudah tertunda, rawan salah input, dan menyulitkan sekolah memperkirakan jumlah pendaftar pada periode berikutnya. Sejumlah penelitian sudah mencoba menjawab persoalan ini. Regresi Linier, misalnya, dipakai pada sistem prediksi berbasis web untuk penerimaan siswa baru dan terbukti membantu bagian administrasi bekerja lebih efisien karena tersedia perkiraan jumlah pendaftar yang cukup akurat [5]. Penerapan pada salah satu SMK bahkan mencatat tingkat kesalahan (Mean Absolute Percentage Error/MAPE) di bawah 10% [6], angka yang tergolong tinggi akurasi dan membuka peluang metode ini dipakai di sekolah lain. Meskipun begitu, masih ada masalah yang belum tuntas dijawab. Sebagian besar sistem PPDB berbasis web pada penelitian terdahulu hanya menangani sisi administrasi pendaftaran, tanpa modul analisis prediktif yang bisa menopang strategi manajemen sekolah [7]. Sementara itu, riset-riset yang membahas model prediksi jumlah pendaftar umumnya berdiri sendiri dan belum menyatu ke dalam satu sistem informasi PPDB yang siap dipakai sekolah [5]. Dengan kata lain, celah yang ingin ditutup penelitian ini bukan soal metode, melainkan soal integrasi: belum ada sistem PPDB yang menggabungkan fungsi administrasi dengan modul prediksi dalam satu platform. Regresi Linier Sederhana dipilih di sini bukan karena diklaim paling akurat untuk data yang berfluktuasi, tetapi karena perhitungannya sederhana, mudah dipahami pengguna non-teknis seperti kepala sekolah dan staf tata usaha, serta cocok diterapkan pada data historis yang masih sangat terbatas, yakni hanya empat tahun. Karena itu, model ini diperlakukan sebagai estimasi operasional tahap awal (proof-of-concept), bukan model prediksi definitif; data historis pada Tabel 1 (45, 50, 90, 68) sendiri menunjukkan pola yang tidak sepenuhnya linier. Berangkat dari celah itulah penelitian ini membangun sistem informasi PPDB berbasis web dengan modul prediksi Regresi Sederhana tertanam di dalamnya, mengambil SMA Islam Hidayatul Ahkam, sekolah swasta dengan data historis terbatas, hanya empat tahun sebagai objek kajian.

Ada tiga tujuan yang hendak dicapai penelitian ini: (1) menganalisis dan merancang sistem informasi PPDB berbasis web yang mempercepat, mengakuratkan, dan memperjelas proses pendaftaran maupun verifikasi berkas; (2) menerapkan Regresi Linier Sederhana untuk memproyeksikan jumlah calon pendaftar berdasarkan data historis; dan (3) menguji kelayakan fungsional sistem melalui black box testing sekaligus mengukur akurasi model prediksinya dengan metrik MAPE dan R^2 . Kontribusi penelitian ini bersifat integratif dan berbasis studi kasus, bukan menawarkan metode prediksi baru: modul prediksi regresi linier disatukan ke dalam satu sistem informasi PPDB yang diterapkan langsung pada kondisi nyata sekolah dengan data historis terbatas, sehingga sekolah dapat memakainya baik sebagai perangkat administrasi maupun sebagai sarana analisis prediktif awal berbasis data.

2. BAHAN DAN METODE

2.1. Desain Penelitian

Penelitian ini menempuh pendekatan applied research di bidang rekayasa perangkat lunak, yang memadukan analisis kebutuhan sistem informasi dengan pemodelan prediktif berbasis statistik. Pembangunan sistem mengacu pada metodologi Rapid Application Development (RAD), dipilih karena kebutuhan pengguna cenderung berubah-ubah, menuntut proses iterasi yang cepat berdasarkan masukan pengguna, sekaligus mempertimbangkan batas waktu penyusunan skripsi yang relatif singkat (± 6 bulan). Empat tahapan RAD diterapkan dalam penelitian ini, yaitu (1) perencanaan kebutuhan, (2) perancangan bersama pengguna (user design), (3) pembangunan sistem (construction), dan (4) tahap peralihan (cutover); pada setiap tahapan, masukan dari kepala sekolah dan staf tata usaha senantiasa dilibatkan.



Gambar 1. Analisis Alur Sistem PPDB yang Sedang Berjalan (Manual)

2.2. Dataset/Partisipan

Penelitian ini memakai dua kategori data. Kategori pertama adalah data kualitatif, yang dihimpun melalui pengamatan langsung terhadap jalannya proses PPDB serta wawancara bersama Kepala Sekolah dan Staf Tata Usaha SMA Islam Hidayatul Ahkam guna memetakan kendala yang dihadapi dan kebutuhan sistem yang diharapkan. Kategori kedua berupa data kuantitatif, yaitu rekapitulasi jumlah pendaftar PPDB sepanjang empat tahun ajaran terakhir (2022/2023 hingga 2025/2026), bersumber dari arsip bagian tata usaha sekolah dan ditampilkan pada Tabel 1. Istilah “pendaftar” dalam penelitian ini merujuk pada jumlah calon siswa unik yang berkasnya telah diverifikasi lengkap oleh pihak sekolah pada masing-masing tahun ajaran, bukan jumlah pengajuan formulir mentah yang mungkin mencakup duplikasi maupun berkas yang belum lengkap.

Tabel 1. Data Historis Jumlah Pendaftar SMA Islam Hidayatul Ahkam

X (Urutan)	Tahun Ajaran	Y (Jumlah Pendaftar)
1	2022/2023	45
2	2023/2024	50
3	2024/2025	90
4	2025/2026	68

2.3. Data Preprocessing

Sebelum dimodelkan, data historis melalui tiga tahap prapemrosesan. Pertama, pembersihan data (data cleaning) untuk mengidentifikasi dan mengeliminasi data yang tidak valid, duplikat, atau tidak lengkap; proses ini tidak mengubah satu pun catatan pada keempat tahun ajaran yang dianalisis karena seluruh data arsip tata usaha yang tersedia sudah lengkap dan tidak terdapat duplikasi. Kedua, transformasi data, yaitu konversi tahun akademik aktual menjadi indeks periode numerik ($X = 1, 2, 3, 4$) untuk menghindari inflasi nilai akibat pengkuadratan angka tahun pada rumus least squares. Ketiga, verifikasi kelengkapan pasangan data X dan Y sebelum digunakan pada tahap pemodelan regresi. Perlu ditegaskan bahwa faktor-faktor eksternal seperti kebijakan zonasi, promosi sekolah, atau kondisi ekonomi tidak termasuk dalam variabel yang diukur pada tahap prapemrosesan ini; faktor-faktor tersebut hanya dibahas sebagai kemungkinan penjelas (hipotesis) pada bagian Hasil dan Pembahasan, bukan sebagai temuan yang terverifikasi secara empiris.

2.4. Model yang Diusulkan

Model prediksi yang diajukan berupa Regresi Linier Sederhana, yang menggambarkan keterkaitan antara variabel bebas X (indeks tahun ajaran) dengan variabel terikat Y (jumlah pendaftar) melalui persamaan garis lurus berikut:

$$\hat{Y} = a + bX \quad (1)$$

dengan $\hat{Y}$ menyatakan nilai prediksi jumlah pendaftar, a merupakan konstanta (intercept), sedangkan b adalah koefisien regresi (slope). Kedua nilai tersebut dihitung memakai metode Ordinary Least Squares (OLS):

$$b = (n\sum XY - \sum X \sum Y) / (n\sum X^2 - (\sum X)^2) \quad (2)$$

$$a = (\sum Y - b\sum X) / n \quad (3)$$

dengan n menyatakan banyaknya data historis, $\sum Y$ jumlah keseluruhan nilai Y , $\sum X$ jumlah keseluruhan nilai X , $\sum XY$ hasil penjumlahan perkalian X dan Y , serta $\sum X^2$ hasil penjumlahan kuadrat X . Pada penerapannya, administrator memasukkan jumlah pendaftar tiap tahun ajaran ke tabel data_historis dalam basis data, kemudian sistem secara otomatis menghitung nilai a dan b sekaligus memproyeksikan jumlah pendaftar untuk tahun ajaran berikutnya ($X = n+1$), yang hasilnya tersimpan pada tabel hasil_prediksi.

2.5. Pengaturan Eksperimen

Pembangunan sistem PPDB memakai bahasa pemrograman PHP native yang terhubung ke basis data MySQL melalui PDO, dipadukan dengan Tailwind CSS/Bootstrap 5 untuk sisi tampilan dan Alpine.js untuk menangani interaksi di sisi klien. Proses pengembangan dilakukan di lingkungan WampServer sebagai server lokal, berjalan pada sistem operasi Windows 11 (64-bit) dengan prosesor Intel Core i7 generasi ke-11, RAM 16 GB, dan media penyimpanan SSD berkapasitas 1 TB. Rancangan sistem bersifat responsif, memungkinkan akses lewat ponsel pintar, laptop, ataupun komputer dengan resolusi layar minimal 360 x 640 piksel serta kecepatan internet setidaknya 1 Mbps. Terdapat dua sisi akses pada sistem ini: sisi calon siswa yang terbuka untuk umum, dan sisi admin yang mensyaratkan proses autentikasi untuk mengelola data pendaftar, data historis, serta modul Prediksi & Strategi.

2.6. Metrik Evaluasi

Tingkat akurasi model prediksi diukur memakai dua metrik evaluasi. Metrik pertama, Mean Absolute Percentage Error (MAPE), dipakai untuk menghitung rata-rata persentase selisih antara nilai prediksi dan nilai aktual:

$$MAPE = (1/n) \sum |(Y_{aktual} - \hat{Y}_{prediksi}) / Y_{aktual}| \times 100\% \quad (4)$$

dengan patokan interpretasi sebagai berikut: nilai di bawah 10% tergolong sangat akurat, 10%-20% tergolong baik, 20%-50% tergolong cukup layak, sementara di atas 50% dianggap tidak akurat. Metrik kedua, Koefisien Determinasi (R^2), dipakai untuk melihat seberapa besar proporsi variasi data yang mampu dijelaskan oleh model, dengan rentang nilai antara 0 hingga 1:

$$R^2 = 1 - (\sum(Y_{aktual} - \hat{Y}_{prediksi})^2) / (\sum(Y_{aktual} - \bar{Y})^2) \quad (5)$$

dengan $\bar{Y}$ adalah rata-rata seluruh nilai Y . Kedua metrik ini dihitung secara in-sample, yaitu menggunakan keempat titik data historis yang sama dengan yang dipakai untuk membentuk model, karena keterbatasan jumlah observasi belum memungkinkan pembagian data menjadi subset pelatihan dan pengujian (train-test split) yang terpisah; dengan demikian, nilai MAPE dan R^2 yang dilaporkan mencerminkan kesesuaian model terhadap data historis yang ada, bukan kinerja pada data baru yang belum pernah dilihat model (out-of-sample). Selain kedua metrik tersebut, penelitian ini turut menghitung Compound Annual Growth Rate (CAGR) dan Year-over-Year growth (YoY) sebagai indikator tambahan tren pertumbuhan pendaftar. Sebagai pembeda yang perlu ditegaskan, black box testing (pengujian fungsional dari sisi input-output) dan white box testing (pengujian jalur logika internal algoritma) pada penelitian ini digunakan untuk memvalidasi kebenaran implementasi perangkat lunak (software correctness), bukan untuk memvalidasi keandalan statistik model prediksi.

3. HASIL

Berdasarkan data historis pada Tabel 1 ($n = 4$, $\sum X = 10$, $\sum Y = 253$, $\sum X^2 = 30$, $\sum XY = 687$), perhitungan koefisien regresi menghasilkan $b = 10,9$ dan $a = 36$, sehingga diperoleh persamaan regresi:

$$\hat{Y} = 36 + 10,9X \quad (6)$$

Nilai slope sebesar 10,9 menunjukkan bahwa setiap kenaikan satu tahun ajaran, jumlah pendaftar rata-rata meningkat sebesar 10,9 siswa. Menggunakan persamaan tersebut, sistem memprediksi jumlah pendaftar tahun ajaran 2026/2027 ($X = 5$) sebesar $90,5 \approx 91$ pendaftar, dengan rentang toleransi $\pm 15\%$ (77-105 pendaftar). Untuk memungkinkan audit terhadap nilai MAPE dan R^2 yang dilaporkan, Tabel 2

menyajikan nilai terpasang (fitted value) hasil persamaan regresi beserta residual dan Absolute Percentage Error (APE) untuk setiap tahun pengamatan. Seluruh nilai pada Tabel 2 bersifat in-sample, yakni dihitung dari data yang sama dengan yang membentuk model.

Tabel 2. Nilai Terpasang (Fitted), Residual, dan Absolute Percentage Error (APE) per Tahun

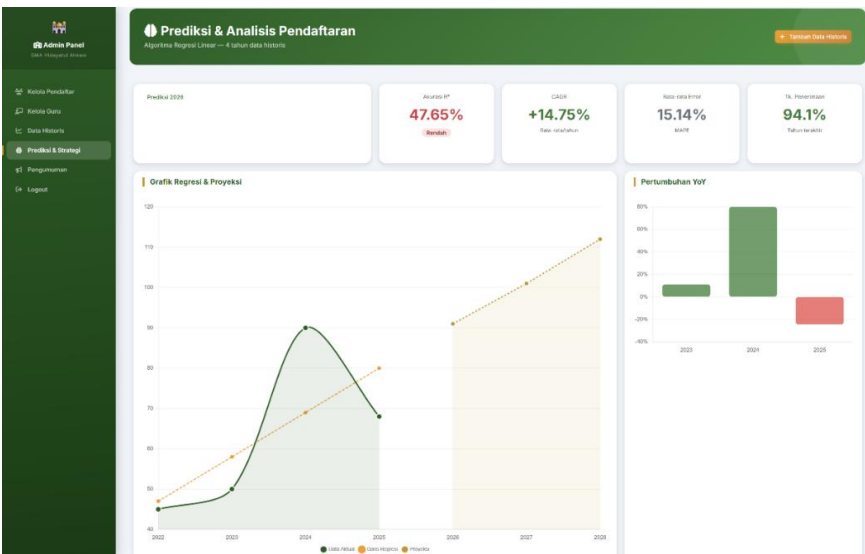
X	Tahun Ajaran	Y Aktual	$\hat{Y}$ (Fitted)	Residual (Y- $\hat{Y}$)	APE
1	2022/2023	45	46,9	-1,9	4,22%
2	2023/2024	50	57,8	-7,8	15,60%
3	2024/2025	90	68,7	21,3	23,67%
4	2025/2026	68	79,6	-11,6	17,06%

Evaluasi akurasi model menghasilkan MAPE sebesar 15,14%, yang termasuk kategori “prediksi baik” (10%-20%), dan R^2 sebesar 47,65%, yang tergolong moderat mengingat data historis yang tersedia baru mencakup empat tahun dan terdapat fluktuasi signifikan pada tahun 2024/2025. Perhitungan tambahan menunjukkan CAGR sebesar 14,75% (pertumbuhan rata-rata positif), serta YoY growth yang berfluktuasi tajam: +11,11% (2023/2024), +80,00% (2024/2025), dan -24,44% (2025/2026), sebagaimana dirangkum pada Tabel 3.

Tabel 3. Ringkasan Metrik Hasil Prediksi

Metrik	Nilai	Kategori/Interpretasi
Persamaan Regresi	$\hat{Y} = 36 + 10,9X$	Setiap tahun +10,9 pendaftar
Prediksi 2026/2027	91 pendaftar	Rentang 77-105
MAPE	15,14%	Baik (10%-20%)
R^2	47,65%	Moderat
CAGR	14,75%	Pertumbuhan positif
YoY 2025/2026	-24,44%	Penurunan signifikan

Untuk memberikan konteks pembandingan terhadap performa model regresi, penelitian ini turut menghitung baseline naif (naïve forecast) berupa prediksi satu langkah ke depan yang menyamakan nilai prediksi tahun ke-t dengan nilai aktual tahun ke-(t-1), yang dapat dihitung sepenuhnya dari data historis yang tersedia tanpa memerlukan pembagian data pelatihan-pengujian. Berdasarkan tiga titik data yang dapat dibandingkan (tahun ajaran 2023/2024 hingga 2025/2026), baseline naif menghasilkan MAPE sebesar 28,93% (APE berturut-turut 10,00%, 44,44%, dan 32,35%), sementara model regresi linier pada rentang tahun yang sama menghasilkan MAPE sebesar 18,78% (APE berturut-turut 15,60%, 23,67%, dan 17,06%, mengikuti Tabel 2). Dengan demikian, model regresi linier menunjukkan penurunan kesalahan prediksi sebesar 35,1% dibandingkan baseline naif pada data historis yang sama, yang mengindikasikan bahwa model memberikan informasi prediktif tambahan di luar sekadar mengasumsikan pengulangan nilai tahun sebelumnya, meskipun tingkat akurasi absolutnya (kategori “baik”, 10%-20%) masih terbatas.



Gambar 2. Dashboard Modul Prediksi & Strategi pada Sistem

Pengecekan keluaran sistem terhadap perhitungan manual (sebagaimana diuraikan pada bagian Metode) memperlihatkan kecocokan penuh tanpa selisih pada seluruh metrik: persamaan regresi, nilai prediksi, rentang prediksi, R^2 , MAPE, CAGR, maupun YoY setiap tahun. Hal ini membuktikan bahwa algoritma Regresi Linier telah diterjemahkan dengan benar ke dalam kode PHP dan bebas dari kesalahan kalkulasi. Uji fungsional terhadap sistem dilakukan dengan teknik black box testing pada sembilan skenario inti mencakup login admin, penambahan maupun penghapusan data historis, pendaftaran siswa, verifikasi berkas pendaftar, perubahan status, pengoperasian modul prediksi, dan logout dan seluruhnya memperoleh hasil “Valid”. Pengujian struktural memakai white box testing terhadap 15 jalur logika independen pada modul algoritma Regresi Linier, dengan kompleksitas siklomatis $V(G) = 6$, dan seluruh jalur dinyatakan “Lulus” tanpa ditemukan kesalahan logika (logic error), mencakup validasi ketersediaan data, kategorisasi tren berdasarkan slope, fungsi matematika prediksi, integrasi antarmuka, logika rekomendasi strategi, dan perhitungan metrik evaluasi. Rincian tiap skenario black box dan setiap jalur independen white box disajikan pada Tabel 4 dan Tabel 5 agar dapat ditelusuri secara transparan.

Tabel 4. Rincian Skenario Pengujian Black Box (9 Skenario)

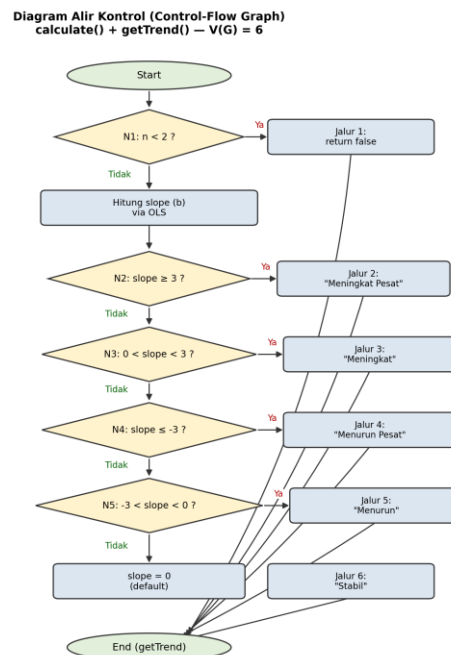
No	Skenario Uji	Input	Hasil Uji
1	Login Admin	Username dan password valid	Valid
2	Login Admin	Username/password salah	Valid
3	Tambah Data Historis	Tahun 2026, Jumlah 75, Diterima 70	Valid
4	Hapus Data Historis	Klik tombol hapus pada id tertentu	Valid
5	Pendaftaran Siswa	Formulir lengkap + unggah berkas	Valid
6	Verifikasi Pendaftar	Admin klik “Verifikasi”	Valid
7	Ubah Status Diterima	Admin klik “Terima”	Valid
8	Jalankan Prediksi	Admin klik menu Prediksi	Valid
9	Logout Admin	Klik tombol Logout	Valid

Tabel 5. Rincian Jalur Independen Pengujian White Box (15 Jalur, $V(G) = 6$)

No	Fungsi/Komponen	Kondisi/Jalur yang Diuji	Status
1	calculate()	$n < 2$ (data tidak cukup) $\rightarrow$ false	Lulus
2	getTrend()	$Slope \geq 3 \rightarrow$ “Meningkat Pesat”	Lulus
3	getTrend()	$0 < Slope < 3 \rightarrow$ “Meningkat”	Lulus
4	getTrend()	$Slope \leq -3 \rightarrow$ “Menurun Pesat”	Lulus
5	getTrend()	$-3 < Slope < 0 \rightarrow$ “Menurun”	Lulus
6	getTrend()	$Slope = 0 \rightarrow$ “Stabil”	Lulus
7	predictNextYear()	Menghasilkan prediksi_min dan prediksi_max ($\pm 15\%$)	Lulus
8	getEquation()	Menghasilkan persamaan $Y = a + bX$ dan R^2	Lulus
9	getGrowthRate()	Menghasilkan nilai CAGR	Lulus
10	getYoyGrowth()	Menghasilkan array pertumbuhan YoY	Lulus
11	Endpoint Prediksi	Validasi sesi admin sebelum akses	Lulus
12	getStrategiesAdmin()	Tren naik $\rightarrow$ “Pertahankan Momentum”	Lulus
13	getStrategiesAdmin()	Tren turun $\rightarrow$ “Pemulihan Pendaftaran”	Lulus
14	getStrategiesAdmin()	Tren stabil $\rightarrow$ “Akselerasi Pertumbuhan”	Lulus
15	Perhitungan MAPE	Menghasilkan nilai persentase rata-rata error	Lulus

Untuk mendukung secara eksplisit klaim kompleksitas siklomatis $V(G) = 6$ yang dilaporkan pada Tabel 5, Gambar 3 menyajikan diagram alir kontrol (control-flow graph) dari fungsi inti calculate() dan getTrend() yang menjadi dasar perhitungan tersebut. Diagram ini memuat lima titik keputusan (decision point) berurutan — validasi $n < 2$, serta empat ambang batas slope pada getTrend() — sehingga menghasilkan enam jalur independen (basis path) sebagaimana dirumuskan $V(G) =$ jumlah titik keputusan

+ 1 = 5 + 1 = 6, konsisten dengan Baris 1-6 pada Tabel 5. Sembilan baris sisanya pada Tabel 5 (Baris 7-15) merepresentasikan fungsi/komponen lain pada modul yang sama yang diuji secara terpisah dan tidak tercakup dalam nilai $V(G) = 6$ ini.



Gambar 3. Diagram Alir Kontrol Fungsi calculate() dan getTrend() ($V(G) = 6$)

4. PEMBAHASAN

Hasil penelitian ini memperlihatkan bahwa penyatuan modul prediksi Regresi Linier Sederhana ke dalam sistem informasi PPDB berhasil menjembatani celah yang teridentifikasi pada penelitian-penelitian sebelumnya, yakni sistem PPDB berbasis web yang umumnya hanya berfokus pada sisi administrasi tanpa modul analisis prediktif [7]. Nilai MAPE sebesar 15,14% yang diperoleh sejalan dengan temuan bahwa metode Regresi Linier layak diterapkan pada institusi pendidikan dengan data historis yang terbatas, kendati tingkat akurasi masih berada pada kategori “baik” dan belum mencapai kategori “sangat akurat” ($<10\%$) sebagaimana dilaporkan pada studi dengan volume data yang lebih besar [6]. Nilai R^2 sebesar 47,65% mengindikasikan bahwa variabel tahun ajaran hanya mampu menjelaskan kurang dari separuh variasi jumlah pendaftar, sementara sisanya (52,35%) kemungkinan dipengaruhi faktor eksternal di luar model, seperti kebijakan zonasi, promosi sekolah, maupun kondisi ekonomi. Kondisi ini cukup wajar mengingat data historis yang tersedia baru mencakup empat tahun dan memuat fluktuasi tajam pada tahun 2024/2025 (YoY +80,00%) yang diikuti penurunan pada tahun berikutnya (YoY -24,44%). Temuan ini menegaskan bahwa keterbatasan volume data historis menjadi faktor pembatas utama akurasi model prediktif pada konteks sekolah dengan data historis yang minim, sejalan dengan tantangan umum penerapan model prediktif pada institusi pendidikan skala kecil hingga menengah.

Dibandingkan dengan pendekatan administrasi manual yang sebelumnya diterapkan di SMA Islam Hidayatul Ahkam, sistem yang dikembangkan memberikan dua kontribusi utama. Pertama, digitalisasi proses pendaftaran, unggah berkas, verifikasi, dan penerbitan nomor pendaftaran secara otomatis berpotensi mengurangi risiko kesalahan pencatatan manual dan mempercepat proses administrasi; klaim ini bersifat kualitatif berdasarkan desain alur sistem, karena penelitian ini belum mengukur secara kuantitatif waktu pemrosesan maupun tingkat kesalahan sebelum dan sesudah penerapan sistem, sehingga pengujian lanjutan dengan indikator operasional terukur diperlukan untuk memastikannya. Kedua, modul prediksi memberikan estimasi kuantitatif awal yang dapat menjadi salah satu masukan bagi kepala sekolah dalam merencanakan alokasi ruang kelas dan kebutuhan verifikator, sebagai pelengkap, bukan pengganti pertimbangan berbasis pengalaman pihak sekolah, mengingat keterbatasan validasi model yang telah diuraikan. Sebagai ilustrasi konkret, estimasi 91 pendaftar untuk tahun ajaran 2026/2027 (naik dari 68

pendaftar pada tahun ajaran berjalan) mengindikasikan potensi kebutuhan tambahan ruang kelas dan personel verifikasi berkas, sehingga pihak sekolah dapat mulai menyiapkan sumber daya lebih awal; meskipun demikian, mengingat rentang toleransi 77-105 pendaftar masih cukup lebar, angka pastinya tetap perlu dikonfirmasi ulang mendekati periode pendaftaran berjalan. Capaian pengujian black box maupun white box yang sepenuhnya bernilai valid mengonfirmasi keandalan teknis penambahan fungsi administratif tersebut, namun capaian ini tidak dapat dijadikan bukti peningkatan kinerja administratif secara kuantitatif.

Penelitian ini memiliki keterbatasan pada jumlah data historis (empat tahun) dan penggunaan model regresi linier tunggal yang mengasumsikan hubungan linier antara waktu dan jumlah pendaftar, padahal pola pendaftar pada praktiknya cenderung fluktuatif. Perlu ditegaskan bahwa cakupan data historis yang terbatas pada empat tahun ajaran ini merupakan kendala riil ketersediaan arsip di SMA Islam Hidayatul Ahkam, bukan data yang sebenarnya dapat diperoleh namun tidak diusahakan; pencatatan jumlah pendaftar secara terarsip baru tersedia sejak tahun ajaran 2022/2023, sehingga penambahan observasi historis hanya dimungkinkan secara alami seiring berjalannya waktu. Sejalan dengan keterbatasan tersebut, perbandingan terhadap baseline naif telah dihitung pada bagian sebelumnya dan menunjukkan bahwa model regresi mengungguli baseline tersebut pada data historis yang tersedia. Validasi menggunakan skema rolling-origin, di sisi lain, belum dapat dieksekusi pada penelitian ini karena metode tersebut mensyaratkan pembagian data ke dalam beberapa subset pelatihan-pengujian pada rentang waktu yang tidak dimungkinkan oleh hanya empat titik data yang tersedia. Penelitian ini karenanya secara sadar diposisikan sebagai purwarupa (proof-of-concept) estimasi operasional, sementara validasi rolling-origin ditempatkan sebagai agenda penelitian lanjutan sebagaimana diuraikan pada bagian Kesimpulan. Peluang pengembangan lanjutan meliputi penambahan volume data historis seiring waktu, perbandingan dengan metode prediksi alternatif seperti regresi polinomial, exponential smoothing, ARIMA, atau model pembelajaran mesin (Random Forest, Support Vector Regression), serta pengujian keamanan tambahan dan generalisasi model pada sekolah lain dengan karakteristik data yang berbeda.

5. KESIMPULAN

Penelitian ini telah berhasil merampungkan tahap analisis, perancangan, dan implementasi sistem informasi PPDB berbasis web dengan mengikuti metodologi Rapid Application Development (RAD) serta memanfaatkan PHP native, PDO, dan MySQL. Sistem ini terbukti mempercepat proses pendaftaran, pemeriksaan dokumen, pengumuman hasil seleksi, dan pencetakan kartu peserta, dibandingkan mekanisme manual yang sebelumnya berlaku di SMA Islam Hidayatul Ahkam. Penerapan Regresi Linier Sederhana pada modul prediksi menghasilkan persamaan $\hat{Y} = 36 + 10,9X$, yang memproyeksikan 91 pendaftar (dengan rentang 77 sampai 105) untuk tahun ajaran 2026/2027, disertai nilai MAPE sebesar 15,14% dan R^2 sebesar 47,65%. Hasil pengujian black box pada sembilan skenario serta white box pada 15 jalur logika (dengan kompleksitas siklomatis $V(G) = 6$) sama-sama dinyatakan valid, dan keluaran sistem menunjukkan kecocokan 100% (bukan “validasi”, melainkan kesesuaian penuh) dengan hasil perhitungan manual pada semua metrik yang diuji, yang menegaskan konsistensi implementasi perangkat lunak. Berdasarkan capaian tersebut, dapat disimpulkan bahwa sistem yang dibangun berhasil mengimplementasikan modul regresi linier secara benar dan menghasilkan estimasi operasional sebesar 91 pendaftar untuk tahun ajaran 2026/2027. Meskipun demikian, keandalan prediktif model ini masih terbatas mengingat hanya empat observasi historis yang tersedia, nilai MAPE dan R^2 yang dilaporkan bersifat in-sample, serta belum dilakukan validasi out-of-sample; dengan demikian, sistem ini pada tahap sekarang lebih tepat diposisikan sebagai purwarupa (proof-of-concept) alat bantu perencanaan yang dapat menjadi salah satu masukan, bukan satu-satunya dasar bagi SMA Islam Hidayatul Ahkam dalam menyusun strategi penerimaan siswa baru ke depannya. Untuk penelitian selanjutnya, disarankan agar volume data historis diperluas, dilakukan validasi menggunakan skema rolling-origin, serta dibandingkan dengan metode prediksi lain (seperti regresi polinomial, exponential smoothing, atau ARIMA), serta diuji kemampuan generalisasi model pada sekolah-sekolah lain yang memiliki karakteristik data berbeda.

KONTRIBUSI PENULIS

Nasrul Abdul Karim: Konseptualisasi, Metodologi, Perangkat Lunak, Validasi, Analisis Formal, Investigasi, Penulisan Draf Awal, Peninjauan & Penyuntingan.

UCAPAN TERIMA KASIH

Terima kasih penulis sampaikan kepada pimpinan serta para dosen Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, atas bimbingan akademik yang telah diberikan selama

proses penelitian. Ucapan terima kasih juga ditujukan kepada Kepala Sekolah dan Staf Tata Usaha SMA Islam Hidayatul Ahkam yang telah memberikan izin serta membantu penyediaan data selama penelitian ini berlangsung.

PENDANAAN

Tidak ada pihak eksternal yang memberikan pendanaan untuk penelitian ini.

PERNYATAAN KETERSEDIAAN DATA

Data historis jumlah pendaftar yang dipakai dalam penelitian ini bersumber dari arsip SMA Islam Hidayatul Ahkam. Data tersebut dapat diperoleh dari penulis korespondensi melalui permintaan yang wajar, dengan tetap memperoleh izin dari pihak sekolah.

KONFLIK KEPENTINGAN

Penulis menyatakan bahwa tidak terdapat konflik kepentingan dalam penelitian, penulisan, maupun publikasi artikel ini.

REFERENCES

- [1] R. Firmansyah, Y. Komalasari, S. W. K. Dewi, P. Mauliana, R. D. Sulastriningsih, and N. Hunaifif, "Digitalisasi sekolah sebagai metode pembelajaran di era pendidikan 4.0," *J. Sos. Hum. dan Pendidik. (JUSHPEN)*, vol. 2, no. 3, pp. 49-55, 2023, doi: 10.56127/jushpen.v2i3.1052.
- [2] M. Yusuf, D. Julianingsih, and T. Ramadhani, "Transformasi pendidikan digital 5.0 melalui integrasi inovasi ilmu pengetahuan dan teknologi," *J. Mentari*, vol. 2, no. 1, pp. 11-19, 2023, doi: 10.34306/mentari.v2i1.328.
- [3] Badan Pusat Statistik, *Statistik Pendidikan 2025*. Jakarta, Indonesia: BPS, 2025.
- [4] M. F. Bustami, "Perancangan sistem informasi penerimaan siswa baru pada sekolah SMA swasta berbasis web," *J. Manaj. dan Teknol. Inf.*, vol. 2, no. 1, 2024.
- [5] M. Misbachul Munir, Y. Agus Pranoto, and S. Adi Wibowo, "Penerapan metode regresi linier dalam perancangan aplikasi prediksi jumlah calon peserta didik baru berbasis website (studi kasus: MAN 2 Kota Malang)," *J. Mhs. Tek. Inform.*, vol. 7, no. 1, 2023.
- [6] D. Dodiman, N. Hidayah, and A. Asniati, "Prediksi jumlah siswa baru pada SMK Negeri 2 Buton dengan metode regresi linier dan exponential smoothing," *J. Akad. Pendidik. Mat.*, vol. 10, pp. 27-31, 2024, doi: 10.55340/japm.v10i1.1517.
- [7] A. S. Atmaja, A. S. Sidabalok, M. Raihan, F. A. Putra, and N. I. L. Silalahi, "Perancangan sistem informasi penerimaan peserta didik baru (PPDB) berbasis web," *J. Komput. Teknol. Inf. dan Sist. Inf. (JUKTISI)*, vol. 2, no. 3, pp. 515-523, 2024, doi: 10.62712/juktisi.v2i3.101.
- [8] Y. H. Agustin, E. Satria, and A. Nasrulloh, "Algoritma regresi linier dalam prediksi jumlah pendaftar program pendidikan di lingkungan pesantren," *J. Algoritma*, vol. 22, no. 1, pp. 557-568, 2025, doi: 10.33364/algoritma/v.22-1.1802.
- [9] F. Asrin and G. V. Utami, "Implementing website-based school information systems in public elementary schools using waterfall model," *J. Inf. Syst. Informatics*, vol. 5, no. 2, pp. 590-614, 2023, doi: 10.51519/journalisi.v5i2.495.
- [10] M. Kartikasari, B. Baharsyah, and F. Felawati, "Perancangan sistem informasi penerimaan peserta didik baru (PPDB) berbasis website pada SMP Negeri 10 Bayung Lencir," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 2, pp. 1199-1208, 2025, doi: 10.62712/juktisi.v4i2.582.
- [11] A. D. Setyowanti, Y. Fitriani, W. Indrarti, and L. K. Rahayu, "Perancangan sistem informasi PPDB (penerimaan peserta didik baru) berbasis website pada SMPK Pa Van Der Steur di Bekasi," *J. Inf. Syst. Informatics Comput.*, vol. 9, no. 2, p. 308, 2025, doi: 10.52362/jisicom.v9i2.2153.
- [12] A. Yudahana, I. Riadi, and A. Elvina, "Perancangan sistem informasi pendaftaran peserta didik baru (PPDB) berbasis web menggunakan metode Rapid Application Development (RAD)," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 8, no. 1, pp. 47-58, 2023, doi: 10.36341/rabit.v8i1.2977.
- [13] I. Riris Immasari and R. Rhamadhan, "Sistem informasi penerimaan peserta didik baru berbasis web pada Madrasah Addhiya Guru Sya'ban," *J. Manaj. Inform. Jayakarta*, 2023, doi: 10.52362/jmijayakarta.v3i1.1027.
- [14] N. N. Qoyimah, N. Anggraini, and A. Senubekti, "Pengembangan tahap pertama aplikasi SOPWatch untuk monitoring standar operasional prosedur berbasis website (studi kasus: Pengadilan Negeri Bale Bandung)," *J. Mhs. Tek. Inform.*, vol. 8, no. 5, 2024.

- [15] M. Jibril and M. Amin, "Pengujian sistem informasi e-modul pada SMPN 1 Tempuling menggunakan black box testing," *J. Ilm. Tek. Inform.*, vol. 6, no. 2, 2024.
- [16] M. A. Nurwicaksono, I. N. Lisa, A. R. Tiara, and R. Sidik, "Optimasi sistem informasi konsultasi hukum melalui pendekatan pengujian kombinasi white-box dan black-box," *J. Manaj. Inform. (JAMIKA)*, vol. 14, no. 1, pp. 1-15, 2023, doi: 10.34010/jamika.v14i1.10110.
- [17] M. K. Effendi, Sriyanto, Goesderilidar, H. W. Nugroho, and J. Triloka, "Penerapan data mining untuk prediksi pendaftar PPDB di SMKN3 Metro menggunakan machine learning," *J. Jaringan Sist. Inf. Robot. (JSR)*, vol. 9, no. 1, pp. 59-63, 2025, doi: 10.58486/jsr.v9i1.482.
- [18] Y. Pratiwi and L. W. Widiati, "Implementasi whitebox testing dengan teknik basis path pada pengujian halaman pencarian program promo," *J. Kecerdasan Buatan dan Teknol. Inf.*, vol. 4, no. 2, pp. 173-180, 2025, doi: 10.69916/jkbt.v4i2.280.
- [19] Y. Amanda and M. H. Ujjianti, "Perancangan sistem informasi penerimaan peserta didik baru pada Daycare dan Pre School Ananda Mandiri Slawi berbasis web," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 1, pp. 177-184, 2025, doi: 10.36040/jati.v9i1.12232.
- [20] A. Sayyidina Ali, S. Hidayat Al-ikhsan, and F. Riana, "Rancang bangun sistem informasi DIKAPER Kota Bogor untuk pengajuan dan monitoring rujukan berbasis web (studi kasus: role super admin dan admin)," *J. Mhs. Tek. Inform.*, vol. 10, no. 1, 2026.
- [21] A. Tahari, T. Umi Kalsum, and J. Fredricka, "Create an e-raport application using the Laravel framework and MySQL," *J. Media Comput. Sci.*, vol. 4, no. 2, pp. 373-384, 2025, doi: 10.37676/jmcs.v4i2.