

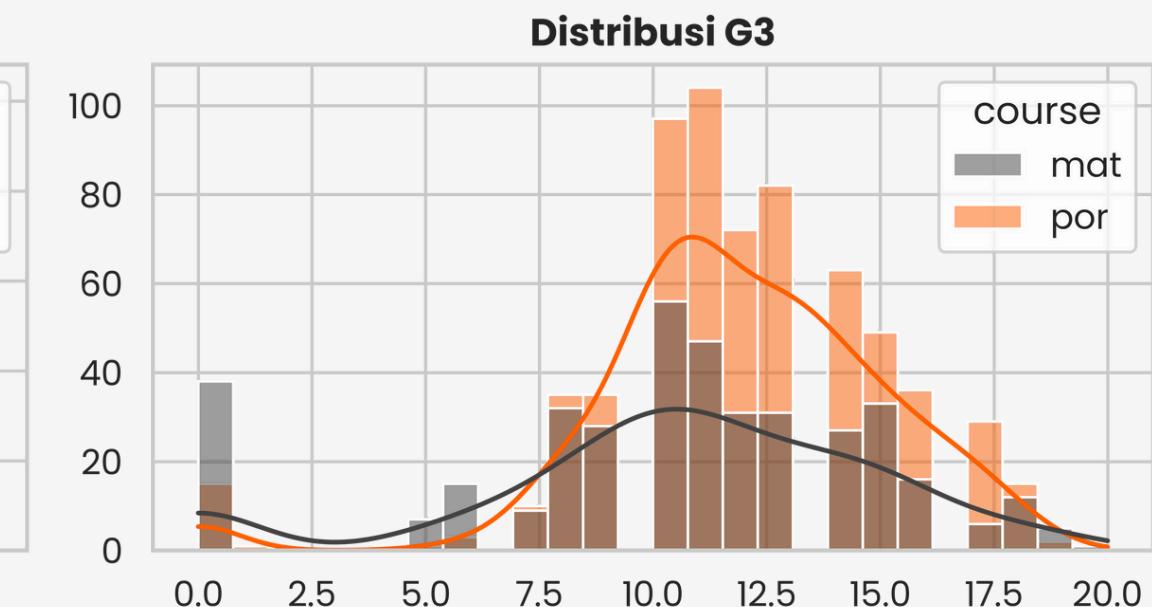
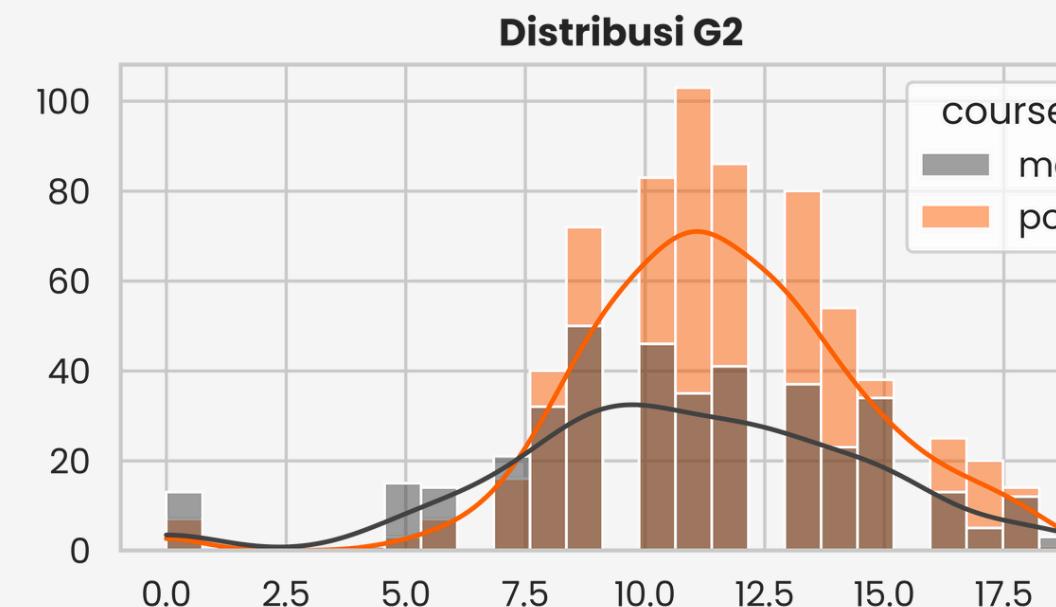
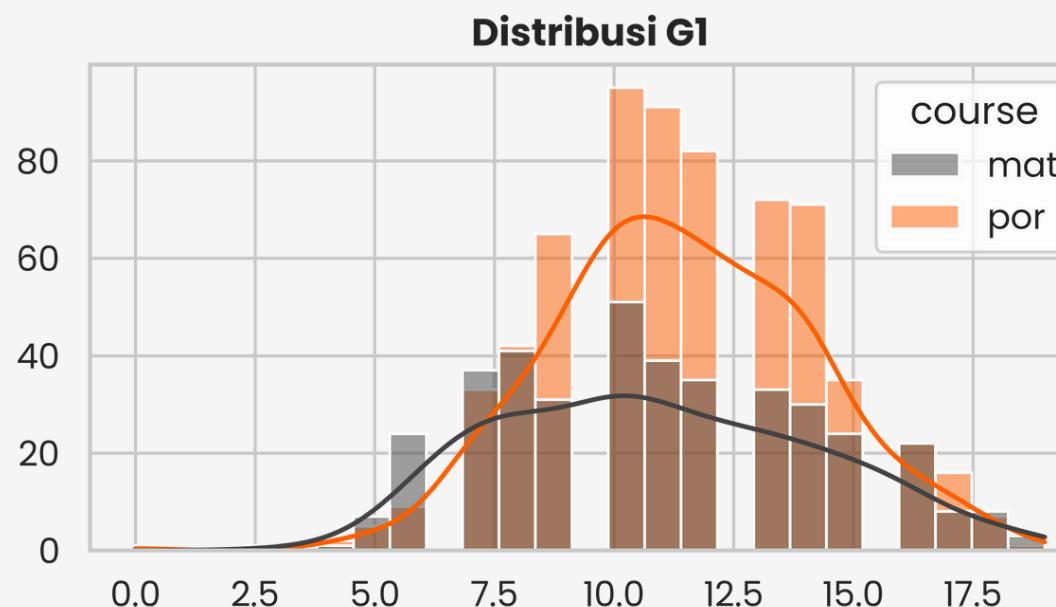
Analisis Prestasi Siswa

Dari Data ke Prediksi



Mengidentifikasi Faktor Kunci dan Membangun Model
Prediktif untuk Intervensi Dini Prestasi Siswa

Nilai Matematika vs Bahasa Portugis: Bagaimana?



Nilai Bahasa
Portugis lebih
bagus



Banyak
nilai nol di
Matematika

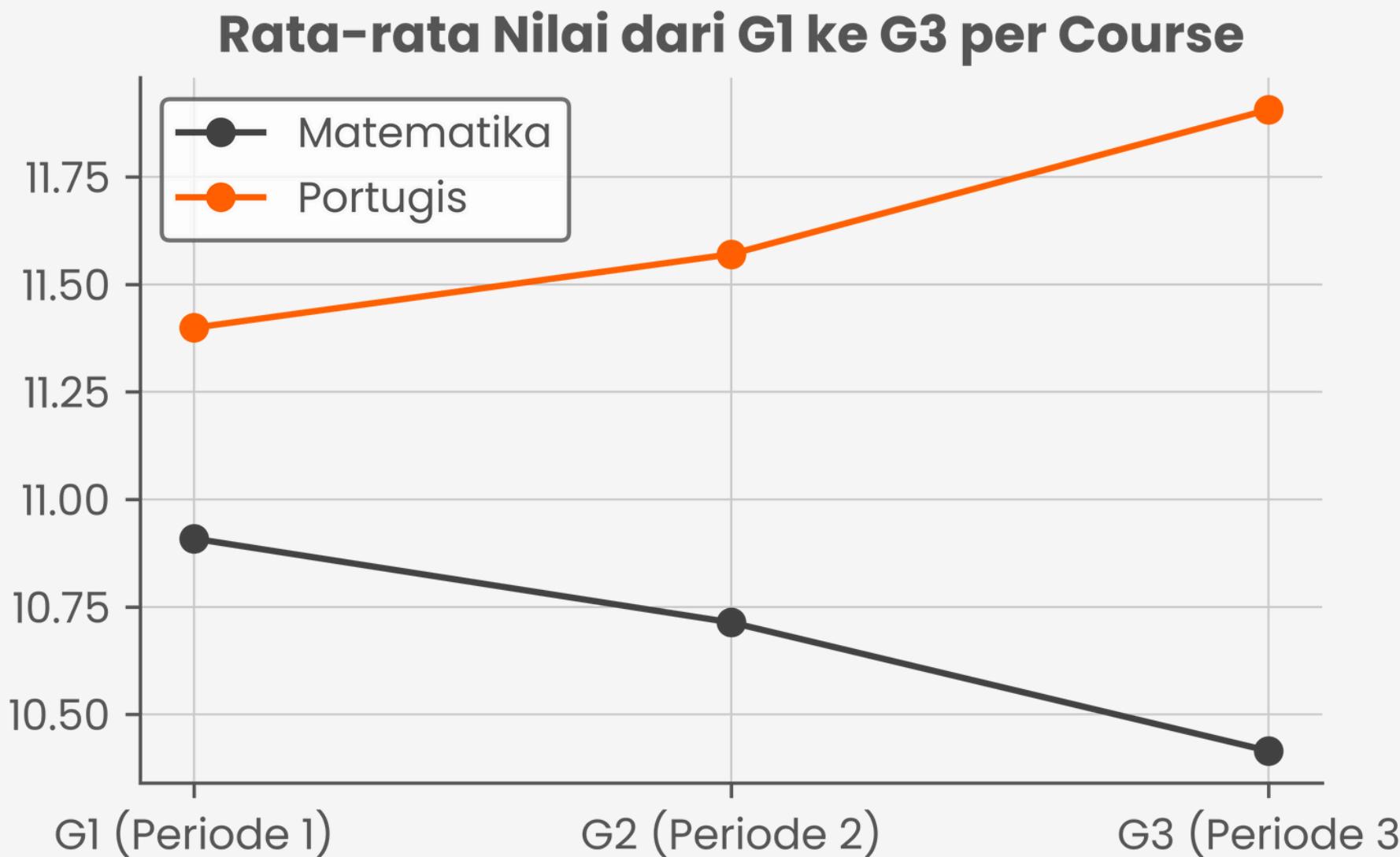


Bahasa Portugis
didominasi nilai
menengah



Matematika
terpolarisasi antara
rendah dan tinggi

Nilai Matematika vs Bahasa Portugis: Dari Waktu ke Waktu



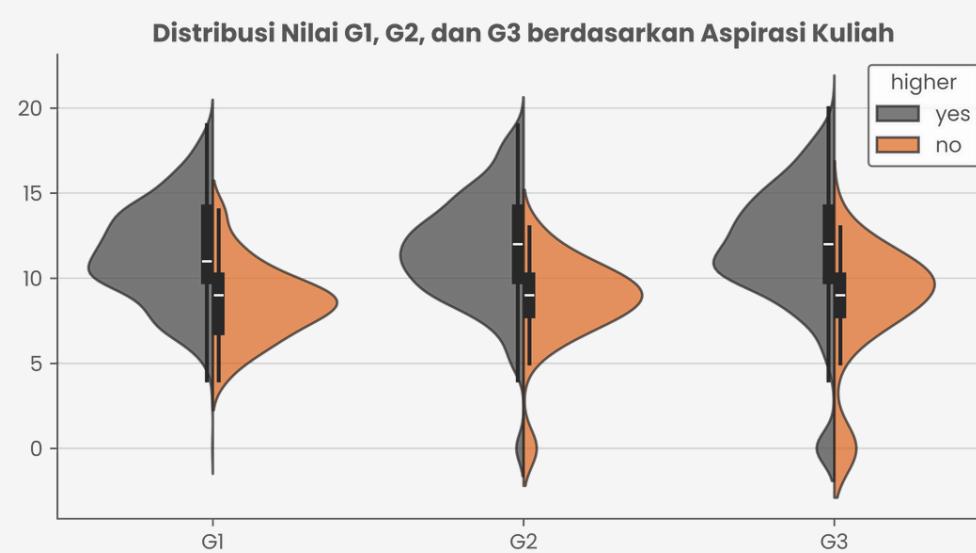
Bahasa Portugis

Nilai rata-rata siswa selalu lebih tinggi dan terus naik dari G1 ke G3. Ini menunjukkan kemajuan yang konsisten.

Matematika

Nilai rata-rata justru turun dari waktu ke waktu, jarak dengan Bahasa Portugis pun semakin lebar di akhir periode.

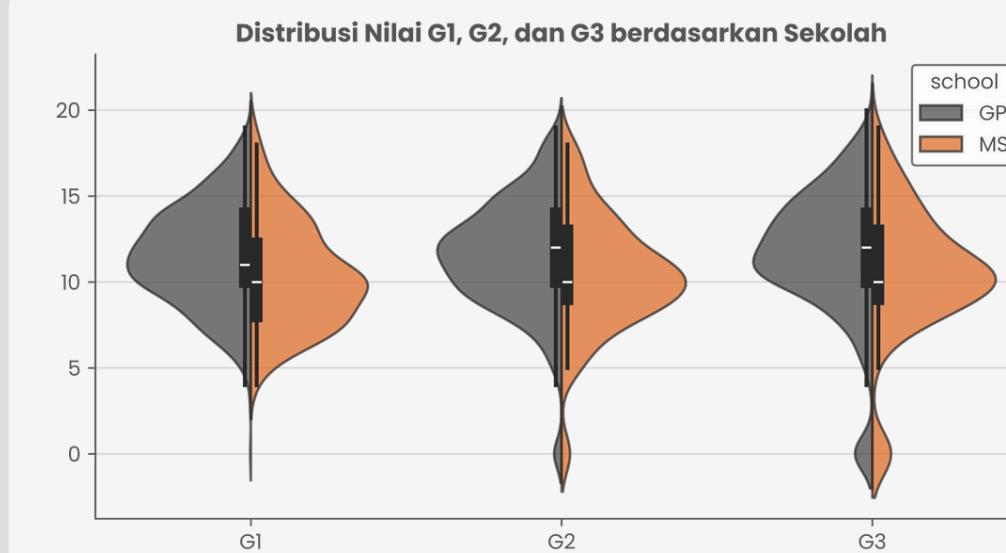
Apa yang Sebenarnya Mempengaruhi Prestasi Siswa?



Aspirasi kuliah jadi faktor terkuat, siswa yang ingin kuliah nilainya lebih tinggi



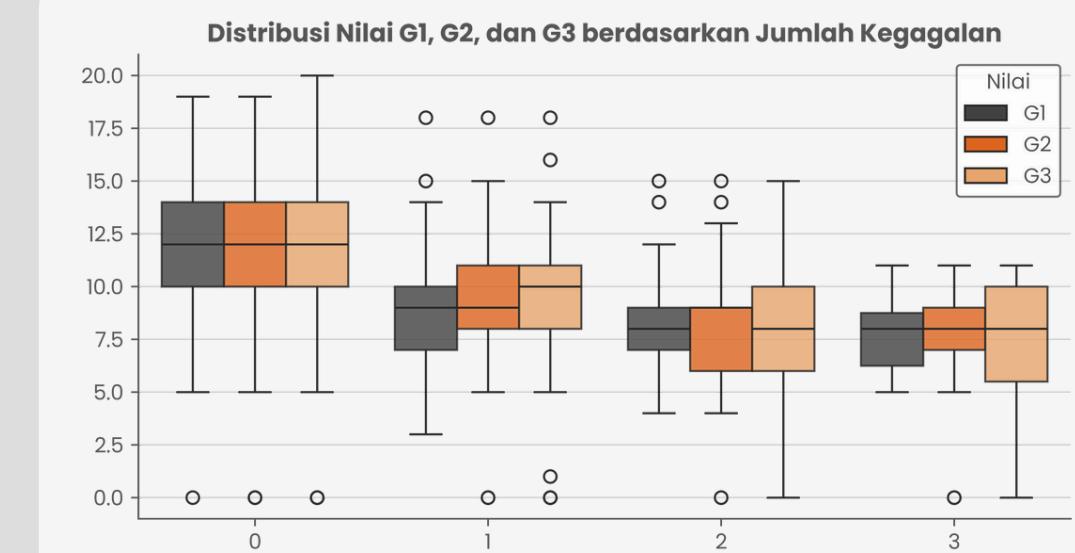
Terbukti signifikan secara statistik



Sekolah Gabriel Pereira dan Mousinho da Silveira punya gap konsisten



Terbukti signifikan secara statistik



Riwayat kegagalan menciptakan siklus negatif yang sulit diputus



Terbukti signifikan secara statistik

Data Preprocessing

Penanganan Data Kosong

Tidak ada data kosong pada dataset ini setelah diperiksa dengan fungsi `.isna()`

Penanganan Data Duplikat

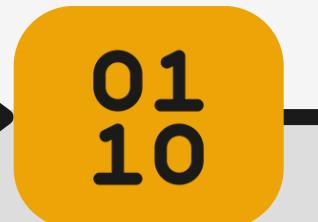
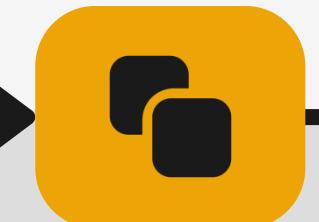
Tidak ada data duplikat pada dataset ini setelah diperiksa dengan fungsi `.duplicated()`

One-Hot Encoding

Semua kolom kategorik nominal diproses dengan One-Hot Encoding karena kategorinya sedikit

Penanganan Pencilan

Ditemukan 40% pencilan dengan metode IQR dan pencilan ditangani dengan capping



Perbandingan Performa Model

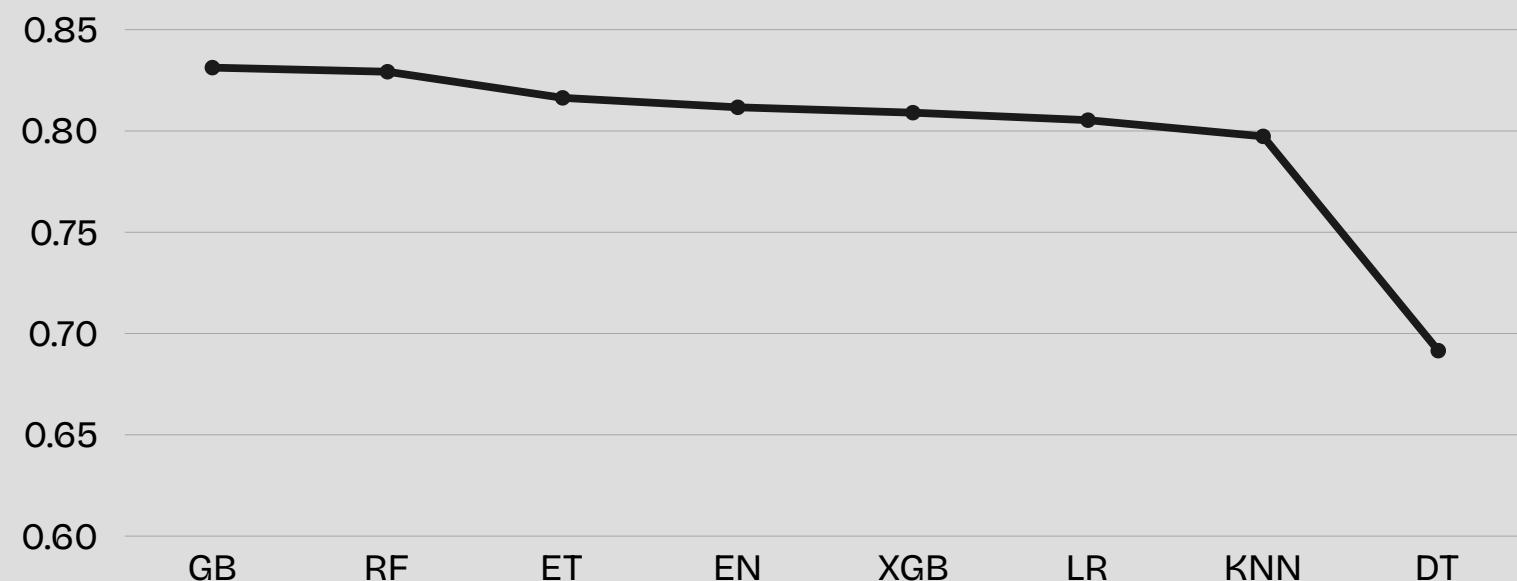
Evaluasi performa model regresi berdasarkan tiga metrik utama: **R²** (koefisien determinasi) di mana semakin tinggi nilainya, semakin baik model menjelaskan variansi data, **MAE** (Mean Absolute Error) dan **RMSE** (Root Mean Squared Error) di mana semakin rendah nilainya, semakin akurat prediksi.

Secara umum, model berbasis **ensemble cenderung unggul** dibandingkan model linear maupun non-parametrik, dengan akurasi lebih rendah dan error lebih tinggi.

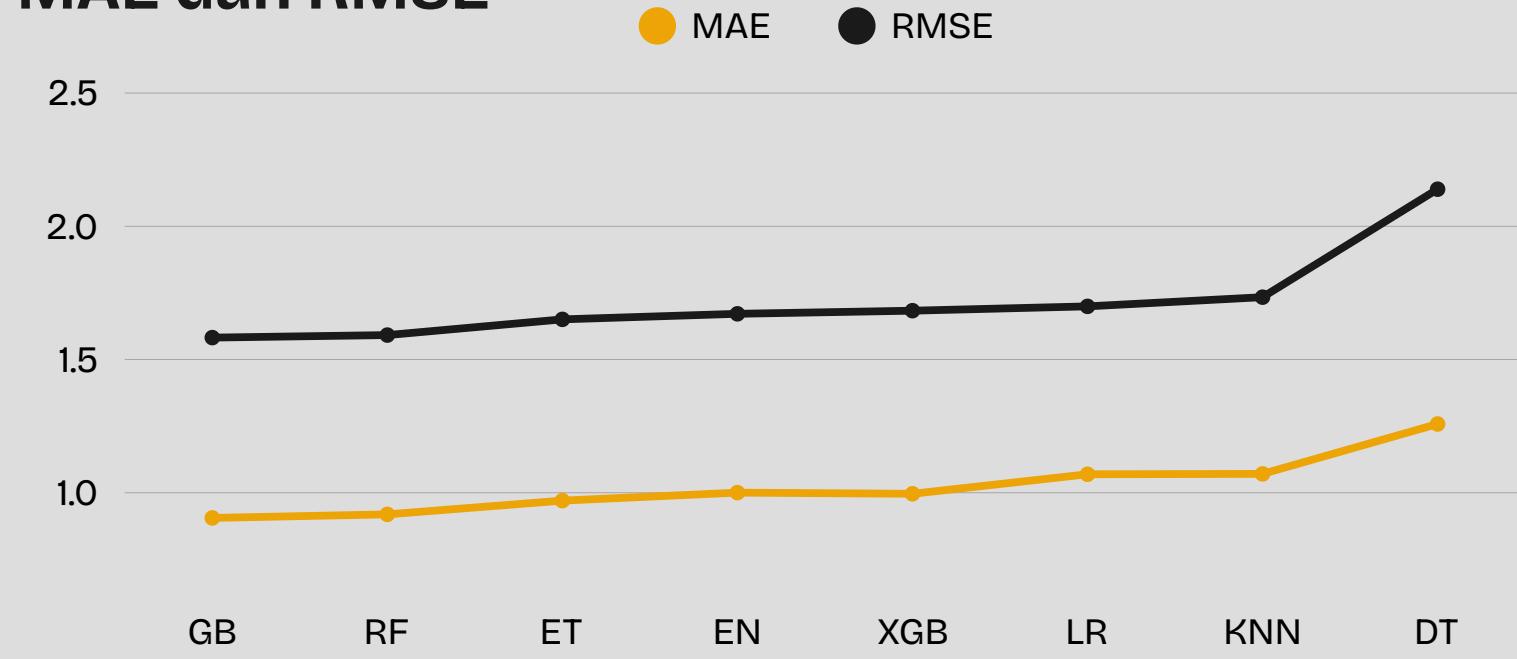


Model terpilih: Gradient Boosting

R² Score



MAE dan RMSE



Hyperparameter Tuning



**Gradient
Boosting**



**R-
squared**



**Grid
Search**

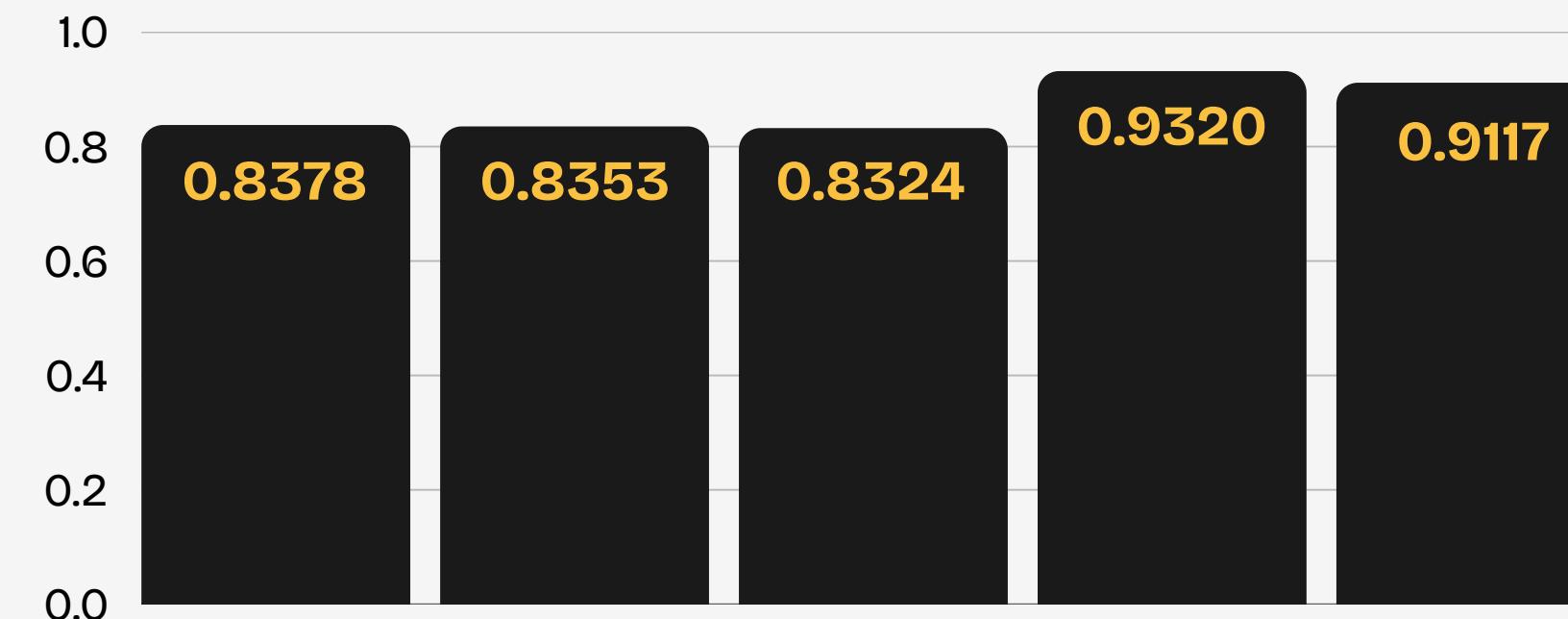


**5-fold
CrossVal**



```
best_param = {  
    'learning_rate': 0.05,  
    'max_depth': 3,  
    'max_features': None,  
    'min_samples_leaf': 4,  
    'min_samples_split': 10,  
    'n_estimators': 100  
}
```

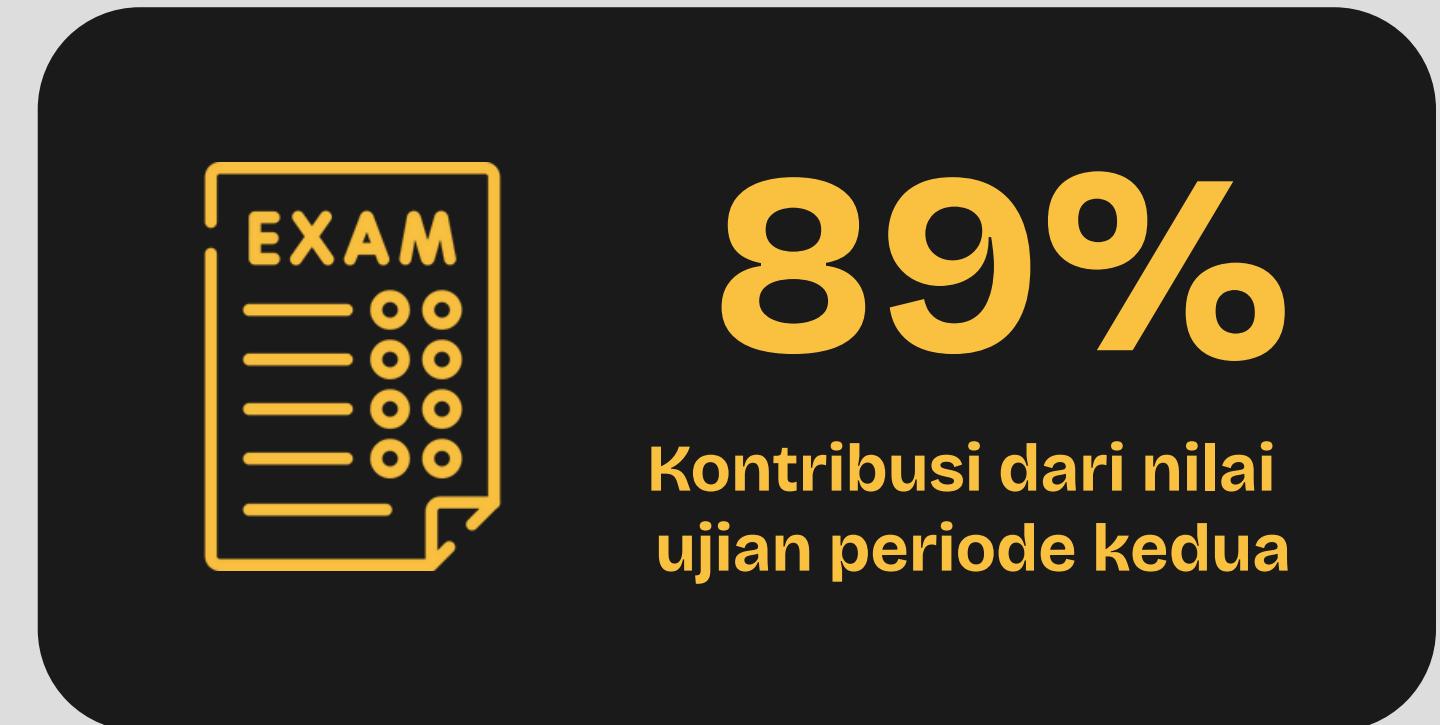
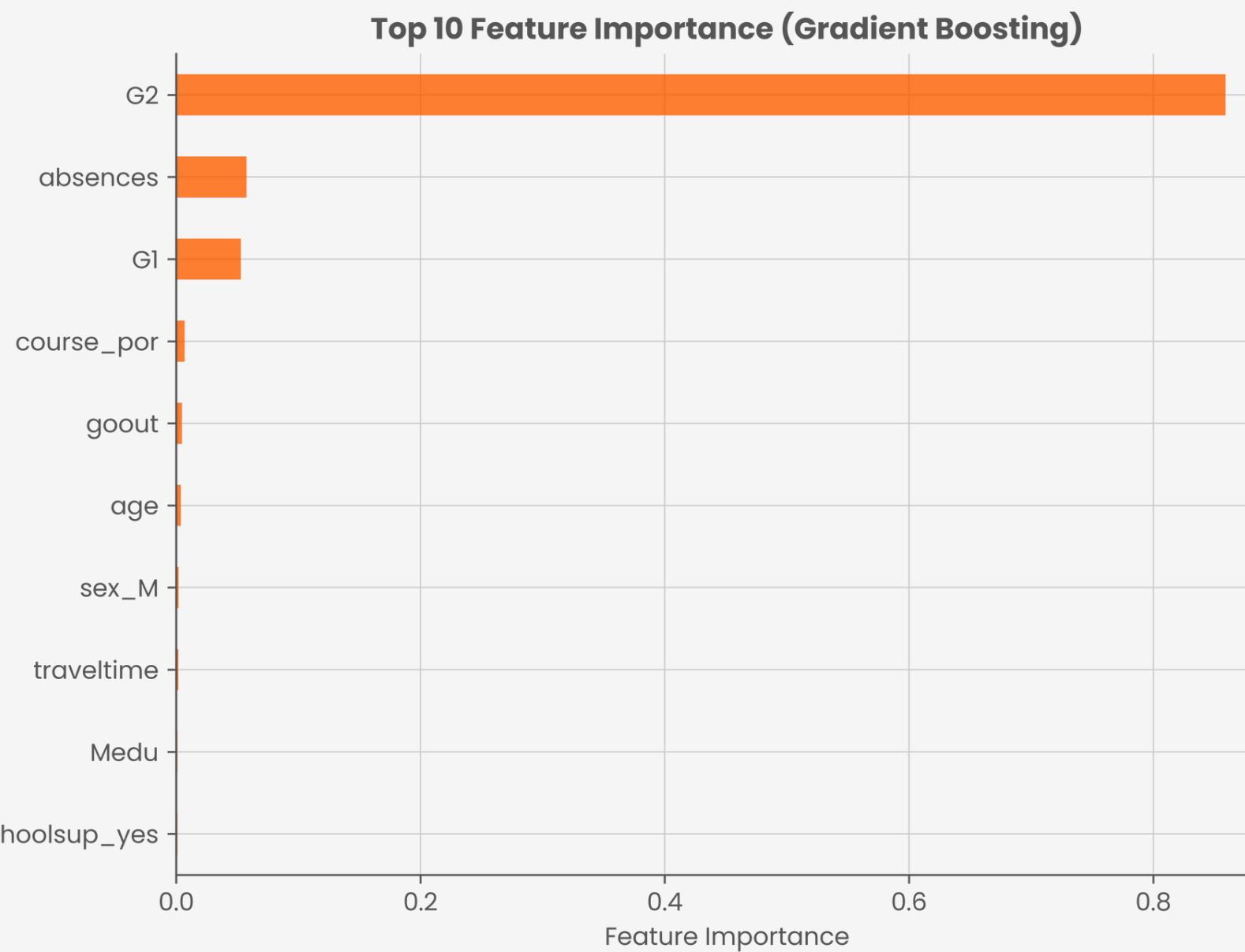
CV R² Scores



**Best
CV Score**

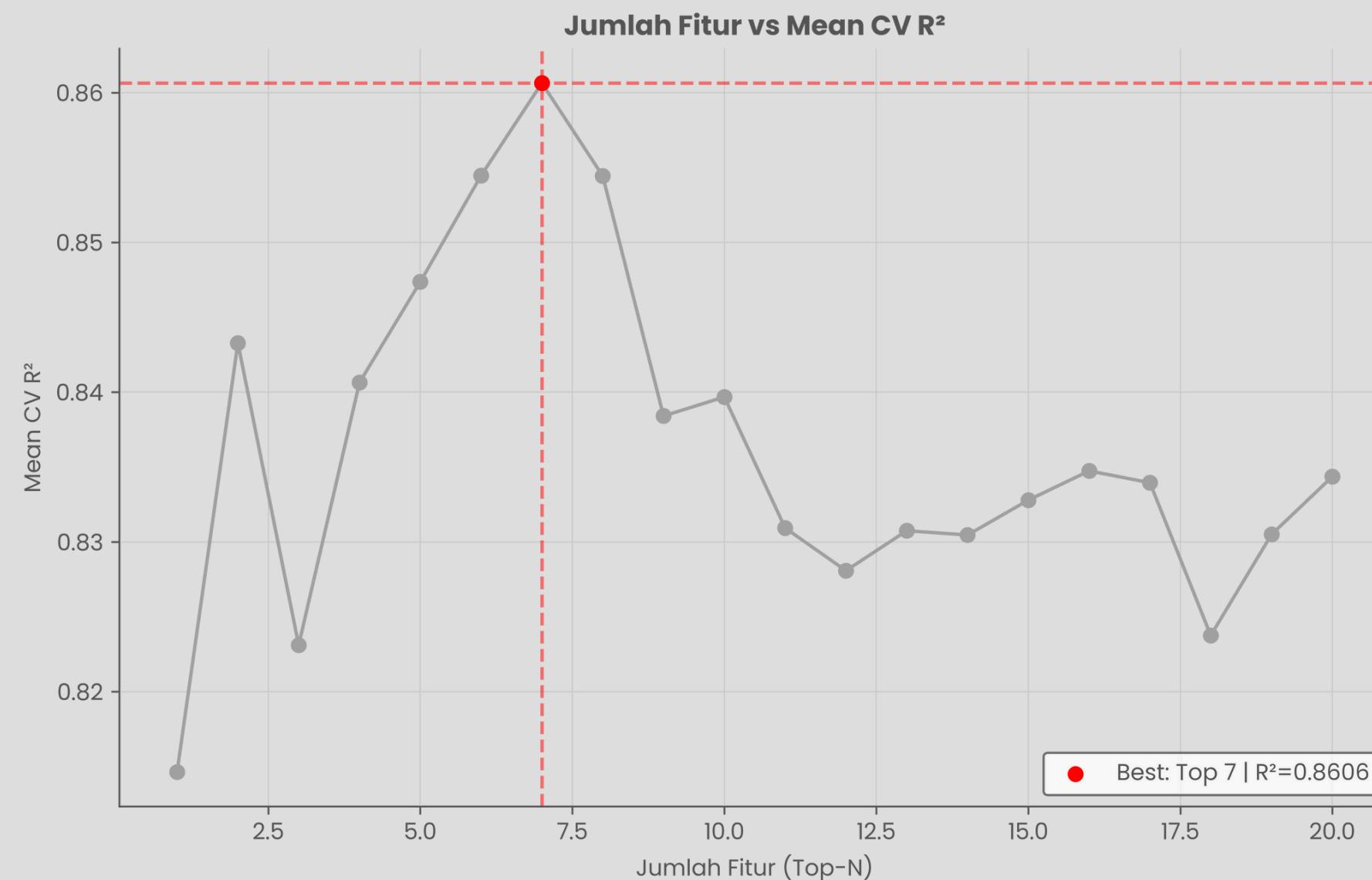
**0.8698
± 0.0860**

Feature Importance



Hasil menunjukkan bahwa **riwayat akademik** adalah **indikator paling kuat** untuk memprediksi nilai akhir, bukan demografi, latar belakang keluarga, ataupun kebiasaan sosial.

Optimisasi Model



Hasil menunjukkan bahwa **jumlah fitur yang paling optimal** untuk mendapatkan nilai **R-squared tertinggi** adalah **7 fitur** dengan skor R-squared sebesar **0.8606**, yaitu G2, absences, G1, course_por, goout, age, dan sex_M.



Top 7

fitur terbaik untuk mendapatkan skor R-squared tertinggi

Residual Analysis

Analisis residual menunjukkan bahwa model regresi memiliki **performa yang memadai** dengan **asumsi linearitas** dan **homoskedastisitas yang terpenuhi**.

0.8748

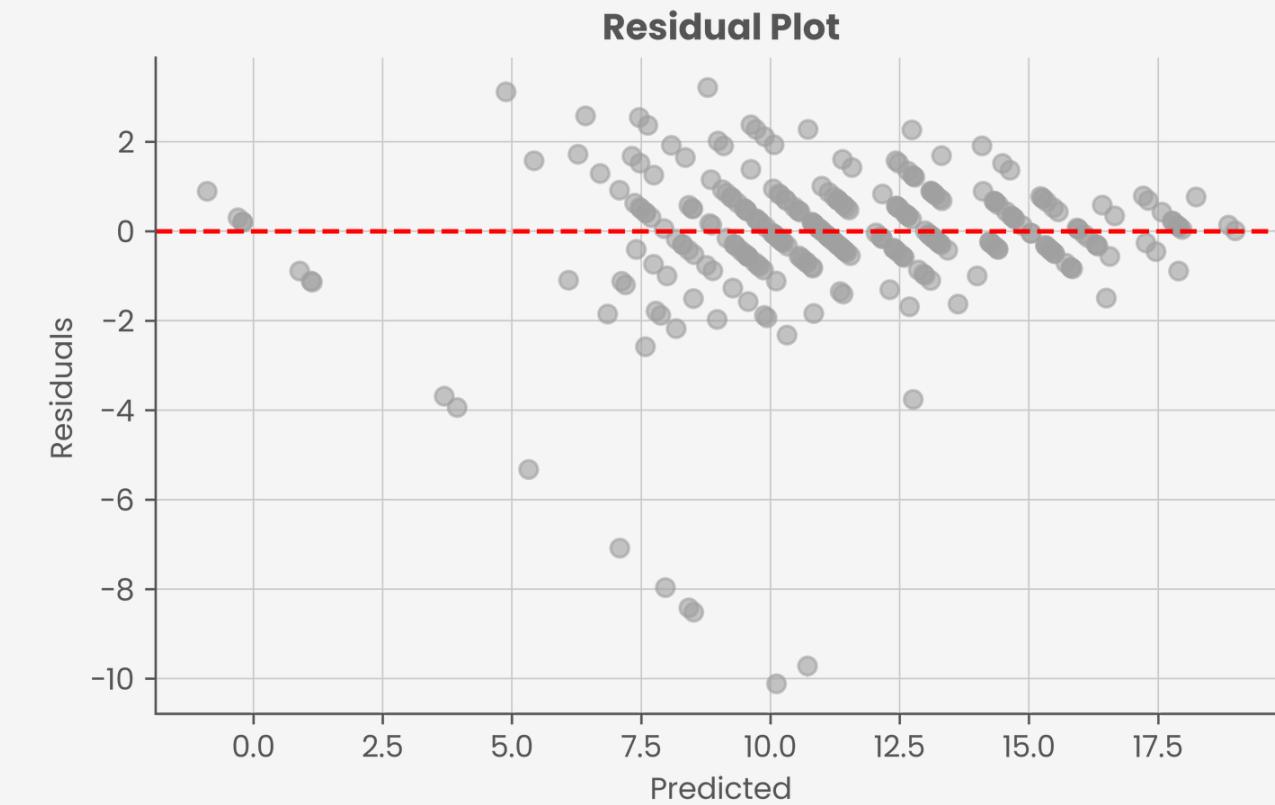
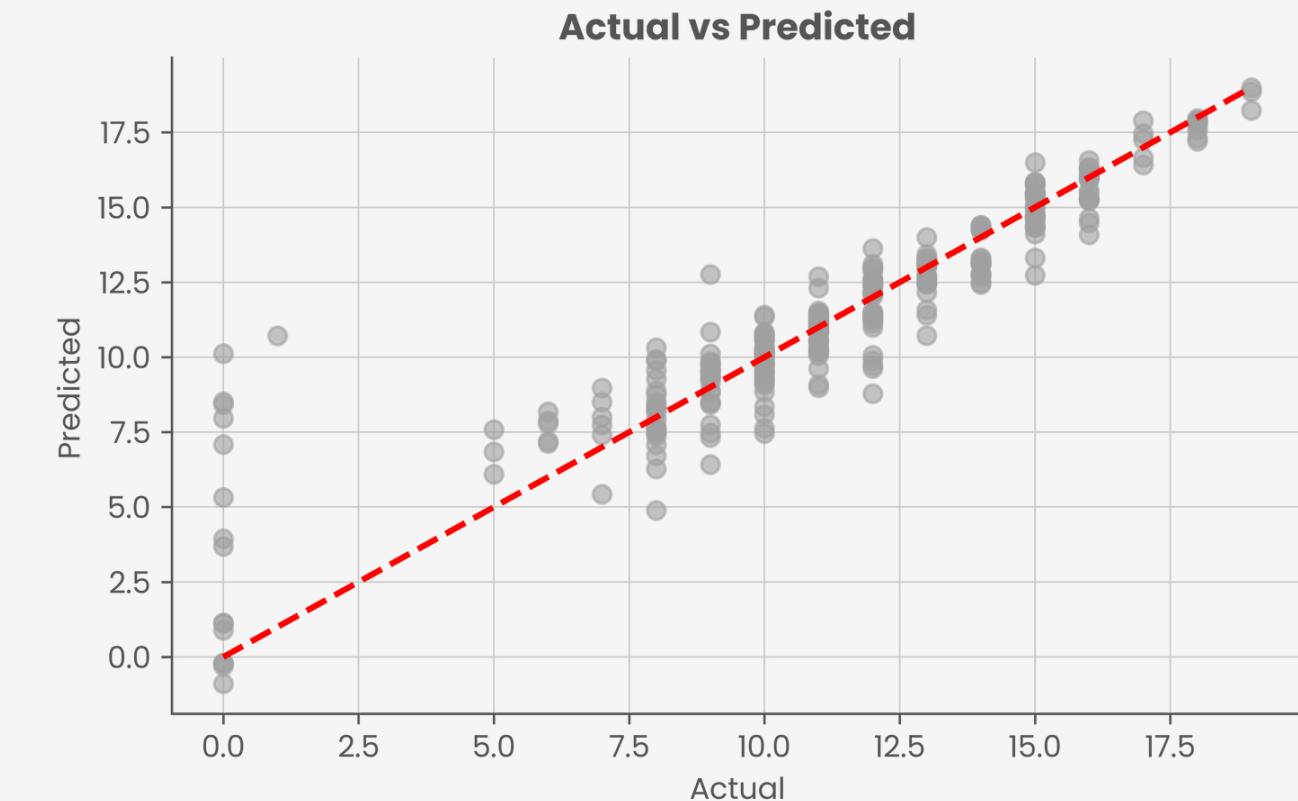
MAE

0.8748

RMSE

0.8748

MAPE



Residual Analysis



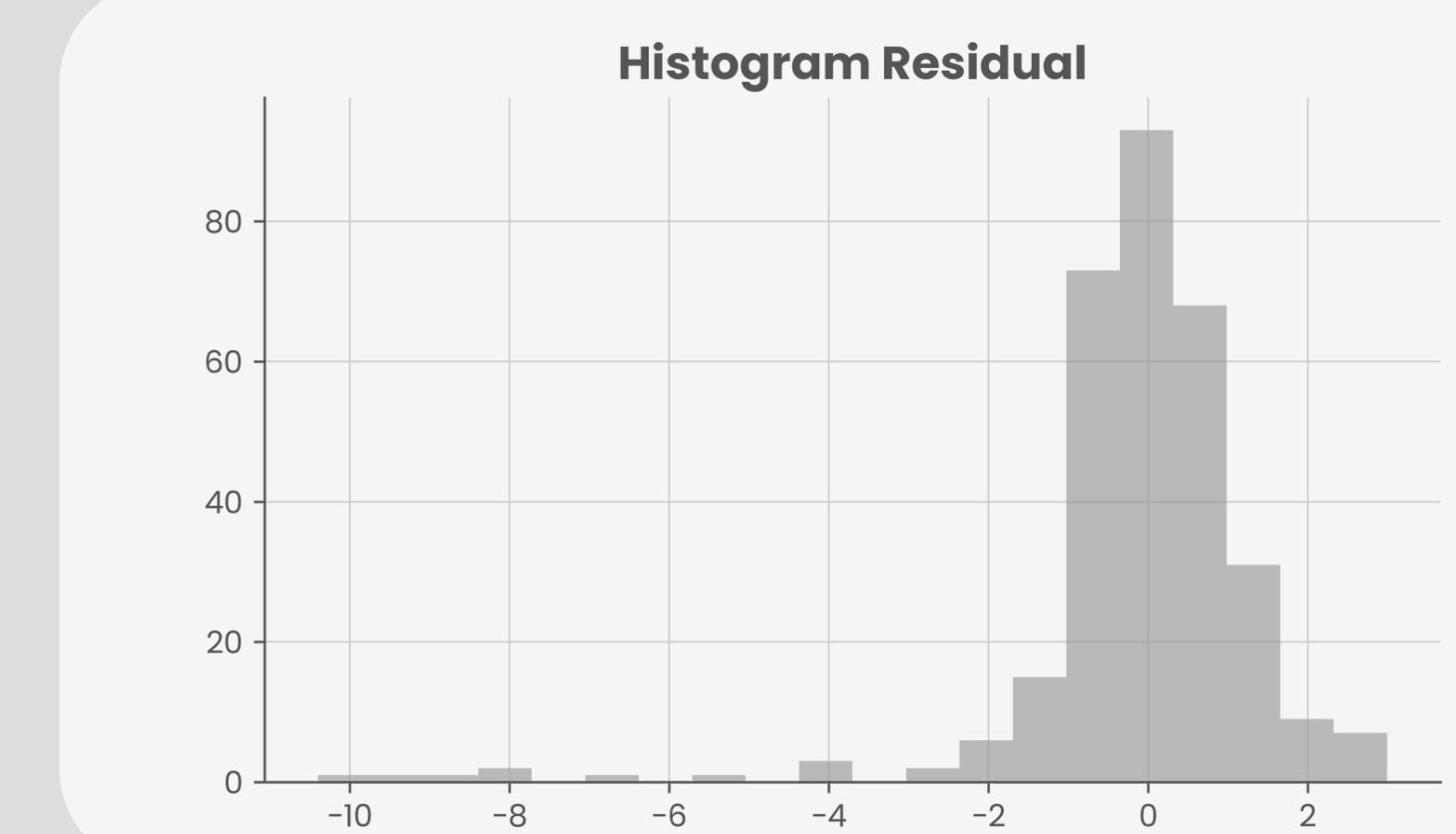
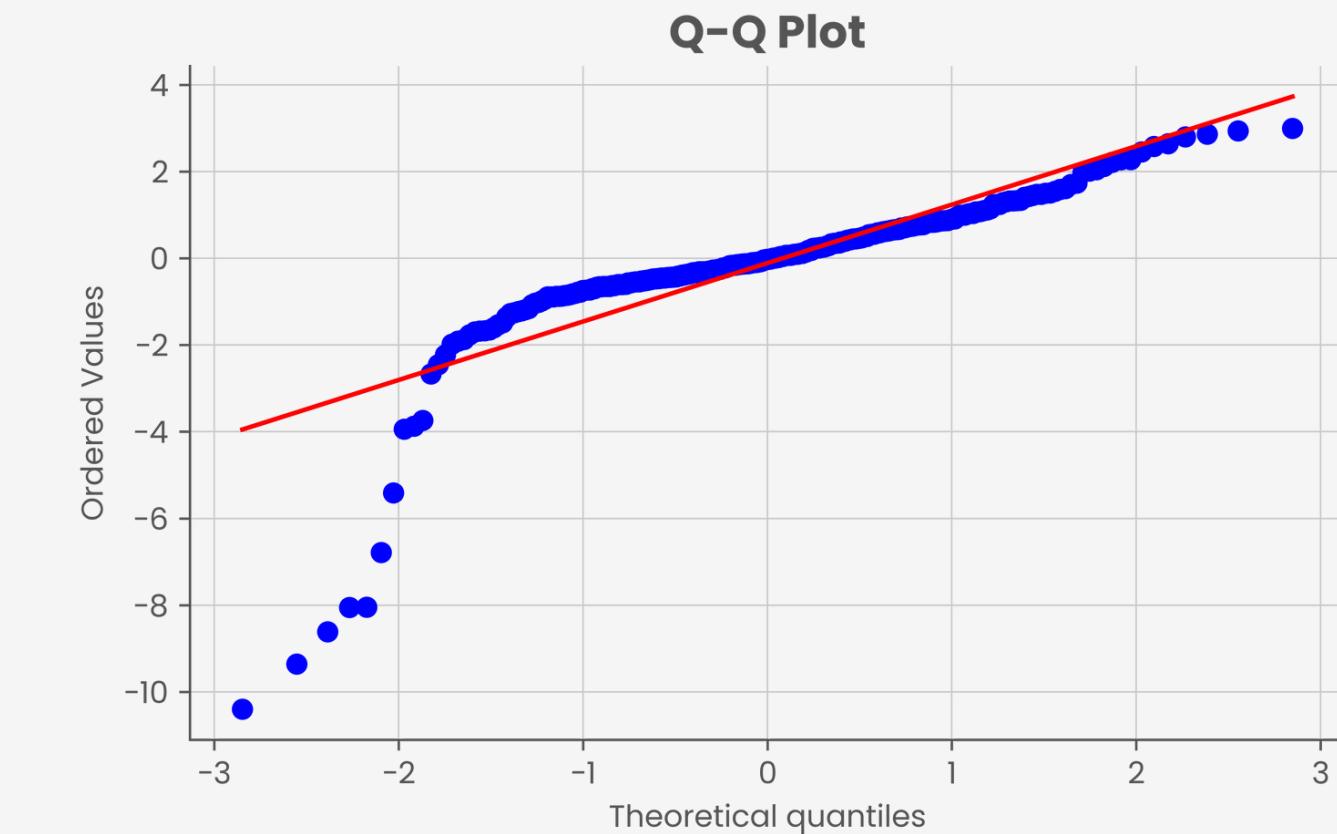
Model memiliki **performa yang memadai** dengan **residual yang sebagian besar mengikuti distribusi normal**.



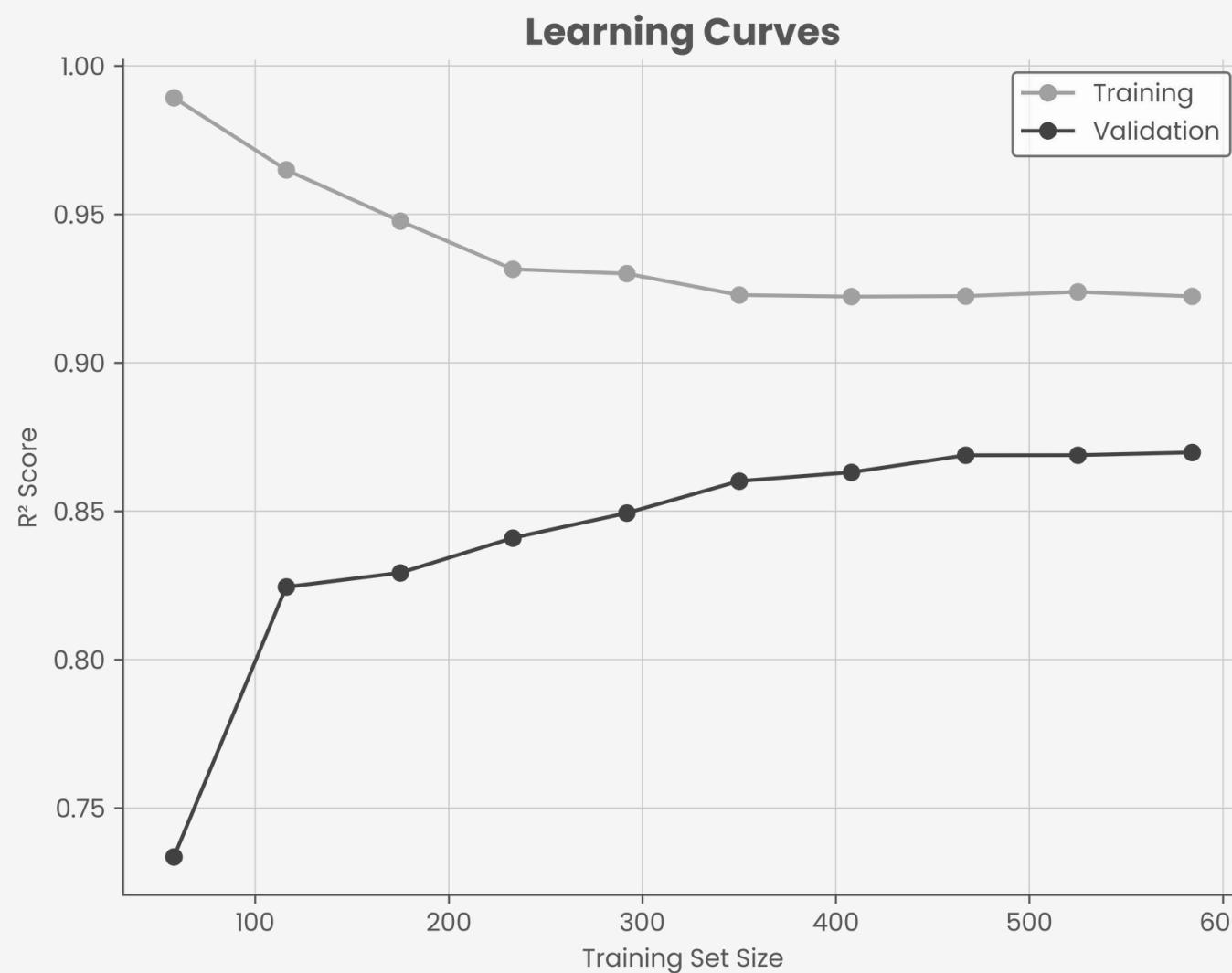
Q-Q plot mengindikasikan **penyimpangan di kedua ujung** distribusi yang menunjukkan adanya **outlier**.



Histogram residual menunjukkan **distribusi yang terpusat di sekitar nol** dengan sedikit skewness ke kiri.



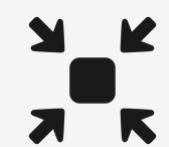
Learning Curves



Model **tidak mengalami underfitting**, training score tinggi dan validation score meningkat bertahap

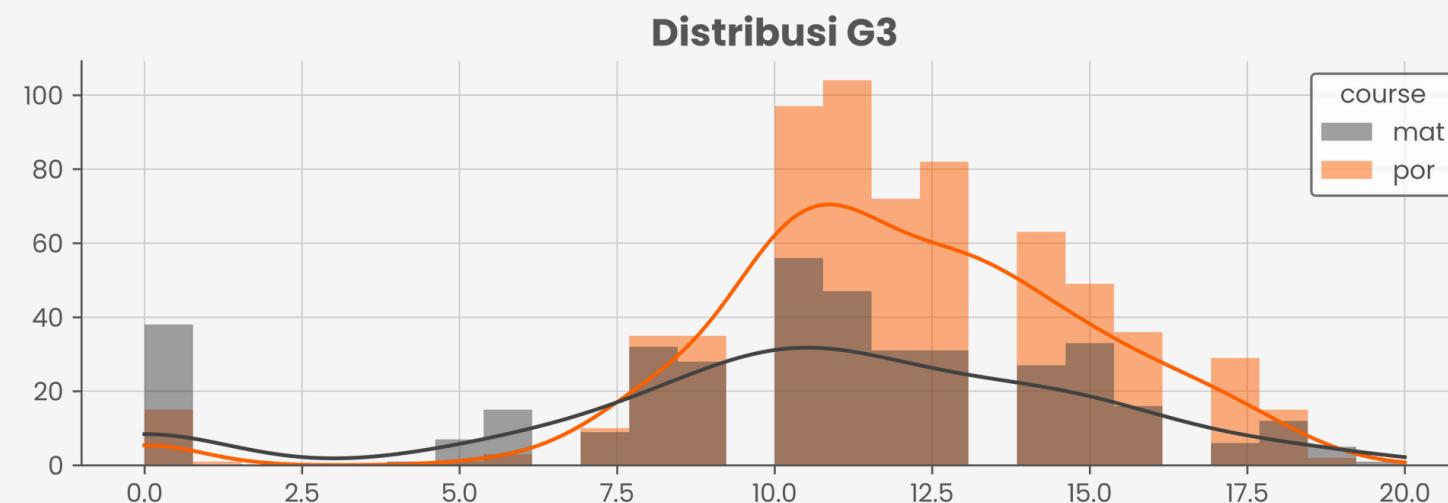


Terdapat **sedikit overfitting** namun **masih wajar**, gap antara training dan validation score tidak terlalu besar



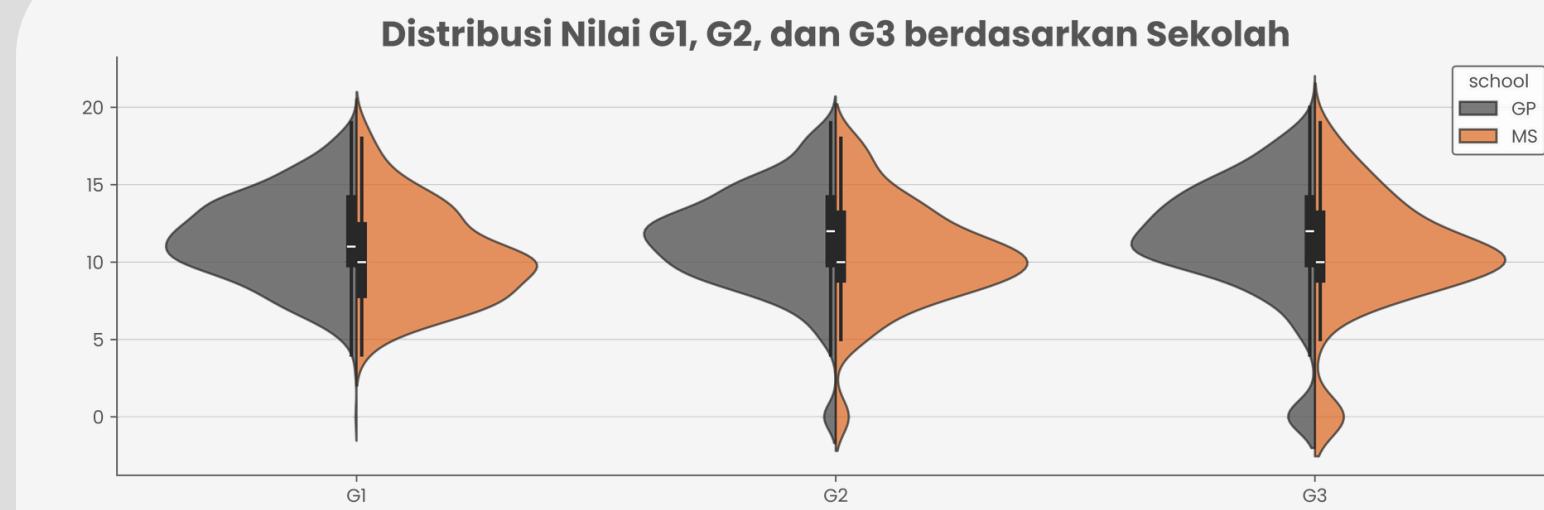
Model menunjukkan **performa yang stabil** di mana validation score **konvergen** tanpa fluktuasi ekstrem

Evaluasi Lanjutan



Evaluasi Pembelajaran Matematika

Karena track Matematika muncul sebagai salah satu kontributor signifikan dalam model prediksi, diperlukan evaluasi terhadap pendekatan pengajaran dan efektivitas kurikulum



Perbandingan Antarsekolah

Sekolah Gabriel Pereira menunjukkan rerata nilai akhir yang lebih tinggi, metode pembelajaran dan manajemen sekolah tersebut dapat dijadikan benchmark untuk pengembangan sekolah lain

Rekomendasi Strategi



Evaluasi Nilai Periode 1 & Nilai Periode 2



Pemantauan dan Tindak Lanjut Absensi



Program Peningkatan Motivasi Siswa

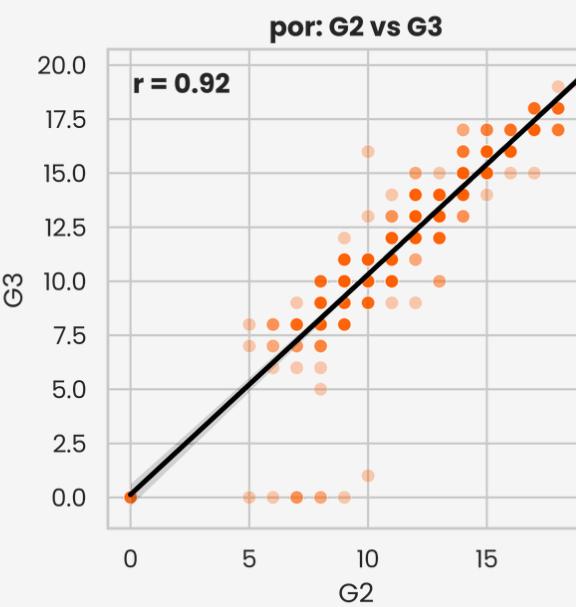
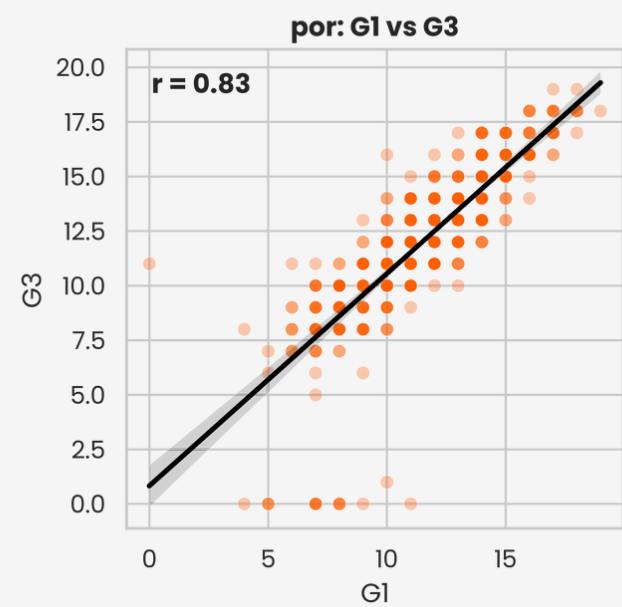
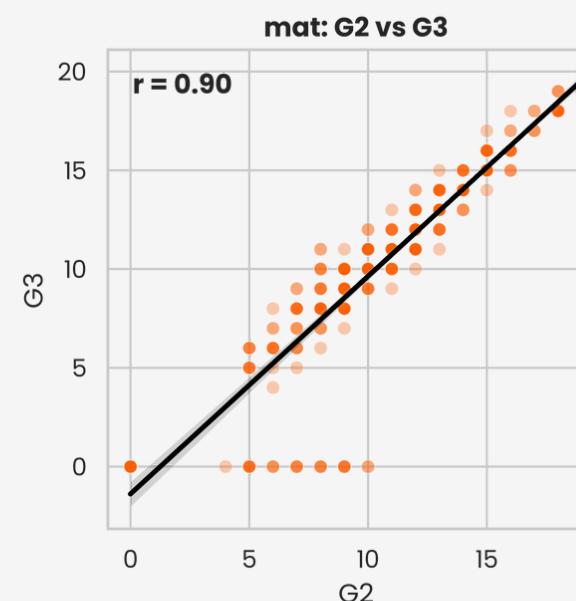
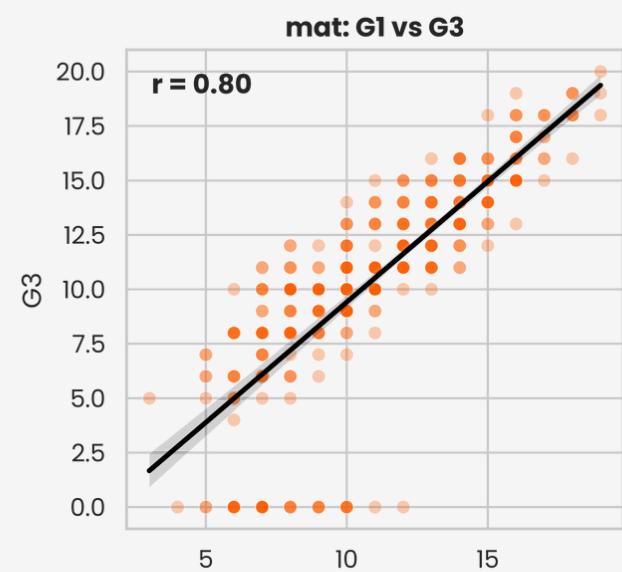


Variasi dan Diversifikasi Metode Pengajaran



Pendampingan Khusus untuk Siswa dengan Riwayat Kegagalan

Evaluasi Nilai Periode 1 & Periode 2



Nilai G1 dan G2 sangat berkorelasi dengan nilai akhir (G3), terlihat dari **grafik korelasi** dan **hasil analisis SHAP**. Ini berarti performa siswa di awal semester adalah **sinyal penting** untuk intervensi dini.



Siswa dengan nilai **di bawah 10** akan diikutkan dalam program **kelas tambahan**

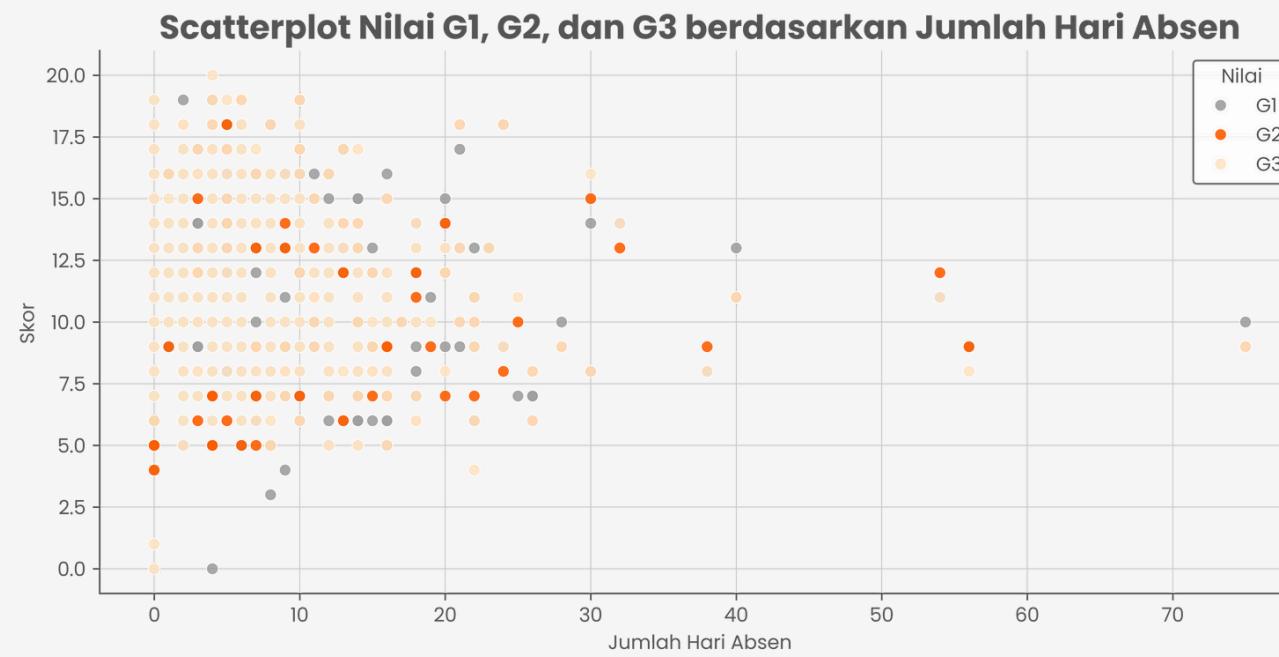


Siswa yang mengalami **penurunan nilai lebih dari 4 poin** dijadwalkan **sesi konsultasi** dengan guru



Evaluasi dilakukan secara **rutin setiap selesai ujian**, untuk memperbarui daftar siswa yang perlu dibina

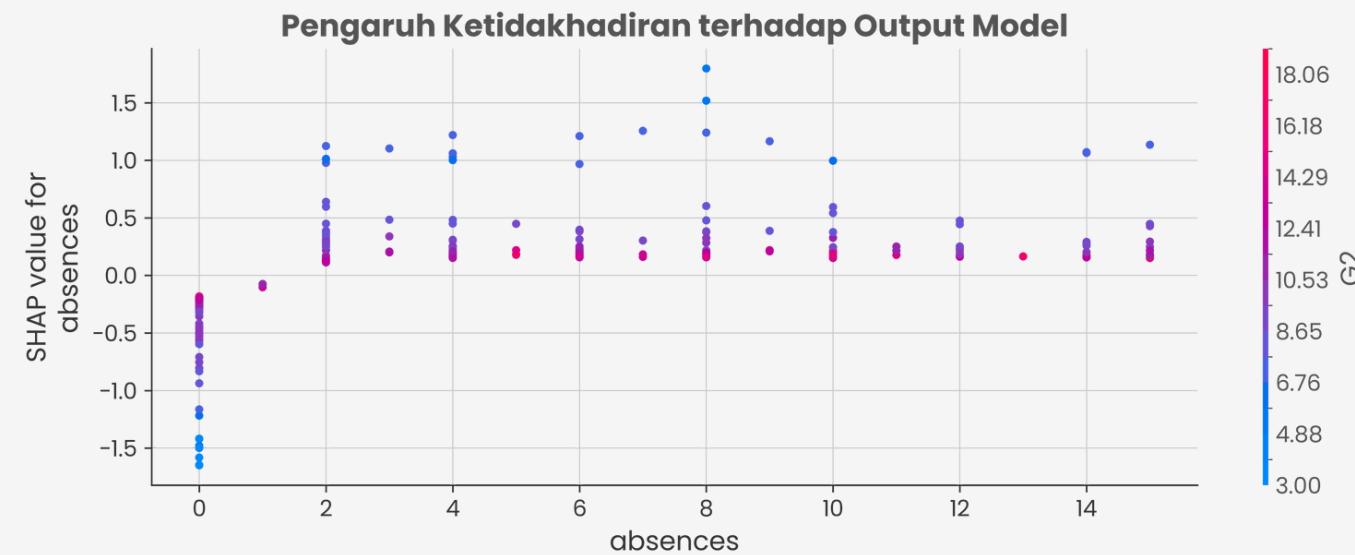
Pemantauan Absensi



Data menunjukkan bahwa **semakin sering siswa absen, semakin rendah kontribusinya terhadap nilai akhir**, meskipun nilai sebelumnya baik. Kehadiran yang konsisten tetap penting.



Absen > 5 hari: Siswa mengikuti pembinaan BK untuk klarifikasi dan bimbingan

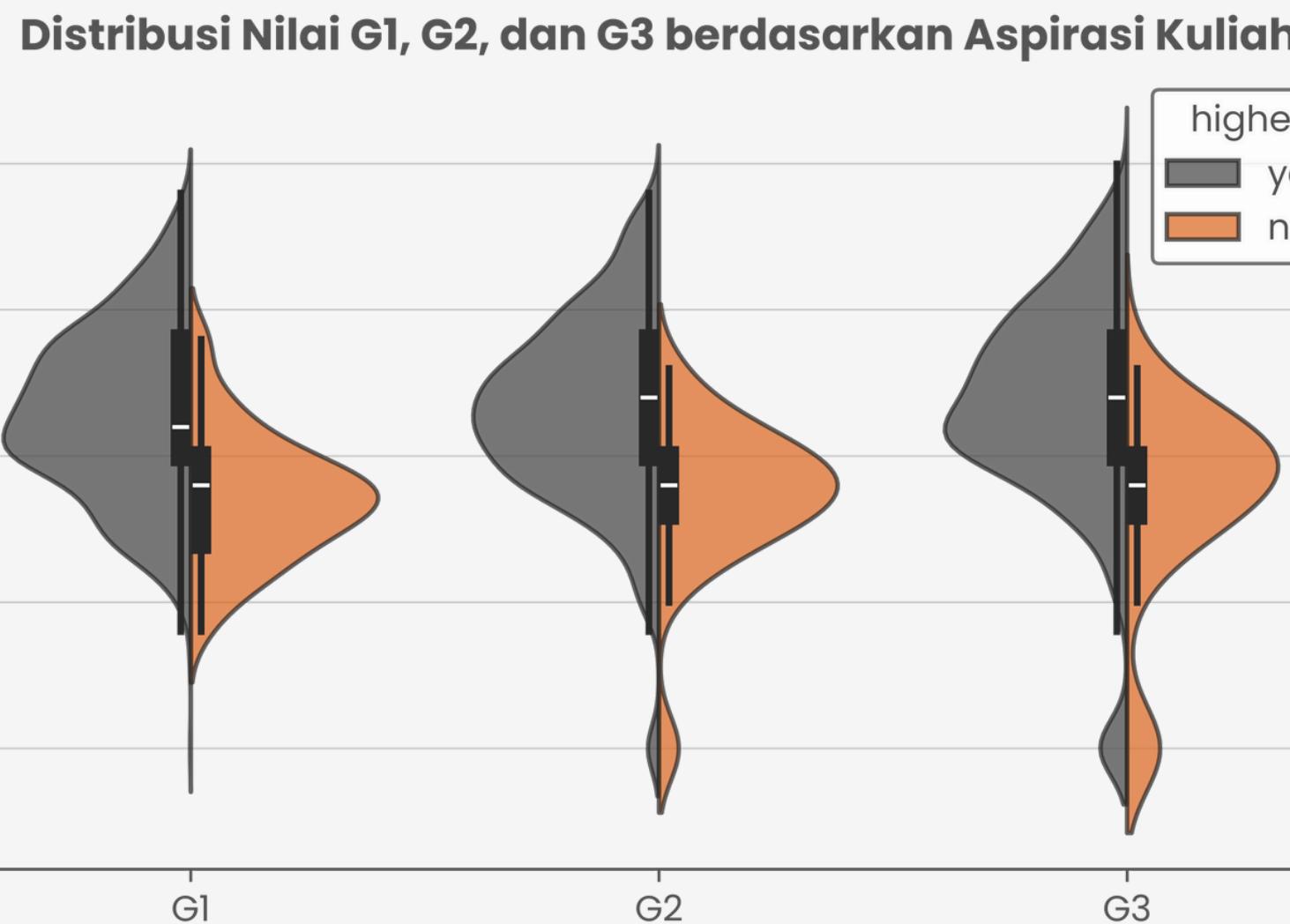


Absen > 10 hari: Siswa dibina oleh BK, dan orang tua dihubungi untuk koordinasi



Absen > 15 hari: Orang tua diminta hadir ke sekolah untuk pertemuan lanjutan bersama BK dan wali kelas

Program Peningkatan Motivasi Siswa



Siswa yang memiliki **aspirasi kuliah** cenderung menunjukkan **nilai akademik yang lebih tinggi**. Ini menunjukkan bahwa **motivasi belajar berperan penting**.



Bimbingan Karir Berkala



Seminar Profesi

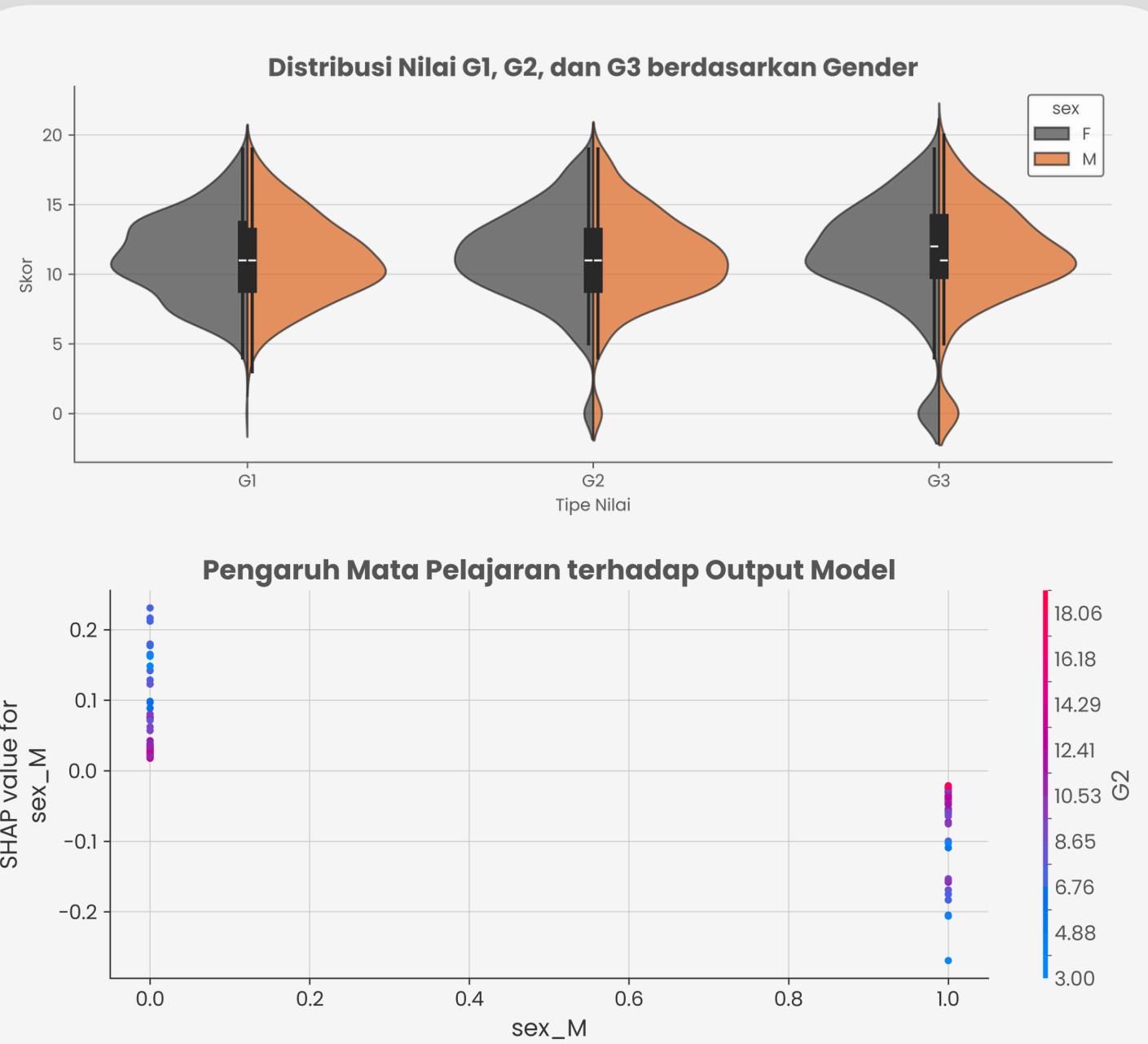


Kelas Inspirasi



Konseling Pendidikan Tinggi

Diversifikasi Metode Pengajaran



Mengurangi risiko penurunan nilai pada **siswa laki-laki** dengan menciptakan pengalaman belajar yang **lebih aktif, relevan, dan memotivasi** sesuai dengan **kecenderungan gaya belajar mereka**



Tanya Jawab Interaktif

Menjaga fokus dan mendorong partisipasi langsung



Gamifikasi Tugas

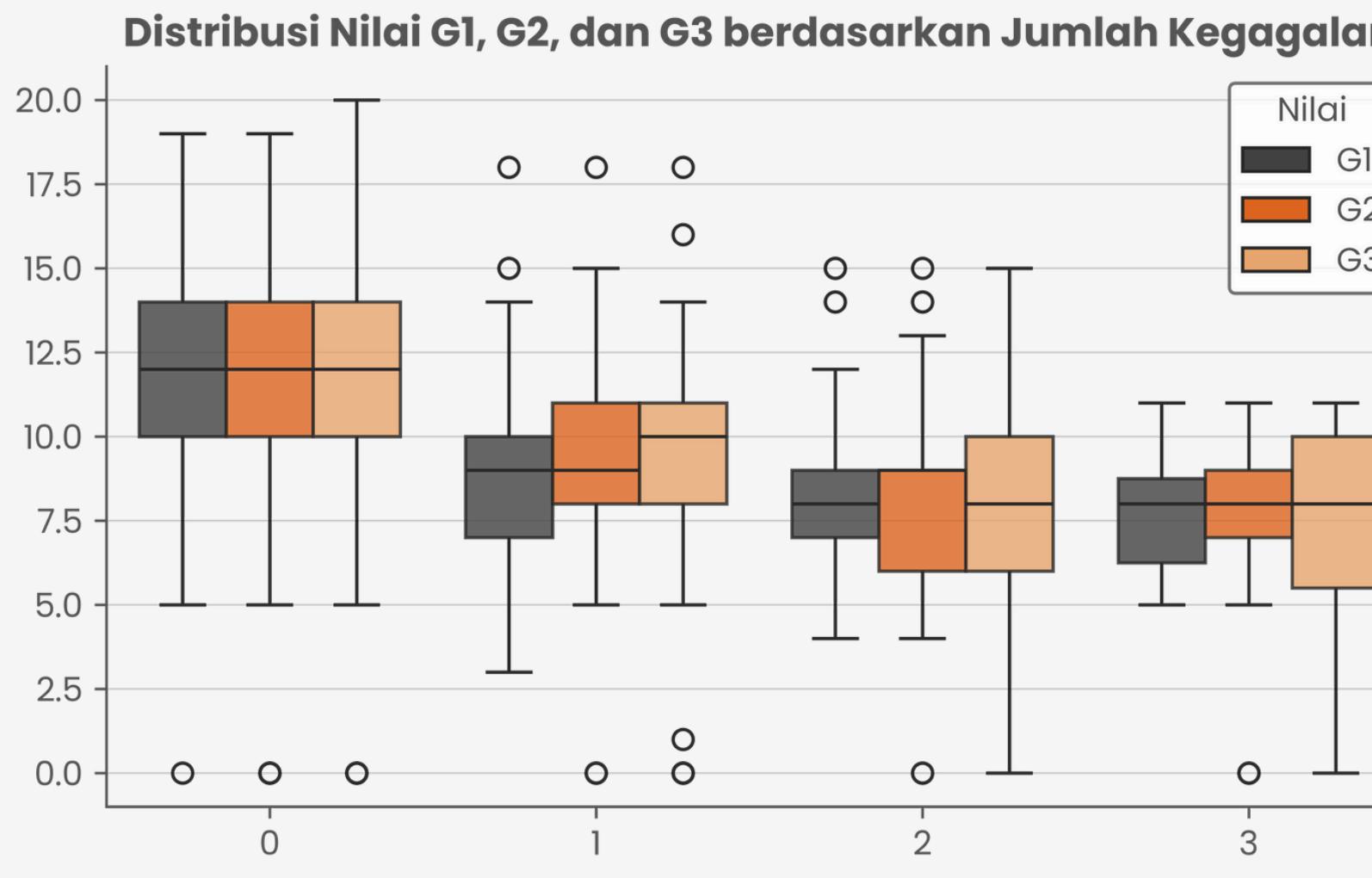
Meningkatkan motivasi lewat tantangan dan kompetisi



Studi Kasus Relevan

Menyisipkan contoh yang terhubung dengan minat siswa

Pendampingan Siswa Berisiko



Siswa dengan **riwayat tinggal kelas** memiliki kecenderungan **nilai akhir yang lebih rendah**. Untuk itu, diperlukan pendampingan yang lebih intensif.



Remedial
Terstruktur



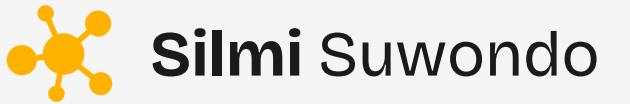
Mentoring
Sebaya



Pemantauan
Berkala



Konseling Belajar
Individual



Terima Kasih



github.com/afsilmis/

afsilmis@gmail.com

